Journal of Digital Imaging

Utilization of a Radiology-Centric Search Engine

Richard E. Sharpe Jr,[1,2] Megan Sharpe,[1] Eliot Siegel,[2] and Khan Siddiqui[2]

Internet-based search engines have become a significant component of medical practice. Physicians increasingly rely on information available from search engines as a means to improve patient care, provide better education, and enhance research. Specialized search engines have emerged to more efficiently meet the needs of physicians. Details about the ways in which radiologists utilize search engines have not been documented. The authors categorized every 25th search query in a radiology-centric vertical search engine by radiologic subspecialty, imaging modality, geographic location of access, time of day, use of abbreviations, misspellings, and search language. Musculoskeletal and neurologic imagings were the most frequently searched subspecialties. The least frequently searched were breast imaging, pediatric imaging, and nuclear medicine. Magnetic resonance imaging and computed tomography were the most frequently searched modalities. A majority of searches were initiated in North America, but all continents were represented. Searches occurred 24 h/day in converted local times, with a majority occurring during the normal business day. Misspellings and abbreviations were common. Almost all searches were performed in English. Search engine utilization trends are likely to mirror trends in diagnostic imaging in the region from which searches originate. Internet searching appears to function as a real-time clinical decision-making tool, a research tool, and an educational resource. A more thorough understanding of search utilization patterns can be obtained by analyzing phrases as actually entered as well as the geographic location and time of origination. This knowledge may contribute to the development of more efficient and personalized search engines.

KEY WORDS: Imaging informatics, internet, teaching, bioinformatics, computers in medicine, education, medical, information resources

INTRODUCTION

The Internet is increasingly utilized in place of traditional textbooks to obtain medical information in a variety of disciplines.[1–6] Physicians access the Internet for a range of purposes, including resident education, communication of information to peers, research, and as a clinical decision-making tool.

Both the medical literature and anecdotal experience suggest that radiologists, in particular, are increasingly using the Internet and Internet-based search engines. As early as 1999, a survey of 210 European radiologists and radiology residents reported that 18% access the Internet daily for radiologic activities. A total of 69% of all respondents had used the Internet at some time for literature searching and 12% at some time for teaching or education. The study further found that 28% of respondents had accessed electronic journals.[4] A 2002 survey of 2,200 physicians found that 65% used the Internet for literature searching, 53% used it to search for medical information, and 45% accessed online journals.[5]

This use has increased significantly in the past decade. In a 2005 survey of 92 radiologists, 97% reported using the Internet for radiology education, and 69% found the reliability of the information to be equal to that of information from traditional sources.[6] The same survey found that 31% of respondents used the Google search engine most frequently to find radiologic information.

[1]*From the School of Medicine, Georgetown University, Washington, DC, USA.*

[2]*From the Department of Radiology, Baltimore VAMC/ University of Maryland, College Park, MD, USA.*

Correspondence to: Richard E. Sharpe Jr, 5545 Woodridge Drive, Waldorf, MD, 20601, USA; tel: +1-202-4257304; e-mail: richsharpejr@gmail.com

Online publication 12 February 2009

doi: 10.1007/s10278-009-9182-z

Physicians find the Internet to be a reliable source of medical information that can be directly applied to clinical problems. Reports have documented the widespread belief among physicians that medical content on the Internet has professional value and that better care can be provided using this information.[5] A recent study described the use of the Google search engine as a diagnostic tool to make the correct "final" diagnosis in 58% (15) of the 26 diagnostic cases published in the case records of the *New England Journal of Medicine* in 2005.[7] Search engines help physicians formulate differential diagnoses and consequently provide better clinical care, especially in difficult cases.

Google, Yahoo, and numerous other generic search engines can be used to provide radiologists with information. However, these generic search engines are designed to provide general results for all users rather than results tailored to the interests of a specific group of users. Use of these generic search engines requires radiologists to sift through a substantial volume of non-radiology material and irrelevant links in order to arrive at relevant information. Consequently, using generic search engines to obtain radiologic information may not provide results that are sufficiently focused and efficient for point-of-care clinical decision making, education, and research.

As one solution, radiology-specific vertical search engines have been developed. These search engines utilize various algorithms to narrow the scope of their results to links that are relevant to the radiologist. Such search engines include Yottalook and GoldMiner, among others.

Yottalook.com is a free search tool developed by radiologists to make Web searches more efficient for medical imaging professionals. This radiology-centric vertical search engine optimizes searches to yield results that are relevant to medical imaging. Yottalook "crawls" the Web to identify medical imaging–centric Web sites and extracts key concepts from each Web page it encounters. It then indexes the key concepts, extracts medical images, and identifies semantic relationships within a website. These relationships are then used by its search algorithm for image retrieval and are also applied to a Google custom search to provide results that are optimized for clinical and research purposes. When one initiates a search, Yottalook applies a natural language query analysis to analyze the search term to better "understand" the purpose of the user's search. It then applies a semantic ontology that incorporates RadLex to expand the user's query by adding synonyms, parent and child terms, and other related terminology.[8] Yottalook also applies its query expansion technology to its Google Books site to provide full-text search capability within medical imaging textbooks. Finally, it offers users "Yottalinks" that provide high-yield editorial content based on the specific search term entered.[9]

We describe searches performed using the website www.yottalook.com over a 3-week period, 7 months after its initial launch. Searches are analyzed to better understand utilization of this radiology centric vertical search engine.

METHODS

For every 25th consecutive Yottalook query from August 1, 2007 to August 24, 2007, the time of search, IP address, and other data were downloaded for retrospective analysis. Institutional review board (IRB) approval was not required according to our local IRB.

Two authors (RES, MJS) sorted the terms into predefined categories including 11 radiology subspecialties (musculoskeletal, pulmonary, gastrointestinal, genitourinary, neuroradiology, vascular and interventional, nuclear, ultrasonography, pediatrics, breast, and cardiac) and three other categories (general radiology, other medical conditions, and other queries). Radiology subspecialty allocation was determined through consensus using criteria defined by the American Board of Radiology in its Diagnostic Radiology Study Guide.[10] Reviewers utilized radiology texts and Medline searches of the primary medical literature to assist in categorization of searches. Search terms that met criteria for inclusion in more than one radiology subspecialty were classified into multiple subspecialties.

Radiology subspecialties were defined as follows: Musculoskeletal radiology included all imaging modalities related to normal physiology or pathology of the musculoskeletal system, including disorders of the bone marrow. Pulmonary radiology included diseases of the lungs, pleura, mediastinum, and all of its structures excluding the heart and great vessels. Gastrointestinal radiology included imaging of the esophagus, stomach, small and large intestines, biliary tract, liver, spleen, pancreas, peritoneal

cavity, and abdominal wall. Genitourinary radiology included searches of the kidneys, adrenal glands, ureter, bladder, urethra, and the male and female genital systems. Neuroradiology was defined as imaging related to the skull, sinuses, mastoids, spine, head and neck, thyroid, parathyroid, and pituitary glands. Vascular and interventional radiology included all imaging of the arteries, veins, and lymphatic structures. Nuclear imaging included all static and dynamic imaging using radiopharmaceuticals. Ultrasonography included use of the modality to investigate the head and neck, thorax, abdomen, pelvis, extremities, breast, scrotum, vascular system, uterus, and fetus. The pediatric category included all imaging of infants and children. Breast radiology included all modalities for diagnosing breast disease. Cardiac was defined as the evaluation of the heart and the great vessels.

The category of general radiology included searches that were radiologic in nature but did not fit into established radiology subspecialties. Terms that did not fit into a specific subspecialty and were medical rather than radiological in nature were classified as other medical topics. Terms not fitting into any of the above-described categories were classified as other queries. The inclusion of specific imaging modalities in search queries was also noted.

International WHOIS databases were used to determine the search origin of queries from their Internet protocol (IP) address. This location was also used to categorize queries into geographical regions using accepted classification systems. Geographic regions included Africa, Asia, Europe, North America, Oceania, and South America. Radiology subspecialty categorization and imaging modality categorization were analyzed by region. After the search origin location was identified, the time of the search was converted to that region's local standard time to analyze searches by time of day.

Search terms were analyzed using standard medical references to check for misspellings. The presence of abbreviations and the language of the search were also noted for each search term.

Data analysis was performed using Microsoft Excel and accepted methods for calculating 95% confidence intervals for proportions with nominal data. Statistical differences were considered significant at the 95% level.

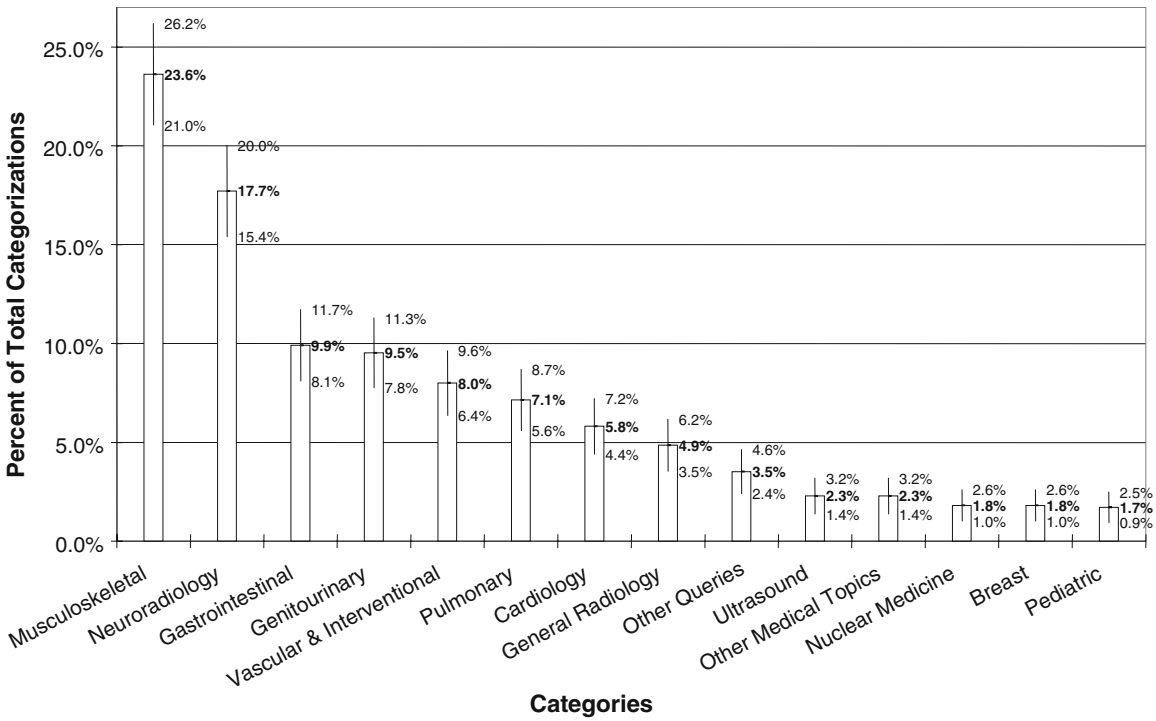

Fig 1. Percent of search queries sorted by predefined categories with 95% confidence intervals.

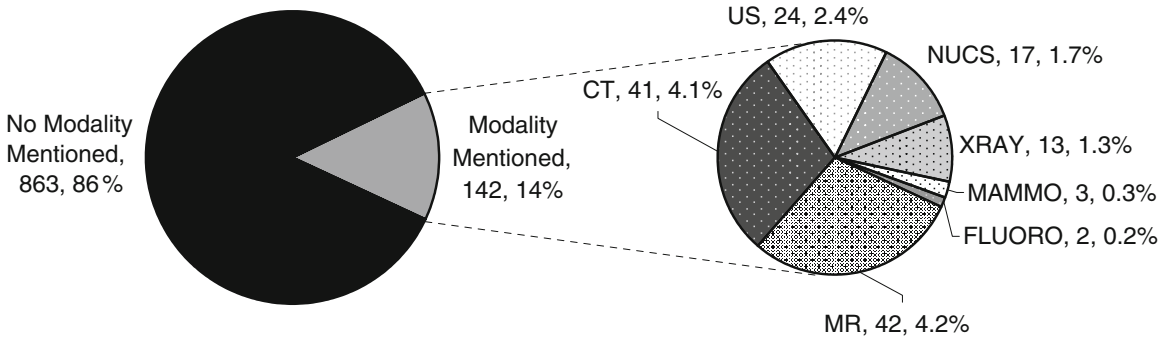

Fig 2. Search query analysis by reference to imaging modality presented as modality, number of times mentioned in search queries, and as percent of overall queries.

RESULTS

During the study period, 25,075 queries were placed at www.yottalook.com, and 1,005 terms were retrospectively categorized in this study. A total of 1,049 subspecialty categorizations were made from the 1,005 included searches. Nine hundred sixty (95.5%) searches were classified into a single radiology subspecialty, 44 (4.4%) terms were classified into two subspecialties, and one (0.10%) term was classified into three subspecialties.

Musculoskeletal imaging was the most frequently searched subspecialty. Neuroradiology was searched significantly less frequently than musculoskeletal imaging and searched significantly more frequently than the other remaining subspecialties. All subspecialties were searched (Fig. 1).

One hundred and forty-two (14.1%) searches specified an imaging modality. Magnetic resonance imaging (MRI) was indicated in 42 (4.2%) searches, and computed tomography was indicated in 41 (4.1%). All common radiological modalities were searched (Fig. 2).

Searches were placed from all continents (Table 1). The largest number (793, 78.9%) of searches originated from North America. Six queries (0.6%) had IP addresses that could not be traced to their origins using WHOIS software. Six hundred seventeen (61.4%) searches were executed between 9 A.M. and 5 P.M. at their local points of origin, whereas only 62 (6.2%) were executed between 12 A.M. and 7 A.M. (Fig. 3). English was the search language in 998 (99.3%) searches. The remaining searches were in Spanish. Abbrevia-

Table 1. Search-Term Categorization by Geographic Region

	Geographic Region											
	North America		South and Central America		Europe		Oceania		Asia		Africa	
Search Term Categories	Queries	Percent total	Queries	Percent total	Queries	Percent total	Queries	Percent total	Queries	Percent total	Queries	Percent total
Musculoskeletal	190	24.0	21	23.6	20	27.4	9	39.1	0	0.0	0	0.0
Pulmonary	55	6.9	6	6.7	4	5.5	2	8.7	1	6.3	2	40.0
Gastrointestinal	74	9.3	14	15.7	9	12.3	1	4.3	2	12.5	0	0.0
Genitourinary	72	9.1	8	9.0	7	9.6	3	13.0	2	12.5	0	0.0
Neuroradiology	148	18.7	12	13.5	12	16.4	5	21.7	4	25.0	0	0.0
Vascular and interventional	61	7.7	11	12.4	3	4.1	1	4.3	1	6.3	1	20.0
Nuclear	18	2.3	0	0.0	1	1.4	0	0.0	0	0.0	0	0.0
Ultrasonography	17	2.1	3	3.4	3	4.1	0	0.0	1	6.3	0	0.0
Pediatrics	15	1.9	1	1.1	0	0.0	0	0.0	1	6.3	0	0.0
Breast	13	1.6	3	3.4	0	0.0	0	0.0	1	6.3	0	0.0
Cardiac	48	6.1	3	3.4	1	1.4	2	8.7	1	6.3	1	20.0
General radiology	33	4.2	1	1.1	9	12.3	0	0.0	1	6.3	1	20.0
Other medical topic	20	2.5	2	2.2	2	2.7	0	0.0	0	0.0	0	0.0
Other query	29	3.7	4	4.5	2	2.7	0	0.0	1	6.3	0	0.0
Total searches by origin	793	78.9	89	8.9	73	7.3	23	2.3	16	1.6	5	0.49

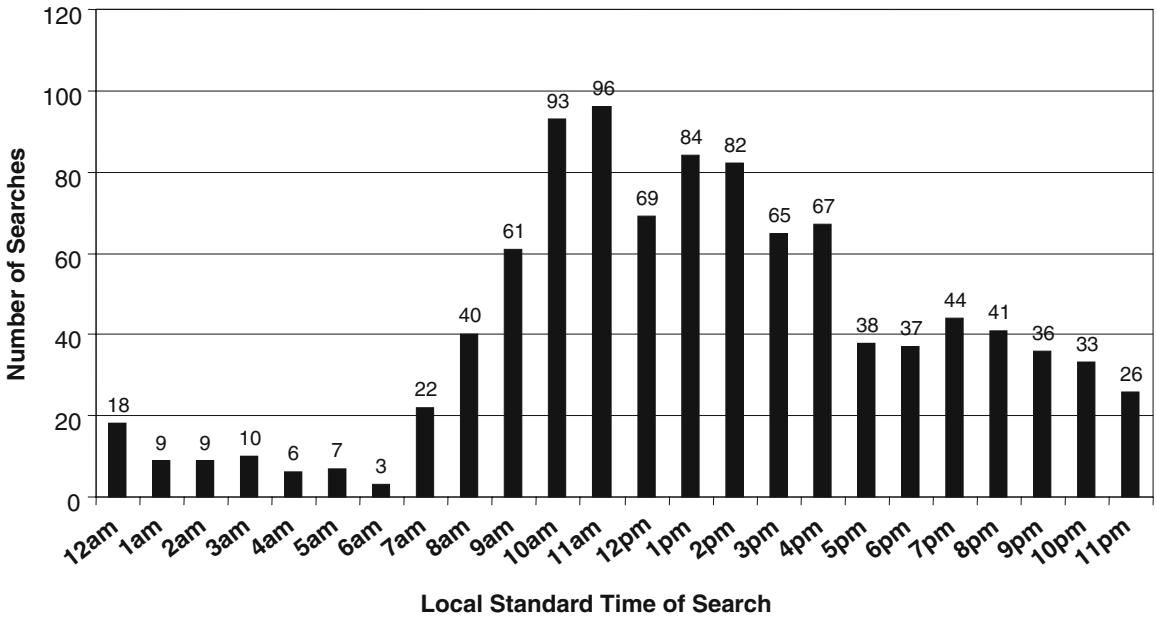

Fig 3. Search engine utilization by local standard time.

tions were used in 186 (18.5%) searches, and 93 (9.3%) searches contained misspellings.

DISCUSSION

More than half of all Internet searches on www.yottalook.com came from three of the radiologic subspecialties: musculoskeletal radiology, neuroradiology, and gastrointestinal radiology. The least frequently searched categories included breast imaging, nuclear imaging, and pediatric radiology, which when combined accounted for approximately 10% of all search queries.

Statistically significant variation was noted in the frequencies of search queries for various radiologic subspecialties. Our study documents the variability in the number of searches for various imaging subspecialties. Radiology-specific search engines can be used for many purposes, including clinical, research, and educational applications, and we would anticipate different results for each of these purposes.

It is not surprising that musculoskeletal imaging and neuroradiology were searched most frequently. These areas are generally considered to represent the "bread and butter" applications of MRI which was the most frequently referenced imaging modality.

The least frequently searched categories of nuclear medicine, pediatric, and breast imaging may represent relatively subspecialized areas that may not be as commonly searched by general radiologists. Additional analysis should be performed to better understand the variability in queries among the different radiology subspecialties.

One interesting observation was that few musculoskeletal searches originated in Asia. In addition, the African search pattern was substantially different from overall patterns. Africa and Europe also had a larger percentage of searches that specified ultrasound. This regional variation may result from variability in the availability of specific modalities and the preferences of practicing physicians and radiologists for various modalities, as well as differences in the prevalence of diseases in these regions. Our data suggest that it may be advantageous to include regional differences in the development of algorithms designed to optimize the effectiveness of these search engines.

Radiology-specific search engines are used globally to collect information to investigate all radiologic modalities, especially during the typical workday. The nature of searches occurring 24 h/day in adjusted local times may suggest that search engines are being used as real-time clinical decision-making tools, particularly for MR and computed tomography questions. Late-night and early morning searches may be originating from on-call attending radiologists, radiology residents, and teleradiologists.

Frequent natural language abbreviations and misspellings suggest that search engines should be prepared to interpret and appropriately modify common searches for increased user efficiency.

The limitations of our methodology include the inability to precisely assess the level of expertise and the professional identity of persons using the search engine. The majority of searches, however, were relatively homogenous. That is, most were performed during the regular business day and a few searches were categorized as "other queries." Therefore, we believe it is likely that most searches included in this study were placed by general radiologists, subspecialty radiologists, radiology residents, students, and other physicians or other radiology professionals. However, the relative proportion of these is unknown. Future studies could evaluate the relative proportions of these users.

We also have limited data from which to assess the practice and search patterns in Asia and Africa because of the smaller number of total searches from these regions.

Another limitation of the study is the fact that it focused on the use of the search engine soon after its initial release. A follow-up analysis would be useful, especially because utilization of the search engine has increased greatly over the past several months.

CONCLUSION

Radiologic searches using a radiology-specific search engine are more frequently undertaken within some radiology subspecialties than in others. Radiologic searches with a radiology-specific search engine vary by subspecialty and modality searched, geographic region of search origin and time of day at which the search was conducted. This variability may reflect the relative caseload of radiologists as well as dispersion and utilization trends of various imaging technologies.

This study demonstrates varying international utilization of specialized Internet search engines across all imaging modalities. We believe current and future Internet-based search engines should take this geographic variation into account in determining their search methodologies to provide optimal and efficient search results for practicing, researching, and teaching radiologists. Further refinements could provide results optimized to the type of user initiating the search or could even utilize the history of a user's previous searches to provide more "personalized" results.

ACKNOWLEDGEMENTS

We would like to thank the creators of Yottalook for their generous support and for providing us with the data to make this study possible. Also, we would like to thank Dr. Nancy Knight for her tremendous support in the revision of this manuscript.

Disclosures Dr. Khan Siddiqui is a Cofounder and Stockholder of iVirtuoso, the parent company that created Yottalook.

REFERENCES

1. Ouellette F: Internet resources for the clinical geneticist. Clin Genet 56(3):179–185, 1999
2. Noxon JO: Dermatology and the World Wide Web. Vet Clin North Am Small Anim Pract 29(6):1461–1473, 1999
3. Evdokimov ME, Aksel'rod BA, Babalian GV: The anesthesiologist and the World Wide Web. Anesteziol Reanimatol 5:85–87, 1999
4. Vorbeck F, Zimmermann C, Vorbeck-Meister I, Kainberger F, Imhof H: Internet use in radiology: results of a nationwide survey. Eur J Radiol 31(2):141–151, 1999
5. Casebeer L, Bennett N, Kristofco R, Carillo A, Centor R: Physician Internet medical information seeking and on-line continuing education use patterns. J Contin Educ Health Prof 22:33–42, 2002
6. Rowell MR, Johnson PT, Fishman EK: Radiology education in 2005: World Wide Web practice patterns, perceptions, and preferences of radiologists. RadioGraphics 27:563–571, 2007
7. Tang H, Ng JH: Googling for a diagnosis—use of Google as a diagnostic aid: Internet based study. BMJ 333:1143–1145, 2006
8. Radiological Society of North America. RadLex. A lexicon for uniform indexing and retrieval of radiology information resources. Available at: www.rsna.org/Radlex/index.cfm. Accessed 5 May 2008
9. Ivirtuoso, Inc. Yottalook—About Us. Available at: www.yottalook.com/about.html. Accessed 10 September 2007
10. American Board of Radiology. Diagnostic Radiology Study Guide. Available at: www.theabr.org/DR_Pri_Study.htm. Accessed 10 September 2007