



Cancer information disparities on the internet: An infodemiological study



Francesco Brigo^{a,b,*,1}, Stanley C. Igwe^{c,1}, Raffaele Nardone^{b,d}, Andrea Orioli^b, Willem M. Otte^{e,f}

^a Department of Neurological, Biomedical, and Movement Sciences, Section of Clinical Neurology, University of Verona, Italy

^b Department of Neurology, Franz Tappeiner Hospital, Merano, Italy

^c Department of Neuro-Psychiatry, Federal Teaching Hospital, Abakaliki-Ebonyi State, Nigeria

^d Department of Neurology, Christian Doppler Clinic, Paracelsus Medical University, Salzburg, Austria

^e Department of Pediatric Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, The Netherlands

^f Biomedical MR Imaging and Spectroscopy Group, Image Sciences Institute, University Medical Center Utrecht, The Netherlands

ARTICLE INFO

Article history:

Received 6 February 2016

Received in revised form 7 March 2016

Accepted 24 April 2016

Available online 27 April 2016

Keywords:

Cancer
Google
Incidence
Infodemiology
Mortality

ABSTRACT

Aim: Most patients with cancers and cancer survivors use the internet to obtain health information and support each other. Our aim was to evaluate whether relationships exist between the information prevalences and search volumes of terms related to various cancers and their actual incidence and mortality figures in the USA and the UK.

Methods: The information prevalences and search volumes of English terms related to various cancers were obtained using the Google search engine and Google Adwords. These data were plotted against actual cancer incidence and mortality data obtained using UK and USA cancer databases.

Results: Breast cancer and melanoma were outliers, with greater levels of information available than expected from their incidence and mortality alone. Conversely, there were disparities between the information prevalences and actual incidences of colorectal and prostate cancers, both in the USA and the UK, indicative of lower levels of information availability online for these cancers.

Conclusion: Disparities in health care-related information exist for some tumor types with similar incidence and mortality. Disparities in virtual health care information may also exist in reality. Infodemiological studies might be useful for planning public health measures to increase the knowledge and attitudes of the general population towards treatable or preventable diseases.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Millions of people use the internet daily as a source of health information. The increasing number of online searches conducted using popular web search engines such as Google generates so-called ‘big data’, which can be analyzed to provide information about online health-related behavior or as a real-time surveillance method to complement traditional data-gathering techniques [1]. This has resulted in a new research discipline, termed “infodemiology”, which is defined as the study of the determinants and distribution of health information [2–4]. Google is the most popu-

lar search engine and is used in seventy percent of all searches [5], and it is likely to be used by both patients and physicians to look for online health-related information.

Cancer is a leading cause of global mortality and is a major source of anxiety for patients. It has been shown that 16–64% of patients with a cancer diagnosis use the internet to obtain health information or to complement their decision-making about their illness and treatment [6]. Wikipedia articles and cancer types are an important source of internet information to patients with cancer. They have a similar accuracy and depth as professionally edited cancer databases [7]. There is also an increasing tendency among cancer survivors to use the internet to obtain informational and emotional support [8].

Malignant cancers that have a high incidence or disease burden (i.e., are associated with mortality or a large impact on a patient’s quality of life), which are preventable, or for which screening tests exist might be expected to have more visibility in the media and

* Corresponding author at: Department of Neurological, Biomedical, and Movement Sciences, Section of Clinical Neurology, University of Verona. Piazzale L.A. Scuro, 10 - 37134 Verona, Italy.

E-mail address: dr.francescobrigo@gmail.com (F. Brigo).

¹ Authors contributed equally.

on the internet than benign tumors. Therefore, those tumors that have a paucity of publicly-available information need to be identified in order to better plan public health campaigns that increase knowledge about these cancers in the general population [2].

In this study, we aimed to obtain the information prevalences and volumes of searches for several terms related to different cancers, obtained by entering multiple search terms into the Google search engine. The correlations between information prevalences and search volumes for these terms and actual disease incidence and mortality figures in the USA and the UK were determined.

2. Methods

2.1. Information prevalence

Information prevalence is a basic but robust infodemiologic indicator that quantifies the absolute or relative occurrence of a certain keyword or concept in an electronic medium, specifically the internet [2]. The information prevalence is straightforwardly obtained by entering a search term, with or without included synonyms, into a search engine (in our case Google) to provide absolute numbers of occurrences of the search term over time. Depending on the search engine, an occurrence can either be the number of documents containing the search term at least once or the number of occurrences of the term in the entire database.

The information prevalences of terms related to several different cancers (prostate, stomach, colorectal, lung, breast, and cervix) were obtained by entering the search terms into the Google search engine (see Table 1). Search terms were entered in English for searches related to the USA. Searches were conducted on Sep 9, 2014.

2.2. Average monthly search volumes

The “average monthly search volumes” were also obtained using the Google Adwords website [available at: <https://adwords.google.com>, accessed on Sep 9, 2014]. Google Adwords is a commercial advertising website that reports data on search volumes (averaged per month) for specific search terms entered into Google, and filters for specific countries can be used to limit the data.

We did not include searches from Bing and Yahoo!, the second and third most frequently used English search engines, as search volumes data is only made available for Google searches.

2.3. Cancer incidence and mortality

The National Cancer Institute (NCI) database [available at <http://www.cancer.gov/statistics/find>] was used to obtain the latest actual (2011) cancer incidence and mortality data in the USA. Data on actual (2011) cancer incidence and mortality in the UK were obtained from the Cancer Research UK Webpage [available at: <http://www.cancerresearchuk.org/cancer-info/cancerstats/>].

2.4. Information seeking/prevalence index

To quantify disparities in virtual health care information we developed a straightforward measure, namely the ‘information seeking/prevalence index’. This index can be generated by dividing a measure of information seeking (i.e., the search volume averaged per month and expressed as thousands) by the information prevalence (expressed as thousands). This index is useful to evaluate whether disparities exist between health-related informative needs and overall information load for different diseases. This is a relative rather than an absolute measure; the highest index value would indicate that the needs of internet users for that specific dis-

ease is higher than the overall amount of information available on the internet for the other diseases.

2.5. Statistical analysis

Information prevalences and search volumes were linearly plotted against actual cancer incidence and mortality figures using the ggplot2 R software package [available at: <http://http://ggplot2.org/>]. To identify the presence of outliers in a visual analysis of scatterplots, the 95% confidence intervals were also determined. An outlier was defined as any value lying outside the 95% confidence intervals.

3. Results

Information prevalences, average monthly search volumes, and cancer incidence and cancer mortality figures are reported in Table 1. The highest information prevalence was for breast cancer (25,700,000 results retrieved) followed by lung cancer (7,190,000 results), melanoma (5,690,000 results), and ovarian cancer (3,760,000 results). In the USA, the highest average monthly search volumes were observed for ovarian cancer (3,760,000 searches), followed by leukemia (all subtypes combined; 165,000 searches) and breast cancer (165,000 searches). In the UK, the highest average monthly search volumes were those related to prostate cancer (60,500 searches) followed by pancreatic (40,500 searches) and breast (40,500 searches) cancers.

For both the USA and the UK data, breast cancer and melanoma information prevalences were upper outliers compared to other cancers in information prevalence versus disease incidence plots (Supplementary Figs. 1 and 4). Conversely, prostate and colorectal cancers were lower outliers compared to other cancers in information prevalence versus disease incidence plots (Supplementary Figs. 1 and 4).

Data on information prevalence for a particular disorder (expressed as thousands) divided by its actual incidence or mortality (expressed as 100,000 people per year) are reported in Table 2. Compared to other cancers with a similar disease burden or higher incidence or mortality, melanoma and breast cancer received high levels of attention, as measured by information prevalences and search volumes. Conversely, there was a disparity between the information prevalence and actual incidence of colorectal and prostate cancers in both the USA and in the UK data. Compared to their actual incidence, these cancers receive less attention on the internet than melanoma and breast cancer. For instance, prostate cancer in the USA has a higher incidence than breast cancer (147.8 versus 124.6 per 100,000 per year), but the overall information prevalence of English search terms related to prostate cancer (7,290,000) is approximately 3.5-times lower than that of breast cancer.

4. Discussion

The main finding of this study is that disparities in health care-related information exist for some tumor types with similar incidence and mortality.

The approach used in the present study has previously been shown to be useful for identifying disease areas in which there may be an information deficit, in order to inform policy and health epidemiological studies [2]. Diseases that have a high incidence or disease burden in terms of both mortality and impact on quality of life, which are preventable, or for which screening tests exist, should have more visibility in the media and on the internet than benign conditions [2]. For cancers with the above-mentioned characteristics, a strict correlation between information prevalence and disease incidence or mortality is desirable from a public health

Download English Version:

<https://daneshyari.com/en/article/3988710>

Download Persian Version:

<https://daneshyari.com/article/3988710>

[Daneshyari.com](https://daneshyari.com)