Oncology Information on the Internet

Yasushi Goto* and Takahide Nagase

Department of Respiratory Medicine, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

*For reprints and all correspondence: Yasushi Goto, Department of Respiratory Medicine, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. E-mail: ygoto-tky@umin.net

Received January 15, 2012; accepted February 21, 2012

Owing to new developments in Internet technologies, the amount of available oncology information is growing. Both patients and caregivers are increasingly using the Internet to obtain medical information. However, while it is easy to provide information, ensuring its quality is always a concern. Thus, many instruments for evaluating the quality of health information have been created, each with its own advantages and disadvantages. The increasing importance of online search engines such as Google warrants the examination of the correlation between their rankings and medical quality. The Internet also mediates the exchange of information from one individual to another. Mailing lists of advocate groups and social networking sites help spread information to patients and caregivers. While text messages are still the main medium of communication, audio and video messages are also increasing rapidly, accelerating the communication on the Internet. Future health information developments on the Internet include merging patients’ personal information on the Internet with their traditional health records and facilitating the interaction among patients, caregivers and health-care providers. Through these developments, the Internet is expected to strengthen the mutually beneficial relationships among all stakeholders in the field of medicine.

Key words: information – Internet – cancer – oncology

INTRODUCTION

The Internet has changed our lives in many respects, increasing its significance and facilitating convenience with recent advances in technology. Today, we can easily access news, maps, music, movies and other information from around the world with the click of a mouse. The Internet also enables real-time communication with other people through text, sound and images. These advances in the flow of information signal the information revolution, which follows the agricultural and industrial revolutions.

The trends and advances in information technology have unexceptionally influenced the field of medicine. One of the first developments facilitating this advance was the emergence of MEDLINE® in 1971, the online version of the computerized medical bibliography, Medical Literature Analysis and Retrieval System (MEDLARS®), first launched in 1964 by the National Library of Medicine (NLM) in the USA. The system was made available to the public in 1997 and renamed PubMed, enabling individuals to access medical information from anywhere. Advances in the so-called evidence-based medicine are closely related to that in information storage. Undoubtedly, the Internet has advanced medicine for physicians as well as patients. For instance, over 80% of the Japanese population has access to the Internet, and 60% seek medical help there (1). This increasing tendency to turn to the Internet to research medical problems seems to be universal (2). However, as with every new technology, the Internet has its pros and cons. While patients can now access medical information online, a significant advantage from early medicine during the age of Hippocrates, without in-depth knowledge of
medicine, the wealth of information may confuse rather than help patients in their decision-making. In this review, we discuss the advantages and disadvantages of seeking cancer information on the Internet, focusing mainly on the users’ experience.

**USE OF THE INTERNET IN CANCER PATIENTS AND CAREGIVERS**

Increasing numbers of patients and caregivers are using the Internet to access the health information (1–5). About 80% and more than 50% of the US (6) and the Japanese populations (1), respectively, turn to the Internet for health information. Some studies even showed that consumers first seek information from the Internet far more than from their doctors (2). However, the role of physicians remains the same regardless of the Internet information, in terms of their prescription of drugs to patients and gaining their trust (2,3).

Why are significant amounts of health information delivered through the Internet? The advantages of obtaining medical information from the Internet include reduced cost, increased convenience, reduced feelings of isolation and stigma, and rapid availability and updates (7). For instance, cancer patients turn to the Internet to obtain information on their disease and to assess the discrepancy between the information they receive from health professionals and their actual needs. More importantly, patient and caregivers turn to the Internet to obtain second opinions, information from other patients, interpret their symptoms, seek information about tests and treatments, identify questions for doctors, make anonymous, private inquiries and raise awareness of specific cancers (4).

The Health Information National Trends Survey (HINTS) routinely collect data on the American public’s use of cancer-related information (8). HINTS was developed by the Health Communication and Informatics Research Branch (HCIRB) of the Division of Cancer Control and Population Sciences (DCCPS) of the National Cancer Institute’s (NCI) Extraordinary Opportunity in Cancer Communications initiative in the US HINTS stores questionnaires from past and present surveys. Based on the question, ‘The most recent time you looked for cancer information, where did you go first?’, the survey found that since 2003, the percentage of patients turning to printed materials is decreasing, while that of patients turning to the Internet is increasing. In 2007, over 50% sought information from the Internet, and only ~25% turned to health-care providers first (9).

Some studies also evaluated the type of content sought from websites and other online sources. A study of the launch and development of a Dutch cancer information source found that half of the visitors found the websites through search engines or links from other sites and that the number of accesses increased over time (10). The study also revealed that most visitors were able to find the information they wanted and that caregivers were the most popular visitors (57%). A similar study by the Korean National Cancer Information Center discovered that about one-third of the patients or caregivers sought information on treatments while the general population sought information on prevention and diagnosis (11).

Many health-care providers understand the patients’ desire for more information (12,13). In 2003, the American Society of Oncology estimated that 30% of the patients use the Internet (12). They also reported that using the Internet has positive and negative outcomes, arming patients with hope and knowledge, while also potentially causing confusion and anxiety. The report also revealed that 44% of the oncologists had some difficulty in discussing Internet information, while 9% reported being threatened while discussing this topic. As of 2011, the situation has not changed much. While many health-care providers recognized the importance of the Internet for patients, they did not routinely refer patients to Internet-based information, mainly due to their unfamiliarity with Internet-based information (13). These conditions may be one of the reasons why patients do not tend to discuss their use of the Internet with their health-care providers (14).

**QUALITY OF CANCER INFORMATION ON THE INTERNET**

The ease of uploading and sending information through the Internet has led to the sharp increase in the available information online. However, the lack of censorship in the uploaded information makes quality the biggest concern, an issue that has been emphasized since 1996 (15). A 1998 report compared obtaining health information from the Internet with drinking the water from a fire hose (16). Since then, many studies on the accuracy or reliability of health information from the Internet followed, using criteria such as accuracy, completeness, readability, design, disclosure and references in assessing information quality. Although these studies used different criteria in their evaluation, the majority (55%) of them highlighted the quality of information from the Internet as an issue (17).

In this paper, we follow up on studies evaluating the available information on cancer on the Internet. A study on health information on breast cancer in both English and Spanish found that only ~50% of English and 25% of Spanish websites provided completely accurate information (18). In addition, the websites in both languages required advanced reading abilities. A similar study on 74 websites related to melanoma found using eight search engines also revealed that most of the sites did not offer complete information on the disease (19). Another similar study on 114 sites on pediatric neuro-oncology found through search engines discovered that 60% were either poor or very poor and required, on average, advanced reading levels (US grade 12 and above), well over the suggested level for adult audiences. Studies also reviewed the accuracy of information from a medical standpoint. Using several criteria, a study
found that only 54% of websites were important (20). Meanwhile, a study on high-ranking websites on lung cancer obtained from a Japanese Google or Yahoo search found that <50% of them were generally inaccurate in their descriptions of Stage IV treatment. In addition, while many commercial sites ranked highly, they generally provide inaccurate information in favor of their product (21). Indeed, commercial interest is a problem in providing information. A study on esophageal cancer websites using major search engines in the UK found that many of these sites provided inaccurate information (22).

Concerns regarding information accuracy are not limited to cancer. A US report also found that information on illicit drugs are also widely available on the Internet (23), which provided potentially harmful recommendations for the management of adverse events. A study on high-ranking commercial sites on mental health or disorders in Japan found that only ~25% of the top 37 sites provided high-quality information (24). Meanwhile, a study of arthritis information discovered that high-quality information is mainly provided by government and academic websites (25).

The inconsistent quality and accuracy of health information on the Internet does not help patients and caregivers, who will turn to anything for help. However, there are many instruments that have been developed to facilitate the search for high-quality, accurate information over the Internet. One of the very first instruments are directories, which are ‘[areas] of a computer disk which [contain] one or more files or other directories’. Internet directory use key words to categorize available information. For example, users can search for websites about lung cancer by entering health information > cancer > lung (Fig. 1A). This categorization and search method is the same as that for files in our personal computers. A problem with this system is that each site chooses its own key words to categorize its information. For example, if this article was categorized under ‘cancer’ and ‘information’, users searching the Internet for other than these key words will most likely not find this article. Thus, using key words does not solve information search issues completely. While the use of multiple key words makes it difficult to categorize each site, the use of less key words makes it difficult for users to find the information they want. Another critical weakness of categorization is that it does not reveal the reliability of a site’s information. For example, under the category of lung cancer, official information from a cancer institute and information from a patients’ blog have the same classification. Thus, there is a need for a better ranking system that indicates reliability.

The second generation of instruments that rank websites is search engines, the most widely used of which is Google. Google was founded by two Stanford University students in 1996. Before the development of Google, search engines ranked results by how many times a searched term shows up on the pages. Google improved on this method by focusing on the relationships among websites (Fig. 1B). Its original program monitored a website to gather information and to create an index. It copies the links from the page and the pages linked to it. These pages are then processed to create an index. Queries on the Google search field will be compared with this index, which then returns the most relevant matches (26). It evaluates the popularity of a site through the pages that link to it and those it links to. Google has been the most popular search engine since 2002. Its automatic indicators of reliability show some promise in helping consumers in their search for reliable information (27).

However, the popularity of sites is not a sufficient criterion for evaluating their accuracy. For example, a study that evaluated the usefulness of sites related to breast cancer on the basis of the quality of their information (28) found that while more popular sites tend to provide more updates on breast cancer research, the quality of their contents, as evaluated by authorship, attribution or reference, display of quality label seal and display of email addresses and other contact information, is not correlated with popularity. The presence of inaccurate information was also more or less the same in both popular and less popular sites. Many studies also showed that rankings by Google or other search engines do not pertain to the quality of information (19–22,29).

Many other instruments have been developed to rate the available health information on the Internet and have been reviewed (30–33). While interobserver usability and validation is important in assessing quality, it is missing in most evaluation tools. A 2005 study evaluated a total of 273 distinct instruments for assessing the quality of health information on the Internet and found that 29% publicly provided their evaluation criteria while 9% used 10 or fewer elements in evaluating websites (33). Furthermore, only seven instruments consisted of elements that could all be evaluated objectively; only one

Figure 1. Directory and search engine emphasizing links. (A) In the directory system, each site is classified with its key words. Site A will be found under health site, cancer and lung. Site A is difficult to be found by another key word. Moreover, differences between Sites A, B, C and D are not counted. (B) In the search engine, ‘crawling’ process assembles the information from the websites and develops an index. By emphasizing links, Site A is recognized as the most popular site (line width is reflecting the traffic concentration between the sites). Each search is referred to the index, and the best-matched result will be shown.
consisted entirely of criteria with acceptable interobserver reliability; and only one instrument used criteria for readability. Many rating tools have remained even after the organizations that developed them have ceased to exist. Thus, these instruments have not been further developed (31).

Wilson (32) classified evaluation tools into five categories: code of conduct, quality label, user guide, filter and third party certification. Each category presents its own advantages, disadvantages and potential beneficiaries. For instance, the guidelines by the American Medical Association (AMA) are classified under ‘code of conduct’ and were developed to assess the quality of online content, potential commercial influence on online content and preservation of personal privacy (34). Although following these guidelines benefits both providers and users, they are not mandatory.

Quality labels are also commonly used by sites to indicate reliability. The VeriSign checkmark is the most recognized trust mark on the Internet, while the Health on the Net (HON) seal is the most widely accepted quality label for health-care websites. HON started in 1995 as a non-profit, non-governmental organization, accredited by the Economic and Social Council of the United Nations (UN). The HON identifies eight ethical codes for medical and health websites, known as the HON Code of Conduct (HONcode) (35). Web publishers need to apply to HON to get certified and use the HON seal. After the evaluation and modification processes, certification is given to these websites. This certification is valid for 1 year after which a re-evaluation is needed for extension. These codes are available in more than 30 languages, and the number of certified websites is increasing. DISCERN is another tool that helps consumers without medical expertise evaluate the quality of health information online (36) by completing a brief questionnaire (37). Evidence showed that this tool helps consumers obtain high-quality information (27) and is effective in evaluating the content quality of websites (20).

Filtering and certification by third parties is another way of establishing the reliability of information on the Internet. For instance, many software have been developed to filter out violent content for children. The VeriSign authentication also certifies whether a website is using cryptographic protocols that provide communication security over the Internet using secure sockets layer (SSL). We are making similar tools for assessing health-care information, in particular, for trusted cancer information (38); however, these tools need to overcome certain obstacles to be effective. First, users should be encouraged to use these tools to further improve their search engine rankings. Secondly, the countless number of websites makes it almost impossible to evaluate the accuracy and relevance of all these health information sources professionally.

COMMUNICATION ON THE INTERNET

The Internet not only stores information but also facilitates communication among individuals. For instance, cancer patients can easily create their own groups on the Internet, which functions as a new community that empowers them in their battle with cancer. Such groups can be formed by recruiting patients for clinical trials or creating advocacy groups. Numerous websites recommend participation in clinical trials (39,40). Internet-based clinical trial matching systems have been tested in breast cancer patients. While patients who participated in such trials felt positive about using those systems, they also required these sites to be credible and private (41). The Life Raft Group made remarkable strides in examining the effectiveness of the drug, Imatinib, available to patients with gastrointestinal stromal tumors (GIST) (42). The usefulness of Imatinib in patients with sarcoma has been tested earlier, which helped facilitate the testing, trial and approval of the drug for GIST patients, as presented in June 2007 at the American Society of Clinical Oncology (ASCO) meeting.

Virtual communities are provided by mailing lists and SNSs. As with websites, the accuracy of the content in the exchanged information in these communities is a concern. A study analyzing the information content provided in communities in breast cancer mailing lists found that 10 of the 4600 postings (0.2%) were false or misleading (43). Of these, seven were identified as false or misleading by other participants and corrected within an average of 4 h and 33 min. Another study performed a similar analysis of the information in communities formed by mailing lists hosted by the Association of Cancer Online Resources (ACOR) found that the manager of the mailing list encouraged appropriate communication among the participants, the most common topics were treatment information and how to communicate with health-care providers, and that participants also requested emotional support (44).

Information from mailing lists is distributed similar to messages while those from websites are presented similar to publication reports. Because of the ease of sending messages in virtual communities, it is often thought that inaccurate information would be more prevalent. However, this study found that this was not necessarily the case, owing to the difference between the goals of websites and mailing lists. Websites are usually created as a permanent source of information for users. As we mentioned previously, commercial interests may influence the accuracy of information provided on such sites. On the other hand, information from virtual communities formed through mailing lists is generated from the communication among senders or participants. While commercial messages can be posted on these forums, they are negated by the personal knowledge and experience shared among the participants.

SNSs are an advanced form of mailing lists where participants represent themselves through profiles and find other users with similar interests. MySpace, Google and Facebook are among the most popular SNSs. SNSs have an advantage over search engines in enabling users to efficiently share and acquire information they are interested in. A 2008 study on virtual breast cancer communities found 620 such groups on
Facebook. The most common goals of these groups were fundraising (45%), awareness (38%), promotion of product or service (62%) and patient/caregiver support (86%) (45).

While there are currently no studies comparing the effectiveness of Internet and face-to-face communication, closely related studies showed that the Internet does not provide sufficient support to patients. A randomized study comparing the impact of Internet and video communication in deciding to take prostate cancer screening tests found that the Internet was less useful because participants did not frequently access the Internet at home (46). A similar study revealed that support through the telephone was more effective in the cessation of smoking (47). Studies have also found that general information does not sufficiently help patients in their decision-making. In a randomized trial examining the usefulness of existing websites on the decision to undergo prostate cancer screening, participants were randomly assigned to look for one of four Internet conditions. The control conditions were links to public prostate cancer-specific websites from credible sources, while the other conditions were online decision aids providing information. The study revealed significant results on prostate-specific antigen test choices, finding that participants looking at public sites tend not to review information and have the lowest knowledge scores (48). Meanwhile, studies assessing the usefulness of the Internet on medical information search have also been increasing (49,50). One such study on breast cancer patients found that their depression and reaction to pain were significantly reduced with the help of electronic support groups (49).

FUTURE OF CANCER INFORMATION ON THE INTERNET

The Internet has changed our daily lives in many respects. Advances in the field of medicine are expected to go in hand in hand with those of the Internet. However, for the Internet to be a reliable source of cancer information, quality assurance of the information available is necessary. Based on the studies we have discussed, information quality has improved in English-speaking countries but is still developing in other regions such as in Japan (21,51). In particular, Japan needs to overcome the language barrier. Since it mainly uses the Japanese language, the available governmental or public information on cancer is limited. Therefore, non-profit and non-governmental organizations can play a significant role in addressing this need. The improvement of the online content in English-speaking countries (52) signals promise.

Guidelines and protocols on the Internet provide quick references for clinicians (53). Website publishers can improve the quality and reliability of online information by promoting evidence on best practices, treatment processes, patient outcomes and potentially avoidable medical errors. The evaluation of published articles online may also gradually change. In addition to Web of Science, Scopus and Google Scholar now also monitor online citations. A 2008 study compared these sites in terms of their citations of published general medical journals and found that while Google Scholar and Scopus found more citations than Web of Science, the citation articles among the databases did not greatly differ (54). However, Google Scholar may have some advantages as the only free-access database.

The available information on the Internet is growing owing to technological advances that efficiently process large amounts of data. Some of these developments pertain to audio and video communication. For instance, Youtube, the world’s largest video-sharing website and a subsidiary of Google, is widely used by organizations including the NCI in the USA, helping it broadcast its messages to its audience. Youtube also helps broadcast personal experiences and stories by patients (55). Another example of the effective use of multimedia is by the Database of Patients’ Experiences (DIPEx) charity, which identifies issues relevant to those who are ill. Its official website, healthtalkonline.org, contains many patient stories. A Japanese version (DIPEx-Japan) was launched in 2007. JPOP-VOICE ((jpop-voice.jp)) has also been providing accurate medical information through videos of patients they interview. Other support groups also share video narratives of patients, such as the Lance Armstrong Foundation and the American Cancer Society’s ‘Stories of Hope’. Currently, we are also in the process of integrating SNS and video message features in our website, ‘Cancer Channel (cancerchannel.jp)’. As with other websites, we are working to resolve online information issues including privacy. As the Internet increasingly uses videos and audios in distributing information, tools for evaluating their quality and reliability are needed.

Another possible but underutilized method of providing patient support is integrating patients. Websites such as PatientsLikeMe.com enables members to share details of their health with other members or, depending on their privacy preferences, with anyone on the Internet. The site was founded by two brothers who lost another brother to amyotrophic lateral sclerosis. The site provides virtual communities for specific diseases. As of December 2011, the site has over 100,000 participants and cites over 1000 conditions. Acor.org is another online community specific to cancer patients. Patients who pro-actively participate in their healing are named e-Patients. e-Patients are those who are not only good at electronic systems, but also equipped, enabled, empowered and engaged (56). They turn to online resources to collect information about their medical conditions, hoping to achieve better outcomes and to build new relationships with their doctors. TED, a non-profit organization spreading free knowledge and inspiration online, presents a resource on the concept of e-Patients (57).

The next step in the advances in medical information is assembling patients’ personal health information from the Internet. Unified standards have been considered in a project by technological leaders such as Microsoft and Google in
this regard. However, while many consumers are aware of the usefulness of the tool, a relucant few used it (10%) (58). Google eventually retired this program in January 2012 owing to the low impact.

It is important to note that not everyone is familiar with the Internet. Although the proportion of those using it is increasing, some issues still need to be resolved to broaden its use. For instance, racial minority groups account for only 10% of all cancer survivor stories on the Internet, which evidences the inequality in available information on the disease (59). Other studies showed that cancer survivors access the Internet at a lower rate than the general population and that higher likelihood of Internet access is associated with younger age, higher education, non-Hispanic, White race/ethnicity, metropolitan residence and better self-rated health (60).

Finally, advances in web technology itself may further change the provision of medical information. A study by Greenwald (61) found that Google was useful in yielding change the provision of medical information. A study by health (60).

Finally, advances in web technology itself may further change the provision of medical information. A study by Greenwald (61) found that Google was useful in yielding information for the diagnosis of rare diseases. However, Internet searches do not yield similar results. For instance, while a simple key word search on Google of ‘biopsy report and laboratory test’ yielded the relevant disease (61), a search using ‘back pain, dyspnea, and hemoptysis’, does not result in the diagnosis of lung cancer. Thus, improved web search technologies are needed to facilitate the diagnosis of diseases. The next-generation Internet named ‘web 3.0’ or ‘semantic web’ will help identify the connections among seemingly unconnected raw data, which then yield hypotheses, conclusions and ultimately, knowledge (62).

**CONCLUSION**

Advances in the field of medicine, especially in the areas of information availability and virtual communities, are in tandem with those of the Internet. To strengthen the mutually beneficial relationships among patients, caregivers and health-care providers, numerous issues need to be resolved, including segregating commercial interests from scientific ones and re-constituting raw information. Further improvements will make the Internet an effective bridge between patients and health-care providers.

**Conflict of interest statement**

None declared.

**References**

6. Sox S. Health Information Online: eight in ten Internet users have looked for health information online, with increased interest in diet, fitness, drugs, health insurance, experimental treatments, and particular doctors and hospitals. Pew Internet Am Life Project 2005:1–22.
Oncology information on the Internet