Seasonality in Seeking Mental Health Information on Google

John W. Ayers, PhD, MA, Benjamin M. Althouse, ScM, Jon-Patrick Allem, MA, J. Niels Rosenquist, MD, PhD, Daniel E. Ford, MD, MPH

Background: Population mental health surveillance is an important challenge limited by resource constraints, long time lags in data collection, and stigma. One promising approach to bridge similar gaps elsewhere has been the use of passively generated digital data.

Purpose: This article assesses the viability of aggregate Internet search queries for real-time monitoring of several mental health problems, specifically in regard to seasonal patterns of seeking out mental health information.

Methods: All Google mental health queries were monitored in the U.S. and Australia from 2006 to 2010. Additionally, queries were subdivided among those including the terms ADHD (attention deficit-hyperactivity disorder); anxiety; bipolar; depression; anorexia or bulimia (eating disorders); OCD (obsessive-compulsive disorder); schizophrenia; and suicide. A wavelet phase analysis was used to isolate seasonal components in the trends, and based on this model, the mean search volume in winter was compared with that in summer, as performed in 2012.

Results: All mental health queries followed seasonal patterns with winter peaks and summer troughs amounting to a 14% (95% CI = 11%, 16%) difference in volume for the U.S. and 11% (95% CI = 7%, 15%) for Australia. These patterns also were evident for all specific subcategories of illness or problem. For instance, seasonal differences ranged from 7% (95% CI = 5%, 10%) for anxiety (followed by OCD, bipolar, depression, suicide, ADHD, schizophrenia) to 37% (95% CI = 31%, 44%) for eating disorder queries in the U.S. Several nonclinical motivators for query seasonality (such as media trends or academic interest) were explored and rejected.

Conclusions: Information seeking on Google across all major mental illnesses and/or problems followed seasonal patterns similar to those found for seasonal affective disorder. These are the first data published on patterns of seasonality in information seeking encompassing all the major mental illnesses, notable also because they likely would have gone undetected using traditional surveillance.

Introduction

Mental illness has one of the largest worldwide disease burdens but receives the least amount of funding per disability-adjusted life-year. Part of the challenge is the limited means of monitoring population mental illness trends. Annual telephone surveys are the principle psychiatric sentinel, but given respondents’ reluctance, because of stigma, cost, and time constraints, mental health items are often limited or omitted. For example, the Behavioral Risk Factor Surveillance System included only major depressive disorder measures (Patient Health Questionnaire) in 2006, 2008, and 2010 at a collective cost of more than $21 million. In addition, time lags between data collection and data availability can last years, during which acute intervening trends can be missed. In place of a running time series, investigators ask respondents when their symptoms were the worst. These prompts, however, lack precision and potentially tap respondents’ cultural beliefs or recall bias.

Until now, the resources have not existed to monitor continuous population interest in mental health problems.
One potential method is monitoring the Internet, the world’s most relied-on health resource.9 Because of mental health’s complexity, stigma, and obstacles to care, patients are likely to investigate their problems online.10,11 For instance, 20% reported searching for mental health information in 2000.12 As a result, these query trends are sufficiently granular to judge mental health information seeking as it changes day to day without the barriers of stigma, high costs, time, or artificial prompts.

Monitoring Internet search queries could be especially valuable to understanding seasonality in seeking information on mental illness. Seasonal affective disorder (SAD), a common mood disorder in which depressive-like symptoms are exacerbated in winter’s low-light environment, is one of the most studied health phenomena.13 Yet, very few studies have assessed how seasonality may exacerbate other mental illnesses. In the only multi-outcome study,14 to date, bipolar disorder, depression, eating disorders, and schizophrenia were stable across the year among Dutch survey respondents.

These results, however, are inconsistent with studies suggesting that bipolar disorder among Barcelonans,15 depression among U.S. children,16 eating disorders among college students in Norway17 and outpatients in Japan,18 and first-episode cases of schizophrenia among Australian men19 were lower during the summer. These inconsistent findings highlight how limited surveillance (and resulting measurement error) can hamper research agendas. This exploratory study investigated seasonal patterns in mental health search queries to both highlight their utility and help clarify the study of seasonality.

Methods

Health-related queries model influenza incidence,20–25 as well as dengue,26 kidney stones,27 and methicillin-resistant Staphylococcus aureus.28 Search query surveillance outside acute disease29–31 and in mental health, however, is rare.32–34 No studies to date have monitored several specific mental illness–related queries or highlighted the implications of these data for mental health.

Query Selection

Mental health queries principally derive from primary (self-diagnosis or treatment) and secondary (those connected to an affected person) sources. Prior studies suggest that queries are attributable to information seeking around incidence, recurrence, or increased severity of mental illness.15 Data were downloaded from Google Trends (google.com/trends), the public database of Google queries. Initially, all mental health queries (identified using Google’s query category feature) were captured that included language suggestive of mental health matters. Separately, all queries occurring in the mental health category that included the terms adhd (attention deficit-hyperactivity disorder); anxiety; bipolar; depression; anorexia or bulimia (eating disorders); OCD (obsessive compulsive disorder); schizophrenia; and suicide were subcategorized into problem-specific categories. This approach, for example, would categorize OCD symptoms, OCD test, or medications for OCD as OCD-related, and the mean trend for these and all similar terms as OCD queries.

Query Volume

Queries were normalized (relative search volume [RSV]) to the period with the highest proportion of searches going to the focal terms (e.g., RSV = 100 is the period with the highest search proportion for queries within a category, and RSV = 50 is 50% of that highest search proportion). This approach corrects for seasonal differences or trending in all queries because of changing Internet use, since weekly RSV estimates are scaled to all queries as a proportion and then rescaled to changes in that proportion.35 Trends were downloaded for the entire U.S. and Australia.

Data Analysis

The continuous wavelet transform was used to identify the significance of and isolate the annual periodic component (seasonality) of each time series, taking advantage of the temporal resolution afforded by weekly query data. In contrast to regression or Fourier decomposition, the wavelet allows examination of the intensity and timing of periodic seasonality and isolation of specific seasonal components of the original time series over time. Wavelet methods are robust to trend and noise, two common time-series analysis problems.36–38 For a comprehensive analytic treatment of the wavelet transform, see Torrence et al.37, for practical examples, see Grenfell38 and Johansson.39

The annual seasonal component of each time series was isolated in order to examine the timing; amplitude (or “magnitude”); and significance of any seasonality. The phase angle (the angular position along the sinusoidal trajectory from −180° to 180°, trough to trough) of the isolated series was calculated in order to estimate and compare timing of peaks and valleys in seasonal trends across nations and years. A difference in phase of 180° between two series indicates that they are completely out of phase (e.g., peaks in the U.S. series correspond to troughs in the Australian series), and a difference of 0° indicates the series are in phase (e.g., peaks in the OCD series in 1 year correspond to peaks in other years). To quantify the magnitude of seasonal peaks/troughs, a comparison was made of the mean search volume in summer months (June, July, and August in the U.S.; December, January, and February in Australia) to the mean search volume in the winter months.40

Results

Mental health queries across all illnesses/problems had pronounced peaks and troughs in the U.S. and Australian time series (Figure 1A). The annual component of the wavelet transformation isolated this seasonality (Figure 1B), detecting that mental health queries in the U.S. were almost entirely (180°) out of phase with Australia, meaning queries followed similar trends but 6 months different in their timing (Figure 1C), and this difference persisted over several years (Figure 1D). However, framing these same data as daylight seasonal cycles suggested that query peaks and troughs were consistent in the winter and summer, respectively, and this pattern was nearly identical between nations. The magnitudes of these patterns
were also very similar between nations. For instance, all mental health queries in the U.S. and Australia were about 14% (95% CI = 11%, 16%) and 11% (95% CI = 7%, 15%) higher in winter than summer. Problem-specific queries showed similar seasonality between nations (Figure 2). Consistencies in the timing of changes between the U.S. and Australia were strongest for ADHD (161°) and at their weakest for eating disorders (106°). The intensity of these seasonal patterns was also similar between nations. For instance, bipolar queries were 16% (95% CI = 12%, 20%) versus 17% (95% CI = 10%, 24%) higher in winter than in summer for the U.S. and Australia, respectively. The absolute difference between nations was less than 3% for half of the outcomes, including ADHD, depression, and OCD.

Over years and within nations, ADHD, bipolar, schizophrenia, and eating disorder queries showed little difference in the timing of their seasonality. Mean phase differences across years were smallest for ADHD and schizophrenia queries in the U.S. at −7.6° (95% CI = −8.9°, −6.4°) and 9.0° (95% CI = 7.6°, 10.4°), corresponding to only a 7–10 day difference in the timing of changes in query trends from year to year. In Australia, bipolar and schizophrenia queries had the smallest mean phase differences at 0.6° (95% CI = −1.9°, 3.1°) and 8.1° (95% CI = 7.2°, 8.9°), or about 1 and 8 days difference in the timing of changes in query trends from year to year.

The magnitude of query seasonality ranked largest to smallest was for: eating disorders, schizophrenia, ADHD, suicide, depression, bipolar, OCD, and anxiety; it was largely consistent across nations. For instance, eating disorder queries (“bulimia” and “anorexia”) had the largest seasonal difference (winter/summer) of 37% (95% CI = 31%, 44%) in the U.S. and 42% (95% CI = 30%, 51%) in Australia. The intensity of seasonal differences was also very high for schizophrenia: 37% (95% CI = 33%, 41%) in the U.S. and 36% (95% CI = 31%, 42%) in Australia. The smallest differences were observed for anxiety at 7% (95% CI = 5%, 10%) and 15% (95% CI = 10%, 19%) in the U.S. and Australia, respectively.

It is important to rule out nonclinical seasonal explanations for seasonal peaks and troughs, particularly media and academic motivations. Some basic investigation and logic does so. News mentioning schizophrenia, for example, was sometimes greater in the summer than winter (e.g., 572 stories in the summer of 2009 compared to 510 in the winter of 2009–2010 using a count of Google News archives), and on average, differences were minimal across mental health. For the current study, it was assumed that search sessions are iterative and that during the course of an online investigation a searcher would include a query with a mental illness root term. However, when symptomatology queries were investigated...
(e.g., hallucination queries compared to schizophrenia queries), the patterns were also seasonal and, at face value, the former appeared likely to be unrelated to academic interest.

**Discussion**

The current results suggest that monitoring queries can provide insight into national trends on seeking information regarding mental health, such as seasonality. Given their relatively anonymous nature, instantaneous availability, and the cost-effective manner by which the data are investigated, query trends have potential as an important adjunct to traditional surveillance.

**Implications for the Study of Seasonal Variation in (Mental) Health**

There has been limited theoretic development regarding seasonality in mental health because the data to inform this thinking have been limited. The current analyses indicate a consistent seasonality in patterns of mental health information-seeking, which may be useful in ongoing research into the biologic, environmental and social pathways influencing mental health, as reflected in these trends.

In the realm of biology, the prevailing theory regarding SAD is based on the retinal–suprachiasmatic–pineal–melatonin axis. Environmental factors, such as changes in daylight hours, influence the circadian system and invigorate physiologic irregularities in the body that affect mood. It is possible that similar mechanisms affect the trends observed in the current study on overall and illness-specific mental health information seeking. Alternatively, vitamin D deficiencies are known to play a role in affective disorders and mediation of brain reactions, with more-intensive sunlight promoting vitamin D absorption. A recent link between vitamin D deficiencies and the severity of pre-existing mental health problems is interesting to consider in light of the seasonal trends described here.

In the environmental realm, omega 3 consumption is higher in the summer and lowest in the winter. Omega 3 deficiency is associated with depression, bipolar disorder and schizophrenia, which may be related to the peaks and troughs observed in the current study. In the social realm, longer summer days create opportunities for social engagement, a well-known health emollient, and thus may contribute to the seasonality in mental health inquiries. Summer also allows for outdoor exercise, and increased physical activity has been associated with...
improved mental health\textsuperscript{52}; so this may contribute also to the seasonality of mental health–related Internet searches. The current results cannot tease out the etiology of search query behavior; they do, however, provide information that may be useful in further hypothesis testing and mechanist formulations.

Implications for Mental Health Treatment
A major challenge in mental health is how to not only assess but also treat mental illness among individuals who do not present for treatment or cannot be reached with telephone surveys.\textsuperscript{53} The Internet is a stigma- and cost-reducing venue to help screen and treat those who search for but may not bring problems to the attention of their clinicians. Internet-based treatment programs show promise\textsuperscript{54}; however, many search engine results are of questionable quality.\textsuperscript{55} Advertisements on search engines to evidence-based programs may link searchers to the best websites. This approach may be especially important for early detection and preventing more severe or opportunistic problems.\textsuperscript{56} Monitoring queries also may make terrestrial care more responsive. Managers can deliver more resources and provide additional screening in the winter based on the current results, as well as monitor trends for other acute needs. This type of data–practice integration is consistent with recent calls for mental health prevention and treatment.\textsuperscript{57,58}

Limitations
Although the current results contain some compelling findings, queries are not a replacement for conventional surveillance. Because queries are analyzed at the population level, they cannot capture demographic profiles like traditional sentinels. There is also a unique validation challenge, as sufficiently granular (daily or weekly) criteria are not available for comparison of queries for some rare illnesses. Thus, seeking information on Google may not correspond to actual mental illness. Yet, nearly all age-by-demographic population categories consume some online health information,\textsuperscript{59,60} suggesting that query trends may be indicative of population changes in information seeking. Still, changes among resource-poor populations greatly affected by mental illness are not captured, because these groups may not have Internet access.\textsuperscript{61} Query archives presently lack the geographic resolution to make inferences about information seeking in less populated areas.

Directions for Future Research
Monitoring queries is an instantaneous, localizable, and cost-effective method of collecting large amounts of data that can be reasonably assumed to correspond with population mental health information seeking. This low-cost and naturalistic method addresses two primary challenges facing mental health. It can produce information for researchers studying the biologic, environmental, and social pathways influencing mental illness. Further, it has potential usefulness for public health and other officials attempting to understand broad patterns of mental illness. If additional studies can validate the current approach by linking clinical symptoms with patterns of search queries (beyond general information seeking), this method may prove essential in promoting population mental health.

JWA received funding for this project from Google.org. BMA was supported also by a Graduate Research Fellowship Program from the National Science Foundation (grant no. DGE-0707427). The funders had no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the paper. The authors thank Carl Latkin and Larry Cheskin for their advice.

No other financial disclosures were reported by the authors of this paper.

References


27. Willard SD, Nguyen MM. Internet search trends analysis tools can provide real-time data on kidney stone disease in the U.S. Urology 2010;76:44–8.


