Proceedings of the 6th INFORMS Workshop on Data Mining and Health Informatics (DM-HI 2011)
P. Qian, Y. Zhou, C. Rudin, eds.

EXPLORING THE FREQUENCY OF QUALITY HEALTH INFORMATION IN ONLINE SEARCH RESULTS

Brent Kitchens
Department of Information Systems and Operations Management
University of Florida
Gainesville, FL
brentkitchens@ufl.edu

Christopher A Harle
Department of Health Services Research, Management and Policy
University of Florida
Gainesville, FL

Shengli Li
Department of Information Systems and Operations Management
University of Florida
Gainesville, FL

Abstract

Health consumers increasingly use online health information to inform their decisions and behavior. Because there may be serious consequences related to decisions based on inaccurate health-related information, it is important to understand the quality of health information that is retrieved by the public in their daily use of the Internet. This study empirically investigates the proportion of quality information that may be expected to be returned by a popular search engine (Google) for typical health-related searches. Results show that approximately half or more of the information returned is of a high quality, as judged by third parties that review and certify online health information. That said, first-page search engine results still include practically significant proportions of web sites whose quality may be questioned. Future research should seek to refine understanding of the amount and variation of online health information quality. Most importantly, this work should attempt to understand and analyze information quality in the context of realistic consumer information search and use behavior.

Keywords: health search quality, online health search, health information accuracy, health information online

Introduction

The Internet, including search engines, is increasingly used as a substitute and supplement for traditional sources of health care information and guidance for the public. More than 6 in 10 U.S. adults use the Internet to seek health information, and the majority of them report that information they found online influenced decisions related to their health [7]. In fact, more people are posing health questions to Google than to their doctors [1], and 46% of those who obtain health information online never discuss this information with a doctor [10].

On one hand, the availability of health information online may lead people to be more informed and make better decisions about their health. However, if Internet searches commonly lead to low quality information, there may be negative effects on people's health. For example, patients with a severe disease may decide to delay or forgo a visit to a physician due to information obtained online, and miss the best opportunity for treatment. Patients who find information online and overestimate their condition may feel pressure to undergo unnecessary medical tests or

procedures. For these reasons and others, it is important to understand whether the Internet allows users to find accurate and credible health information.

A review of 79 studies examining the quality of health information online determined that 70% of reviewed studies find issue with quality of online health information, despite differences in methods and rigor [4]. Several methods have been proposed for assisting consumers in assessing the quality of health information online; however there is no agreed-upon standard for such assessment [8, 14]. Consequently, many resources are available to aid consumers in finding accurate health information online.

One source, the not-for-profit Health On the Net Foundation (HON), has created the Health On the Net Code of Conduct for the provision of authoritative, trustworthy Web-based medical information, and certifies sites against this code of conduct based on request [9]. Certification criteria include the authority of included information, attribution to sources of information, justifiability of claims, and a clear distinction of advertisements and site content, among others. A search page is made available on the HON site that allows users to search for medical information, and only results from sites that have been HON certified are returned.

Another resource, MedlinePlus, is a part of the US National Institute of Health [13]. MedlinePlus also hosts a search page that allows users to search for medical information online and only returns results from trusted sources. However, MedlinePlus does not provide a certification upon request, as does HON. Instead, MedlinePlus attempts to choose and link to a limited number of sites that provide authoritative, but not redundant, information on medical topics. Like HON, there are specific criteria that must be met by a site for it to be linked to by MedlinePlus, including the quality, authority, and accuracy of the health content; differentiation between content and advertising; and reliability of the other sites to which the evaluated site has links.

Despite the existence of these and other sites that direct users to health information that has been determined to be of higher quality, it has been shown that only 27% of people who obtain health-related information online begin their search at a health-related website, as opposed to 65% who commence their search at a popular search engine [6] (such as Google). In the context of patients searching for health information, general search engines might not always place the most accurate or trustworthy information in the first several positions, but rather the most visited or referenced sites. It has been shown that people tend to limit their attention to the first page of (or even the first few) results when using search engines [3], and that the traffic to ranked results diminishes rapidly with decreasing rank [5]. If general search engines return low quality results on the top positions, the likelihood of patients being misled will increase.

In a previous study [11], we investigated the level of quality in results returned to users when searching a large range of health-related terms on Google. The purpose of this paper is to provide further insights to consumers, search engines, and policy makers by extending that work to consider the frequency with which consumers search various terms. To do this, we will assess the proportion of quality results returned by general search engines for a typical health-related search.

Data

Consistent with the methods of Laurent and Vickers [12], we created our data set using the list of 2,069 terms in the health topics index of MedlinePlus as of March 9, 2011. MedlinePlus is a resource provided by the National Library of Medicine directed at patients, not professionals in the medical field. Therefore, we expect that the terms and topics covered by MedlinePlus will be similar to terms that consumers will use when searching online for health information. The set includes terms divided into categories of body location/systems, disorders and conditions, diagnosis and therapy, demographic groups, and health and wellness. All terms were included in our analysis, regardless of similarity between them, as during exploratory searches it was found that Google returns different results or ranking of results even for very similar terms.

For each term we performed a Google search using a blank profile with no history or cookies stored, and location was set to USA to prevent inclusion of local results such as physicians' offices. Each search returned at least 10 results on the first page: a mixture of primary results, advertisements, and special results for images, videos, books, etc. We included only the primary results and excluded advertisements and special results. A maximum of 10 primary results were returned for each term, but for a small number of terms, the inclusion of special results reduced primary results to less than 10 on the first page. In total, we obtained 20,487 results over 2,069 terms.

No objective measure of quality is available for health information online, so to measure quality we used certifica-

tion or evaluation by the HON Foundation and MedlinePlus as a proxy. We used both sites in combination as it appears that, while there is some overlap, the two entities tend to certify or evaluate different types of sites. For each result, we manually extracted the domain to determine if that domain was certified or evaluated by HON or MedlinePlus as containing high-quality information. From the 20,487 unique results, we were left with 5,249 unique domains.

For each of the domains identified, we compared the domain to a list of those found to have a certificate issued by HON, counting each matching domain as HON certified, and therefore high-quality. As MedlinePlus only links to sites determined to meet their criteria for quality and does not issue official certificates to the sites, we used the MedlinePlus search feature to search for each domain to determine if such links exist. Domains for which no results were returned were determined not to be certified by MedlinePlus, while domains with direct links returned were labeled as certified.

However for MedlinePlus, there was a third category of domains – certain searches returned links to sites which in turn linked to the domain in question. While this may not be as strong of an indicator of quality as a direct link, one of MedlinePlus' stated criteria for evaluation is the reliability of the other sites to which the evaluated site has links. Therefore being referenced by a MedlinePlus certified site may also be an indication of quality. We labeled these domains as referenced and performed our analysis below both with and without these domains considered as high-quality. To summarize, domains (and therefore results) were categorized as certified or not certified by HON, and certified, referenced, or not certified/referenced by MedlinePlus.

For each of the 2,069 terms in our population, we calculated the proportion of high-quality results returned by considering only certified results to be high-quality. This calculation was repeated considering both certified and referenced results to be high-quality. Detailed results for this dataset using equal weighting of terms were reported previously [11]. In the current study, we extend these results by considering the frequency with which each term is searched by a typical user. Clearly certain terms are searched more frequently than others. For example, some terms are more well-known or represent more common health concerns and thus are more likely to be used in health consumers' searches.

To determine an expected search frequency for each term, we used the Traffic Estimator tool which Google provides to users of its AdWords service (https://adwords.google.com/o/Targeting/Explorer). The tool is meant to assist advertisers in identifying keywords to which they would like to associate an advertisement. It provides an estimate of the number of global monthly searches that the advertiser may expect to be made for an entered term. We determined the global monthly searches estimated by Google for each of our 2,069 terms and used these as weights for our analysis. For a small number of terms (134), the Traffic Estimator tool could not provide an estimate, and these terms were excluded from weighted analysis. Additionally, we excluded the term "sex" from our weighted calculations, as it was an extreme outlier with regard to search frequency, with global estimated searches almost 50 times that of the next most popular term.

Analysis

The quality proportion for each term was calculated using both certified only and certified/referenced measures of quality. These proportions were calculated using full page search results, as well as results at or above cutoff ranks 1 through 9 (for example, the proportion for cutoff 3 was calculated as the number of quality results in results ranked 1, 2, and 3, divided by 3). These proportions were the individual data points used in our analysis. We calculated the weighted average and 95% confidence limits for these proportions (table 1).

At all levels, the weighted results are lower than the previous un-weighted results, suggesting that more popular terms return lower proportions of quality results. However, at the page level, even using the certified only weighted measure, the proportion of quality results returned remains at approximately 50%. An interesting trend that may warrant further investigation is that quality proportion seems to increase much more sharply with rank when weighted measures are considered, even though at the page level, the level of quality is similar between weighted and un-weighted measures. This may suggest that terms which are searched more frequently have similar overall quality proportions but tend to return lower quality in the most highly ranked positions.

	Weighted [1]		Not Weighted [2]	
Cutoff	Certified & Referenced	Certified Only	Certified & Referenced	Certified Only
1	37.0 - 41.4	33.9 - 38.2	51.5 - 55.9	48.3 - 52.7
2	45.7 - 48.8	42.1 - 45.3	58.8 - 61.8	55.3 - 58.5
3	50.6 - 53.2	47.1 - 49.8	61.6 - 64.0	57.4 - 60.0
4	53.7 - 56.2	49.9 - 52.5	62.9 - 65.1	58.1 - 60.5
5	55.3 - 57.8	50.9 - 53.5	63.5 - 65.5	58.0 - 60.2
6	55.5 - 58.1	50.6 - 53.3	63.3 - 65.3	57.2 - 59.4
7	56.1 - 58.6	50.6 - 53.3	62.8 - 64.8	56.1 - 58.3
8	55.7 - 58.2	50.1 - 52.7	62.1 - 64.1	54.8 - 56.8
9	55.3 - 57.7	49.2 - 51.7	61.4 - 63.4	53.5 - 55.5
Page	54.2 - 56.6	47.9 - 50.4	60.4 - 62.2	52.1 - 54.1

Table 1: 95% confidence limits for mean quality proportion (%) at or above each possible rank

Discussion

In this study, we find that even when considering the disproportionate frequency with which users search for various health-related terms, approximately half or more of first page results returned by popular search engines are high quality. While this leaves room for improvement, severe concern that may exist among health care providers about the general inaccuracy of health-related information online may be unwarranted. However, it does seem that highly popular terms may produce lower quality results than less popular ones. This may be the result of search engine optimization and other techniques that lower quality sites may use to increase their search ranking for popular keywords. Alternatively, it is possible that more popular terms are those which are more familiar to the average user, and thus are more likely to be discussed in user-generated web content, which may not be considered high quality (either in general or by our proxies).

Regardless of the cause, this apparent discrepancy between more and less popular health search terms may be an area of interest to those who seek to provide consumers with quality health information online. This might include governmental or non-profit agencies, physicians, or even popular search engines themselves. Search engines usually seek to provide the most relevant results possible, but presumably when users search for health information, they are interested in obtaining results that are not only relevant, but accurate and trustworthy. Since health search appears to represent a significant amount of activity online, search engines competing for customers may be interested in increasing the quality (i.e. accuracy) of results returned for health-related terms.

Continued study of this topic is merited. Few prior works empirically investigate the quality of results in health related search on popular search engines, although this appears to be the primary way in which users obtain health information online. Our studies were limited in that we only considered results from Google for a single corpus of terms. We also restricted our focus to primary results by turning off local results and ignoring special results for books, scholarly articles, etc. Inclusion of these results in future studies would make analysis more difficult, but would also increase realism as these results are presented to the average user with little differentiation from primary non-local results. Future studies might also compare health-related queries over time to determine factors that influence rank of search results. For example, it is conceivable that certification or evaluation by HON or MedlinePlus may be related to improvements in search ranking for a site. If this were the case, sites may have more incentive to provide high quality information in order to receive certification and boost their search engine rankings.

It is also possible that our measurement of quality proportion using HON and MedlinePlus is an underestimate of

[1] Weighted proportions calculated without the term "sex," an extreme outlier in search frequency (weight), as well as 134 terms for which the Traffic Estimator tool provided no estimate.

[2] Un-weighted results reported previously in [11]

the actual quality obtained by consumers performing health-related search. There may be high quality sites in our result set which are not certified or evaluated by HON or MedlinePlus. Additionally, there is a distinction between quality and relevance which this paper has not investigated. As many terms in our population may be related to topics other than health, searches for these terms will return at least some results which may be considered as not health-relevant. These results may be distinguished from health-relevant results which are not considered low quality. However, we have not as of yet determined a method of systematically making this distinction and instead have considered these results as low quality. Consumers in a realistic search situation, however, will likely be able to make this distinction, or even alter their queries to eliminate non-relevant results. Consumers may even be able to distinguish high quality results from low quality – for example, it has been shown that medical universities and governmental sources of online information are trusted more than other sources such as insurance companies and community health organizations [2]. Therefore the proportion of quality information obtained by consumers online may be much higher than estimated above.

Conclusion

This study finds that approximately half or more of first-page search results may be expected to provide accurate and credible information in a typical consumer search for health-related information on Google. These results suggest general search engine use will often return quality health information. That said, because search engines also return a meaningful proportion of information that may be of lesser quality, researchers, policy makers and information providers should continue to explore the relationship between search engine results rankings and the quality of health information that consumers use regularly to understand many aspects of their health and healthcare.

References

[1] Cohen, E., 2010, "Your top health searches, asked and answered," CNN Health, retrieved on May 26, 2011 at http://www.cnn.com/2010/HEALTH/10/21/top.health.searches.answered/.

[2] Dutta-Bergman, M., 2003, "Trusted online sources of health information: differences in demographics, health beliefs, and health-information orientation," Journal of Medical Internet Research, 5(3), e21.

[3] Eysenbach, G., Kohler, C., 2002, "How Do Consumers Search For and Appraise Health Information on the World Wide Web? Qualitative Study Using Focus Groups, Usability Tests, and In-Depth Interviews," British Medical Journal, 324, 573-577.

[4] Eysenbach, G., Powell, J., Kuss, O., Sa, E.R., 2002, "Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web: A Systematic Review, the Journal of the American Medical Association," 287(20), 2691-2700.

[5] Feng, J., Bhargava, H.K., Pennock, D.M., 2007, "Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms," INFORMS Journal on Computing, 19(1), 137-148.

[6] Fox, S., 2006, "Online Health Search 2006," Pew Internet & American Life Project, retrieved on May 26, 2011 at http://www.pewinternet.org/~media//Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf.

[7] Fox, S., Jones, S., 2009, "The Social Life of Health Information," Pew Internet & American Life Project, retrieved on May 26, 2011 at http://www.pewinternet.org/~media//Files/Reports/2009/PIP_Health_2009.pdf.

[8] Gagliardi, A., Jadad, A.R., 2002, "Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination," British Medical Journal, 324, 569-573.

[9] "Health on the Net Foundation," retrieved on May 26, 2011 at http://www.hon.ch/.

[10] Humphrey, T., 2010, "'Cyberchondriacs' on the Rise?," Harris Interactive, The Harris Poll #95.

[11] Kitchens, B., Harle, C.A., Li, S., "Quality of Health-Related Online Search Results," Decision Support Systems, Forthcoming.

[12] Laurent, M.R., Vickers T.J., 2009, "Seeking Health Information Online: Does Wikipedia Matter?," Journal of the American Medical Informatics Association, 16(4), 471-479.

[13] ”MedlinePlus - Health Information from the National Library of Medicine,” retrieved on May 26, 2011 at http://www.nlm.nih.gov/medlineplus/.

[14] Wilson, P., 2002, ”How to Find the Good and Avoid the Bad or Ugly: a Short Guide to Tools for Rating Quality of Health Information on the Internet,” British Medical Journal, 324 598-602.