Information-seeking behaviour for epilepsy: an infodemiological study of searches for Wikipedia articles

Francesco Brigo¹,²,a, Willem M. Otte³,⁴,a, Stanley C. Igwe⁵, Harald Ausserrer², Raffaele Nardone²,⁶, Frediano Tezzon², Eugen Trinka⁶,⁷,⁸

¹ Department of Neurological, Biomedical, and Movement Sciences, University of Verona, Italy
² Department of Neurology, Franz Tappeiner Hospital, Merano, Italy
³ Department of Pediatric Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, The Netherlands
⁴ Biomedical MR Imaging and Spectroscopy Group, Image Sciences Institute, University Medical Center Utrecht, The Netherlands
⁵ Department of Neuro-Psychiatry, Federal Teaching Hospital, Abakaliki-Ebonyi State, Nigeria
⁶ Department of Neurology, Christian Doppler Klinik, Paracelsus Medical University, Salzburg, Austria
⁷ Center for Cognitive Neuroscience, Salzburg, Austria
⁸ Department of Public Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i.T., Austria

aAuthors contributed equally.

Received April 24, 2015; Accepted July 16, 2015

ABSTRACT – Millions of people worldwide use the internet daily as a source of health information. Wikipedia is a popular free online encyclopaedia used by patients and physicians to search for health-related information. Our aim was to evaluate information-seeking behaviour of English-speaking internet users searching Wikipedia for articles related to epilepsy and epileptic seizures. Using Wiki Trends, which provides quantitative information on daily viewing of articles, data on global search queries for Wikipedia articles related to epilepsy and seizures were analysed. The daily Wikipedia article views on syncope, psychogenic non-epileptic seizures, migraine, and multiple sclerosis served as comparative data. The period of analysis covered was from January 2008 to December 2014. Overall, the Wikipedia article “epilepsy and driving” was found to be more frequently visited than the articles “epilepsy and employment” or “epilepsy in children”. Since January 2008, the Wikipedia article “multiple sclerosis” was more often visited compared to the articles “epilepsy”, “syncope”, “psychogenic non-epileptic seizures” or “migraine”; the article “epilepsy” ranked 3,779 and was less frequently visited than “multiple sclerosis”, ranked at 571, in traffic on Wikipedia. The highest peak in search volume for the article “epilepsy” coincided with the news of a celebrity having seizures. Fears and worries about epileptic seizures, their impact on driving and employment, and news about celebrities with epilepsy might be major determinants in searching Wikipedia for information.

Key words: epilepsy, infodemiology, internet, web, Wikipedia
Millions of people use the internet daily as a source of health information. The increasing number of online searches, which are stored in query databases and conducted using popular web search engines, such as Google, generates so-called “big data”. These data provide valuable information about online health-related behaviour. The search dynamics, in particular, may be very useful as a real-time surveillance method to complement surveillances based on traditional data-gathering techniques (Ginsberg et al., 2009). Big data analysis has resulted in a new research discipline, termed “infodemiology”, which is defined as the study of the determinants and distribution of health information (Eysenbach, 2009).

The mother of big data systems, for search-related behaviour on the internet, is the freely accessible Google Trends service (Google Trends, 2014). This service has mainly been used in infectious disease monitoring (Carneiro and Mylonakis, 2009; Corley et al., 2010). However, other disciplines have started to catch up on this source of information and interesting examples have recently been provided in the field of neurological disorders (Bragazzi, 2013; Otte et al., 2013; Brigo et al., 2014a, 2014b, 2014c; Brigo and Ausserer, 2014; Brigo et al., 2015).

The use of big data analysis in this field is important and timely considering that neurological disorders are major determinants of the global burden of disease. There are multiple gaps in understanding the many issues related to brain disorders. Access to reliable big data through services such as Google Trends is useful to obtain more reliable information on this increasing global public health problem and to improve our understanding of related patient and public opinion. However, the Google Trends service is increasingly being criticized over the highly overestimated (up to 200%) flu prevalence (Butler, 2013; Lazer et al., 2014). The reason for this failure is not fully understood but has been partly attributed to the fact that Google does not make public the specific search terms it uses as raw data to analyse search queries. The particular algorithm the company uses to convert search frequencies into, for instance, “flu prevalence”, is also unknown and subject to change. The perceived flaws in the Google Trends service necessitate complementary services to provide a similar level of information, but with a higher degree of reliability and transparency.

A very good candidate to complement the Google Trends service in analysing online behaviour related to neurological disorders is Wikipedia. A first recent study has shown promising results, in terms of accuracy, based on the statistics of visiting Wikipedia’s page on the prediction of flu rates (McIver and Brownstein, 2014). Since its launch in 2001, the free online encyclopedia has become the most popular general reference site on the internet, and it is a popular source of health care information. However, Wikipedia web search behaviour for epilepsy has not yet been studied.

The aim of this study was to evaluate information-seeking behaviour of English-speaking internet users searching Wikipedia for articles related to epilepsy and epileptic seizures. In particular, we aimed to identify major determinants in searching Wikipedia for information, as a deeper understanding of these aspects may prove useful in improving and refocusing public enlightenment campaigns on epilepsy. We hypothesized that data on Wikipedia searches would reveal peaks corresponding to epilepsy-related news headlines. Access to reliable big data through services such as Wikipedia Trends would be useful to obtain more reliable information on epilepsy, which represents a global public health problem, and improve our understanding of related patient and public opinion.

Methods

Background

Wikipedia contains approximately 30 million articles, which are available in up to 287 languages. It includes over 4.5 million English articles. With 18 billion page views and nearly 500 million unique visitors a month, English Wikipedia ranks fifth place among the most visited websites globally (Wikipedia, 2014). Wikipedia offers a wealth of detailed information on an almost limitless range of topics. Consequently, it provides a very relevant source of public information on several aspects, including health. Furthermore, statistics and trends based on the amount of usage of particular articles make the Wikipedia environment an area of interest for researchers, for instance, to track which topics are “trending” in the public sphere (McIver and Brownstein, 2014).

Data collection

First, we selected the following Wikipedia articles related to epilepsy: “epilepsy and driving”, “epilepsy and employment”, “epilepsy in children”, “epileptic seizure”, “seizure types”, and “seizure threshold”. These terms were entered on the “Wikipedia Trends” webpage (available at: http://www.wikipediatrends.com/; accessed on 11 December 2014). Analysis was restricted to these articles, as no traffic statistic information on other articles related to epilepsy is available at http://www.wikipediatrends.com/. For this reason, the analyses did not include other terms such as “seizure”, “seizure disorder”, or “convulsion”.

Epileptic Disord, Vol. 17, No. 4, December 2015
As a comparator, we evaluated the number of visits to Wikipedia using the following terms: “syncope”, “psychogenic non-epileptic seizures”, “migraine”, and “multiple sclerosis”, which were also entered on the "Wikipedia Trends" webpage. “Multiple sclerosis” was chosen as it is a chronic condition with episodes of acute worsening (relapsing-remitting multiple sclerosis), hence similar to epilepsy (a chronic condition with acute, paroxysmal events [seizures]), but with a much lower prevalence (worldwide prevalence of multiple sclerosis ranges from 5 to 100 per 100,000 subjects [WHO, 2014]). Conversely, migraine was chosen as a comparator because it is a condition characterized by episodes of neurological involvement, hence similar to epilepsy, but with a much higher prevalence and incidence.

The information on article viewing is available using Wikipedia Trends, which provides traffic statistics of English Wikipedia articles from January 2008 (http://www.wikipediatrends.com/). Trends for several articles are presented in a single chart, and it is possible to analyse results using linear, logarithmic and normalized (by main page traffic) scales. Furthermore, we collected data (available at http://stats.grok.se) on the ranking of Wikipedia articles on “epilepsy”, “syncope”, “psychogenic non-epileptic seizures”, “migraine”, and “multiple sclerosis” among all those available in Wikipedia since January 2008. We entered these keywords on the http://stats.grok.se webpage. Next, we used Wikipedia Trends to identify the highest peaks in search volumes of the article “epilepsy” and then used the Google Trends service to identify possible relationships with news headlines published online. The keyword “epilepsy” was entered on the Google Trends main page (available at: http://www.google.com/trends; accessed 11 December 2014). Our search covered the time between 1 January 2004 and 11 December 2014. If no peak-related news headlines were found, we searched Google, entering the keyword “epilepsy” on the Google webpage (available at: “http://www.google.com”; accessed 11 December 2014) and using a custom date range and news filters to find news headlines related to the term “epilepsy”.

All searches in Wikipedia Trends were conducted on 11 December 2014 by one author (FB), in order to ensure reliability and replicability of the quantitative data retrieved by searching http://www.google.com/trends and http://stats.grok.se. Possible correlations between highest peaks in search volumes of the article “epilepsy” in Wikipedia and news headlines related to the term “epilepsy” in Google Trends or Google were independently and narratively evaluated by three authors (FB, SCI and WMO). As a comparison, we also evaluated possible correlations with news headlines before an increase in search volumes for the term “epilepsy” was observed (i.e. at the base of the search peaks); any disagreement was resolved through discussion.

Results

Overall, the Wikipedia article “epilepsy and driving” was more often visited than the articles “epilepsy in children” or “epilepsy and employment” (figure 1).

Figure 1. Wiki Trends graph depicting tendency over time of Wikipedia search queries for the articles “epilepsy and driving” (upper line), “epilepsy and employment” (arrow) and “epilepsy in children” from January 2008 to December 11th 2014. Results are expressed as logarithmic relative normalized search volume numbers.
Over time, search volumes for the Wikipedia article “epileptic seizure” were greater than those for “seizure types” and “seizure threshold” (figure 2). From January 2008, the Wikipedia article devoted to “multiple sclerosis” was more often visited than the articles “migraine”, “epilepsy”, “syncope”, or “psychogenic non-epileptic seizures” (figure 3).

The article on “epilepsy” ranked 3,779 in traffic on http://en.wikipedia.org/, and was therefore more frequently visited than the article “migraine” which ranked 3,944, but was less frequently visited than the article “multiple sclerosis”, which ranked 571. No information on ranking for Wikipedia articles on “syncope” or “psychogenic non-epileptic seizures” was available due to an insufficient volume of searches.

The highest peak in search volumes for the Wikipedia article on “epilepsy” (with 15,348 views) was observed on 29 March 2013 and coincided with the news headline published online: “(the rapper) Lil Wayne recovering in hospital after a seizure” (15 March 2013). In the weeks before an increase in search volumes for the term “epilepsy” was observed (i.e. at the base of the search peaks), no news headlines on celebrities with epilepsy were found. No news headline which
was likely to act as a driver for searches of the article “epilepsy” in Wikipedia were found other than those related to celebrity news events on epilepsy.

Discussion

This study is the first attempt to study information-seeking behaviour of English-speaking internet users searching Wikipedia for articles related to epilepsy and epileptic seizures.

Surprisingly, we found that the Wikipedia article on “multiple sclerosis” was much more frequently visited than the article on “epilepsy”, despite the fact that multiple sclerosis is a less common neurological disorder. The high search volumes of the Wikipedia article devoted to multiple sclerosis might be due to an increasing number of magnetic resonance imaging scans prescribed for a wide range of neurological symptoms with consequent increased incidental findings of white matter lesions, mostly ultimately unrelated to multiple sclerosis. However, these lesions may be a source of concern for physicians, patients and relatives, leading to increased internet searches, including Wikipedia visits.

Conversely, the Wikipedia article devoted to syncope was much less often visited than the article on epilepsy. This finding is in contrast to a previous report which showed that syncope-related terms have a greater number of hits on Google compared to epilepsy (Brigo et al., 2014b). This finding is in accordance with the high incidence of syncope in the general population (corresponding to a lifetime cumulative incidence of 30-40% [Ganzeboom et al., 2006]). However, the total number of Google searches related to syncope was found to only slightly exceed that related to epilepsy (Brigo et al., 2014b). This discrepancy might be due to the fact that the population using Google to obtain health-related information might be different from Wikipedia users. On the other hand, considering that seizures and epilepsy are much rarer events than syncope (estimated cumulative lifetime incidence of seizure is estimated at 4% in the general population [Haueter et al., 1993; Forsgren et al., 1996]), the number of Google and Wikipedia searches related to epileptic events is surprisingly high, and it suggests that internet searches may mirror patients’ fears and worries about seizures (and epilepsy), which are probably perceived as more serious events than syncope. A similar explanation may be offered for Wikipedia articles devoted to migraine (as for epilepsy above) in terms of search trends and ranking.

As expected, the Wikipedia article on “epileptic seizures” was more often visited than those focusing on more specific aspects of epileptic seizures, such as “seizure types” and “seizure threshold”. Conversely, the high volume searches for the Wikipedia article “epilepsy and driving” is in accordance with previous studies showing that lack of a driving license or limitations in driving are major determinants of health-related quality of life in patients with epilepsy, followed by concern for employment (Gilliam et al., 1997; Martin et al., 2005; Luoni et al., 2011).

In a previous infodemiological study, we evaluated changes in Google search behaviour occurring in English-speaking countries over time for terms related to epilepsy and epileptic seizures (Brigo et al., 2014a). Most terms associated with the search queries were related to symptoms of seizures, especially tonic-clonic seizures, and to seizures occurring in children, a population in whom seizures (including febrile ones) are an understandable cause of concern among parents and relatives. Similarly, in the present study, the Wikipedia article “epilepsy in children” was the most frequently visited epilepsy-related Wikipedia article after “epilepsy and driving”.

Furthermore, in both this infodemiological study of information-seeking behaviour of Google users (Brigo et al., 2014a) and in the present study, highest peaks in search queries were temporally related to news about famous people (especially in show business) suffering from epilepsy or epileptic seizures. No formal correlation between highest search peaks and celebrity news events related to epilepsy was possible, hence this relationship remains speculative and should be formally addressed in specific studies. However, apart from headlines related to celebrity news events on epilepsy, we failed to find other news headlines which were likely to have acted as a driver for searches for the article “epilepsy” in Wikipedia. Although not derived from a rigorous formal analysis, this finding is, however, consistent with previous observational studies on cancer (Chapman et al., 2005; Twine et al., 2006; MacArthur et al., 2011; Mcalfe et al., 2011; Juthe et al., 2015), multiple sclerosis (Brigo et al., 2014c), and Parkinson’s disease (Brigo, 2014), showing that a celebrity diagnosis can significantly influence public health behaviour, leading to increased public interest in disease diagnosis or prevention. Furthermore, this finding suggests that celebrities who publicly announce their epilepsy diagnosis might effectively promote epilepsy awareness programmes and increase public knowledge and reduce stigma related to such a diagnosis.

The present study appears to confirm a previous similar study on information-seeking behaviour of Wikipedia and Google users indicating that emotion towards epilepsy might play a significant role in the use of the internet as a source of health information, and that news about celebrities with epilepsy is probably a major determinant in online searching for epilepsy-related information.
These findings may be taken into account when planning public health measures aimed at improving the knowledge and attitudes of the general population towards epilepsy. In particular, informative campaigns with high emotional impact or with famous people acting as “testimonials” might be considered to effectively promote epilepsy awareness programmes and increase public knowledge related to this disorder. The relevance of infodemiological studies evaluating the online seeking behaviour for information on epilepsy and epileptic seizures is supported by the fact that nowadays an increasing number of patients have ready access to online information about health and disease (information on epilepsy is just “only a mouse click away” [Donner and Buchhalter, 2014]), and if health care providers do not provide the required information, patients may seek it online. In one survey conducted in men affected by epilepsy, the internet was the third source of information on epilepsy after the general physician and the neurologist (Sare et al., 2007), whereas other surveys found that 54 to 77% of patients with epilepsy seek online information on how to self-manage their epilepsy (Escofery et al., 2008). Epileptologists should be aware of the specific information needs of patients with epilepsy and work towards providing them.

This study has some limitations. First, our search was conducted using English terms, thus the findings are likely only to reflect online behaviour of English-speaking users. We had originally intended to evaluate search trends for Wikipedia articles in other languages, but we were unable to do this since traffic statistics are only available for English Wikipedia articles. Second, Wikipedia is probably less used by lay people compared to the Google website. Therefore, the available data are likely affected by non-representative sampling bias. Furthermore, Wikipedia Trends does not provide information on changes in search volume by country. A previous infodemiological study assessing the online Google seeking behaviour for the term “epilepsy” showed that the greatest search volume occurred in developing countries, where there is reduced public knowledge of epilepsy (Brigo et al., 2014a), however, it is unknown whether this regional seeking pattern also occurs for Wikipedia searches related to “epilepsy”. Another limitation is that our study analysed Wikipedia search volumes only. Although most searches are conducted through the Google search engine (NetMarketShare, 2014), moreover, Google search behaviour for the term “epilepsy” has previously been studied (Brigo et al., 2014a), it is unclear what types of information are provided by other search engines or searched by people using other search engines. Furthermore, we have no definitive information on who actually uses Wikipedia, other informative webpages on epilepsy, or search engines (patients, relatives, or physicians), or what their reasons are for searching. It is also possible that Wikipedia users used search key words slightly differently from those analysed, combined more terms, or shortened their search terms. Similarly, it is possible that availability of information on other sites could influence the number of searches in Wikipedia.

All these aspects may greatly influence the online seeking behaviour for information on “epilepsy” and may hamper the possibility of formally analysing the temporal patterns to obtain correlation coefficients or predicting fluctuations in search trends related to this condition (in contrast to the use of big data derived from the analysis of Google and Wikipedia Trends for real-time surveillance of disease outbreaks) (Carneiro and Mylonakis, 2009; Martin et al., 2014; McIver and Brownstein, 2014; Google Flu Trends, 2015).

At least at present, and to our knowledge, there is no way to overcome these shortcomings; online seeking behaviour for certain terms can only be partially explored and this is subject to limitations. Finally, speculation regarding online search behaviours inevitably carries the significant risk of ecological inference fallacy, i.e. deducing inferences about the nature of single individuals from inference for the group to which those individuals belong (Schwartz, 1994; Wakefield and Shaddick, 2006). Anyone performing or analysing data from this kind of study should be aware of this risk, intrinsically inherent to infodemiology. Infodemiological studies remain, however, the only way to shed light on the web search behaviour of millions of people worldwide.

Despite these limitations, our study indicates that fears and worries about epileptic seizures, their impact on driving and employment, and news about celebrities with epilepsy might be major determinants in searching Wikipedia for information.

Disclosures.
None of the authors have any conflict of interest to disclose.

References
Bragazzi NL. Infodemiology and infoveillance of multiple sclerosis in Italy. Mult Scler Int 2013: 924029.


Wikipedia. http://en.wikipedia.org/wiki/%C4%B0K%C4%BDER%0D%C4%B0A (accessed 28/10/14).