Know Your Market: Use of Online Query Tools to Quantify Trends in Patient Information-seeking Behavior for Varicose Vein Treatment

Asheesh K. Harsha, MD, J. Eric Schmitt, MD, PhD, and
S. William Stavropoulos, MD

ABSTRACT

Purpose: To analyze Internet search data to characterize the temporal and geographic interest of Internet users in the United States in varicose vein treatment.

Materials and Methods: From January 1, 2004, to September 1, 2012, the Google Trends tool was used to analyze query data for "varicose vein treatment" to identify individuals seeking treatment information for varicose veins. The term "varicose vein treatment" returned a search volume index (SVI), representing the search frequency relative to the total search volume during a specific time interval and region. Linear regression analysis and Kruskal-Wallis one-way analysis of variance were employed to characterize search results.

Results: Search traffic for varicose vein treatment increased by 520% over the 104-month study period. There was an annual mean increase of 28% (range, −18%–100%; standard deviation [SD], 35%), with a statistically significant linear increase in average yearly SVI over time ($R^2 = 0.94$, $P < .0001$). All years showed positive growth in mean SVI except for 2008 (18% decrease). There were statistically significant differences in SVI by month (Kruskal-Wallis, $P < .0001$) with significantly higher mean SVI compared with other months in May (190% increase; range, 26%–670%; SD, 15%) and June (209% increase; range, 35%–700%; SD, 20%). The southern United States showed significantly higher search traffic than all other regions (Tukey-Kramer, $P < .00001$).

Conclusions: There have been significant increases in Internet search traffic related to varicose vein treatment in the past 8 years. Reflected in this trend is an annual peak in search traffic in the late spring months with an overall geographic bias toward southern states. Rigorous analysis of Internet search queries for medical procedures may prove useful to guide the efficient use of limited resources and marketing dollars.

ABBREVIATIONS

ANOVA = analysis of variance, GIS = Google Insights for Search, SD = standard deviation, SVI = search volume index

The Internet has become an important primary source of medical information (1,2) for patients; > 184 million

U.S. residents are regular Internet users, and a large percentage (70%–80% of U.S. adults) use the Internet to research information on their health conditions (1,4,5). This process of health information seeking online is not only a form of self-reporting medical illness but also has been shown to represent an indirect measure of disease prevalence, as has been shown for kidney stone incidence (6–8) and influenza outbreaks (9,10). In the era of direct-to-consumer Internet advertising for services including elective medical treatments, patients are increasingly being exposed to a wide variety of treatment options and treatment providers while investigating their medical condition. Patients who use the Internet for health information are 60% more likely to contact a health provider than patients who do not use the Internet (4).

Treatment of symptomatic varicose veins is an area of elective interventional practice in which patients elect to proceed with treatment and choose their provider.

Varicose veins affect 23% of adult Americans (22 million women and 11 million men). Since their emergence as efficacious and cost-effective therapies (11–14), endovenous laser therapy and radiofrequency ablation have increasingly become a component of practice of numerous physicians; endovenous procedures performed increased from 80,000 to 195,000 during the period 2004–2007 (11,15). Patients' desire for less invasive treatments, improved reimbursements to practitioners, and lack of requirement for surgical skills have resulted in a larger pool of physicians offering endovenous therapy (11,16). In the era of declining reimbursements (17), increasing provider competition for patients, and limited financial and personnel resources (17,18), developing and maintaining a successful venous practice requires the ability to plan for changes in market trends and allocate resources strategically based on annual and seasonal changes in volume (19).

Although traditional methods of market analysis may have a primary role in market research and practice management, Internet search data are a powerful repository of information that has the potential to reveal overall population interest and identify market trends in near real time with minimal expense and time to the investigator. The leading Internet search provider (20), Google (Google, Inc, Mountain View, California), has developed a publicly available Web-based tool, Google Insights for Search (GIS), which provides normalized Web traffic for user-defined search terms in a numerical and graphic format. Several studies have established the epidemiologic value of GIS for disease (6,7,10,21–23); we sought to examine the use of Internet search data to identify and characterize the temporal and geographic population interest in varicose vein treatment.

MATERIALS AND METHODS

GIS (http://google.com/insights/search) is a publicly available online tool that provides aggregated search volume for search terms that receive a significant amount of traffic for user-designated geographic regions and time restrictions. GIS also provides regional variations in search activity by country, state, or metropolitan area as well as a forecast function that is purported to predict future trends based on analysis of previous data. GIS search data are available only beginning January 1, 2004. In the present study, GIS was queried from January 1, 2004, to September 1, 2012, for the United States using lay terminology related to varicose vein treatment in English. With the objective of identifying potential patients interested in varicose vein treatment, the initial phrase employed was "varicose vein treatment." Additional search terms that GIS identified as related

terms that received significant search volume included "varicose veins treatment" and "vein removal." All three terms were queried in combination.

This query returned search volume results in the form of a search volume index (SVI). SVI represents the relative search frequency of a particular search term normalized to the overall search traffic from the reference population and time period. This analysis indicates the likelihood of a random user to search for a particular search term from a certain location at a certain time (24). GIS enforces a threshold of traffic for search terms, and terms with low search volume are not displayed. GIS also eliminates repeated queries from a single Internet Protocol address to eliminate the possibility that results are artificially impacted by multiple queries by a single user. Results are displayed on a scale of 0–100 with 100 representing the maximum relative search volume during the queried time period. Data points consisted of repeated measurements of search activity with a value provided for each week in the year over the 8-year study period. Each week of the study was grouped into its corresponding month, and the SVI was averaged to create an average SVI for each month. For example, the average June SVI represented the SVI for all weeks occurring in June from January 2004 to September 2012.

In addition, GIS provides regional search volume for each state using Internet Protocol addresses from server logs to establish the origin of web queries. A single SVI value was provided from data aggregated from each state over all 8 years. To analyze geographic differences in search volume, states were grouped into the following regions: South, West, Midwest, and East according to U.S. Census Bureau regional identification (25).

The data were imported into the R statistical programming environment for analysis (The R Project for Statistical Computing, 2012). Linear regression analysis and Kruskal-Wallis one-way analysis of variance (ANOVA) were employed to characterize annual search trends and peak search activity. One-way ANOVA also was employed to characterize differences in SVI from different geographic regions. Specific regional group differences were identified via post hoc Tukey-Kramer analysis. An α of .05 was considered statistically significant.

RESULTS

Normalized search traffic for varicose vein treatment and related key words increased by 520% over the 104-month study period with SVI increasing from an average of 10 in 2004 to 62 in 2012. There was a year-over-year mean increase of 28% (range, −18%–100%; standard deviation [SD], 35%), with a statistically significant linear increase in average yearly SVI over time ($R^2 = .94$, $P < .0001$) (Fig 1). All years showed positive growth in mean SVI except for 2008, which showed an 18% decrease (Fig 1).