The Zeitgeist of Online Health Search

Implications for a Consumer-Centric Health System

Daniel P. Lorence, PhD, JD,[1] *Liza Greenberg, RN, MPH*[2]

[1]Department of Health Policy and Administration and The School of Information Science and Technology, The Pennsylvania State University, University Park, PA, USA; [2]URAC, Washington, DC, USA.

BACKGROUND: Health care queries rank among the most frequent topics of information-seeking activity initiated by users of commercial search engines. The quality of information located through existing search engine technology has received little attention, especially when considering the widely varied knowledge levels of internet users.

OBJECTIVE: This study sought to create a benchmark technology assessment of online health search trends and practices, with corresponding evaluation of its applicability within the Federal Health Architecture (FHA) plan for a nationwide, interoperable health information infrastructure.

DESIGN: Exploratory technology assessment, analyzing focus group participants' views on barriers to effective health information searching, using existing commercial search engine technologies and methods.

SETTING AND PARTICIPANTS: Focus group, national leaders in electronic health care (e-health).

RESULTS: A variety of web-based assessment tools are available for consumers to be able to identify reliable health websites; however, many may be too difficult for the layperson to use or understand. Existing search technologies are increasingly powerful, although the expanding volume of information on the internet suggests the need for better mediated searching. Search engines provide consumers a means for quickly bypassing information that appears too technical for their individual knowledge level, and at times, searchers often overlook critical information most relevant to their needs. Overall, existing search technologies need to be more interactive, visible, and context-driven, and supported by better technology assessment methodologies, scalability of information, and enhanced access by underserved subgroups.

CONCLUSION: Future technology assessments are needed to provide structure for interoperability of health information systems, especially where consumers, providers, and payer systems intersect. State-of-the-art search engine technologies are still not widely available to those who can benefit most from them.

KEY WORDS: consumer health; technology assessment; health policy; disparities; communication; health behavior; medical informatics; underserved populations; education technology; patient-centered care.
DOI: 10.1111/j.1525-1497.2005.00295.x
J GEN INTERN MED 2006; 21:134–139.

There is little doubt that seeking health information is one of the most common reasons why consumers use the internet. Of Americans with internet access, 62% have used the Web to obtain health or medical information, and about 6 million Americans go online in search of health information on an average day.[1] The promise of the internet to empower consumers and increase their ability to make intelligent health care and lifestyle choices seems clear. The overall quality of information on the Web, however, remains highly variable. Health information on websites can range from personal testimonials, to product-promotion sites, to sites containing evidence-based health content. These sites may not be updated regularly. Most studies in this area suggest that information on health websites is often incomplete or outdated, and that some sites are far better than others.[2,3] In effect, a consumer's online health experience may be greatly influenced by the site from which they obtain their information.

An added concern: when consumers do access health information on the Web via a search engine, most do not later discuss the information with a health care provider.[1] For many health consumers, the internet may be the primary (or even sole) source of health information. As a result, consumers often find and act upon health information of variable quality, with little input from health professionals.

Search engines and directories play a central role in facilitating access to online health information. Eighty-one percent of consumers seeking health information online do so through a search.[4] For many, search engines are the main source of guidance for information that will be used to help make decisions.[5,6] The criteria used to identify and rank health-related websites, likewise, vary among search engines, and may be unknown or ignored by most consumers.[7,8] Content factors as well as paid placements may also affect search results. There is a growing need for more information about how searches could be targeted more effectively to assist consumers in accessing quality health information.

Online Search Within a National Health Information Infrastructure

Although not contemplated at the initiation of this study, the announcement of plans for a National Health Information Infrastructure (NHII) created a new focus on health consumer information. While not designed to address the NHII specifically, findings from this project thus take on an added dimension, given the consumer-centric focus within the NHII.[9] A posthoc review of focus group findings therefore examined, to the extent possible, the implications within the context of the evolving NHII plan.

Integrating Search Process and Search Outcomes

Exploratory studies to date have focused mostly on the quality of the information found on websites. We sought to understand

---

The authors have no conflicts of interest to declare for this article.

Address correspondence and requests for reprints to Dr. Lorence: Department of Health Policy and Administration and The School of Information Science and Technology, The Pennsylvania State University, 114 Henderson Health and Human Development Building, University Park, PA 16802 (e-mail: dpl10@psu.edu).

Manuscript received April 18, 2005
Initial editorial decision July 12, 2005
Final acceptance September 13, 2005

the broader environment of online search by incorporating search technologies as well as information-seeking behaviors, mindful that retrieval of information from the Web involves a stepwise series of tasks, rather than an effortless review of website material. Searching in this context encompasses information search decisions, search strategies, results compilation, site selection, and site content review. This search process model and related framework are outlined in Figure 1.

Research Objective

The ultimate goal and aim of this project was to define a research and development agenda that can be implemented by a cross-disciplinary group of stakeholders to improve consumers' ability to locate quality health information on the Web.

METHODS

Study Design

This was a descriptive study based on expert focus group participation. Stakeholder meetings and postconference review occurred throughout 2003 to 2004. The meetings were designed for national leaders in e-health to clarify and define the issues and challenges raised by online health search, with a focus on how best to enable consumers to access quality online health content reliably. The study was a part of a larger technical review of the effectiveness of web-based search technologies and strategies related to seeking consumer-friendly health information on the internet. Volunteer focus groups were chosen to collect data because they provided a means for understanding the complex behaviors related to health information seeking. They also allowed an exploration of the degree of consensus among group representatives about the barriers to effective health information search, and resulted in data collection that was respectful of opinions from a diverse group of stakeholders.

Sample

Sixty participants were recruited from among identified leaders in web-based health care delivery and support, and were randomly assigned to 1 of 2 focus groups. Inclusion criteria included being in a management or senior technical position related to delivery of web-based health information. Participants were sent an invitational letter to attend a group session. The final sample consisted of 53 participants in 1 of the 2 sessions.

Literature Search Methods

Literature responding to selected keywords related to search engine use, e-health, and consumer health in the context of quality, cost, and access to health care were identified using the Cumulative Index of Nursing and Allied Health Literature, MEDLINE, PubMed, Expanded Academic ASAP, Lexis-Nexis, Proquest, Ingenta, and related databases in health care, information science, and computer science. Issues were identified that discussed how consumers can more readily find relevant health information using web-based media and search tools. Where initial searches revealed poor topic coverage, associated reference lists, books, and media that were considered to inform the topic were included. The literature was then examined for common themes and concepts and was reported using descriptive analysis. An initial 8-member panel examined the content of literature on access to web-based information, as well as the quality and usability of such information by consumers.

A modified Delphi method (see Fig. 2) was used to help structure information flow and promote feedback, which was used to condense detailed comments and prioritize the issues most relevant to web-based health care.[10,11] This process sought to create groups that would take advantage of expert opinion, experience, and intuition, and allow full use of the limited information available, when full scientific documentation was often lacking. Participants reviewed the literature provided and answered a series of prioritizing questions, formulated as hypotheses, and experts stated their estimates when they thought these hypotheses would be fulfilled. Each round of questioning was followed with the feedback on the preceding round of replies, with the option to respond anonymously. The experts were encouraged to revise their earlier answers in light of the replies of other members of the group, converging toward narrowed topic consensus. After several rounds, median scores determined the final discussion topics. The results were used to finalize a list of most important issues, present and future, related to searching for health information on the Web. From this feedback, a focus group guide was developed and sent to participants 3 weeks prior to the group meeting.

Two focus groups were held: 1 in San Francisco and 1 in Washington, DC.

After informed consent and a discussion of focus group ground rules, the facilitator used an interview agenda to elicit responses and generate discussion. Group comments were transcribed, and the facilitator recorded notes on whiteboards and flip charts. The facilitator asked directed questions and used appropriate open-ended prompts when needed to generate group discussion, validate key points, and to indicate

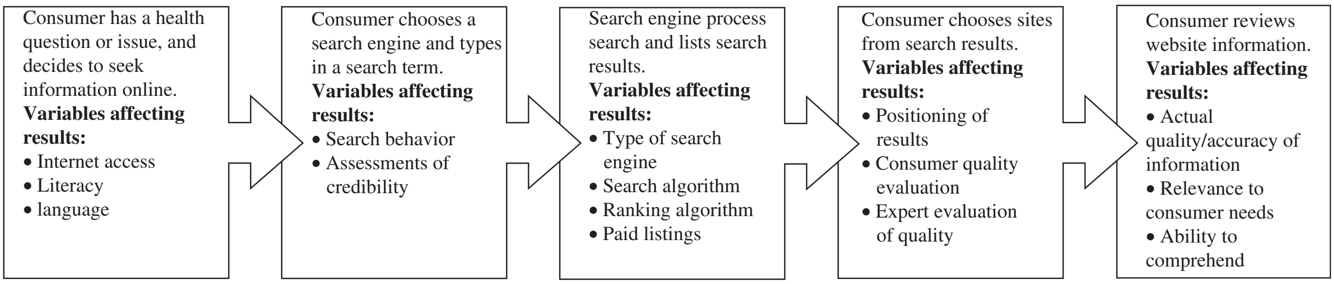

FIGURE 1. Incremental search process model.

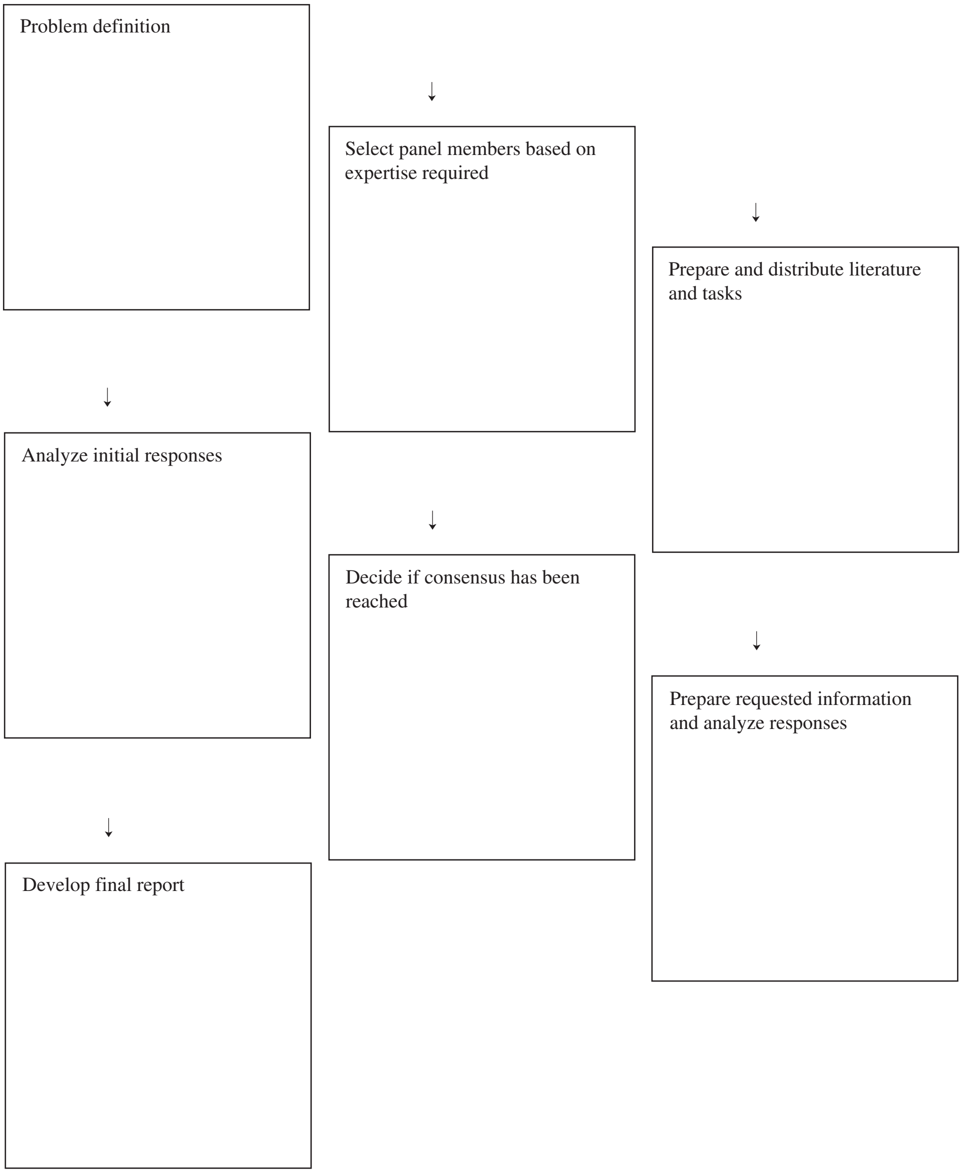

FIGURE 2. Summary of the Delphi process.

priorities. At the close of each group, the facilitator summarized the main discussion points and solicited additions or changes. The meeting ended when all participants agreed that the key issues had been noted. Each group session lasted approximately 7 hours each.

Analysis

Participant feedback was analyzed to identify the most salient beliefs from the focus group data and to understand facilitation and barriers for implementing proposed guidelines. Feedback was grouped by theme within each topical construct. The facilitator organized responses hierarchically by the level of importance (high to low) based on the priorities and emphasis given, and independently reviewed data categories. Study leaders resolved any discrepancies and grouped results into higher-level categories.

RESULTS

The basic issues reviewed within this study are summarized as follows:

Issue 1: Are There Sufficient Tools Available for Consumers to be Able to Identify Reliable Websites?

Findings from this discussion suggest that a number of criteria have been proposed as checklists for the consumer to use when evaluating a website, but it is questionable whether they would use such instruments in a Web environment, given the tendency to scan and click key terms rather than carefully read all content. The need for tools to enhance recognition of quality websites was therefore deemed a high priority to facilitate consumer searches. Such tools may be implemented by websites themselves, for example, through increasingly sophisticated coding to highlight quality content, and toward consumers, to direct them more effectively to relevant, high-quality information. In addition, as there are currently multiple tools for evaluating the quality of websites, a coordinating body will be necessary to standardize and validate the effectiveness of these tools.

Implications for the NHII are that, across integrated and connected systems, consumers will need gateways to filter information, increase its relevance, and provide expert assessment regarding the validity of available sources. There will also be a demand for more sophisticated models for providing useful and relevant information to consumers via customized approaches. Such approaches could potentially be embedded in search algorithms. Given the stated goal of standardization and integration of systems under NHII, priorities relating to the quality of websites could serve as the beginning of a national research agenda on consumer health information.

Issue 2: Are Existing Search Technologies the Best Way for Consumers to Search for Information?

There is clear consensus among e-health leaders that commercial search engines are becoming increasingly important as a tool for locating and organizing information from the vast resources of the internet. The volume of information on the Web, however, is so great that consumers may need different types of mediators, such as search engines or librarians, to help manage the volume of information. Consumers are often unaware of the limitations of their search strategies, or are averse to using either electronic or human search support systems. In a national framework specific to health care, it may be possible to develop search technology that more effectively steers consumers to quality results. While librarians may have the skills needed to conduct better-mediated searches, there is doubt as to whether they could fill such a role in most existing information environments.

A number of participants suggested that more information about search algorithms and quality factors identified in the algorithms, specifically for health, would be of value. Unfortunately, there is little motivation or desire for search companies to reveal such proprietary information at this time. A national policy-making body, if formed, might possess the clout to negotiate with commercial providers in order to yield more health care-friendly searches.

Within the NHII agenda, an emerging group of new technologies could be used to link the electronic health record (EHR) to provider-specific sources, or the broader Web, and mark quality content or assist search engines in learning as they search. Search engine providers are also developing technology to search for semantic–neutral synonyms, which may enhance health searches conducted by laypersons. These technologies may ultimately be more effective than requiring consumers to either apply filters or modify their search strategies.

Nationally, there is consensus that with projected near-term advances in search engine technology, it may be possible to identify quality proxies that could improve page rankings of quality websites. Search engines could, for example, give higher ranking to official sites maintained or approved by Health and Human Services for diseases, or piggyback onto credibility assessments provided by groups such as Healthfinder.gov. Most search technologies have yet to be tested and evaluated. Ultimately, adoption of technological solutions depends on the ability of the health services research community to validate the correlation between proxies and quality content. The Office of National Coordinator of Health Information Technology might serve to facilitate additional research needed to assess the effect of such technology, differences in search engines, and the relationship of search results with consumer needs and actions.

Issue 3: Independent of the Search Process, Do Most Consumers Possess Enough Knowledge to Understand What They Find?

There was consensus that a national lead, possibly in the Department of Health and Human Services, will have to be designated for health search literacy. The lead organization should collaborate with other agencies and national organizations on a leadership summit for health search literacy. The summit could leverage interest on chronic disease and the need for consumer information on managing chronic health. Interested entities include the office surgeon general, The U.S. Office of Disease Prevention and Health Promotion (ODPHP), the Centers for Medicare and Medicaid Services), the National Institutes of Health, and the other health organizations.

Funding sources, including foundations and the U.S. Department of Health and Human Services, will likewise need to establish a comprehensive, long-term research agenda for internet health information. The ODPHP has already established goals for health communication that can drive much of the agenda. The California Health Care Foundation has also identified a number of recommendations for improving internet-based health information that can be incorporated into the agenda. Development of methodologies to test the effect of online information on consumer action is a needed first step. The methods should be statistically valid for nationwide and global comparisons, as well as for specific segments of the population. We may also need to develop and use new technology to conduct research on search behavior, for example. What outcome measures are appropriate for health information implementation?

Under an NHII, the implications are that a planned EHR infrastructure could be linked to specified resources already maintained by education organizations, provider groups, and health services researchers. For people with disabilities, for example, tools and approaches based on consumer segment research, tools and approaches for persons with special needs can be built and integrated. High-priority consumer segments may include the disabled, low-literacy, and non-English-speaking groups. The integrated EHR, as described in the most recent NHII, may be able to take advantage of existing resources by supporting human-mediated searches, incorporating reading-level rankings as a quality factor in website evaluations, and identifying Web design factors that enable users to access the site.

The power of a shared provider-consumer EHR implies that provider organizations should also be used to educate provider members on the value of offering "information therapy"[12] and provide them with tools to assist and steer consumers to high-quality sites. These groups could potentially also provide education for physicians on how to work with patients on internet searching, perhaps using search education brochures and training Web pages. Health organizations will also need to educate health website developers on how to make information findable and how to meet the content level of their intended users. For example, health websites may be able to tailor their reading level and navigation features to promote access. There will be a need for education organizations, in collaboration with health organizations, to develop school-based or publicly available health search curricula for those who would be amenable to such training, for example, parents/grandparents who are not internet savvy. For younger consumers and frequent searchers, educational components can be embedded in the search itself through pop-up clarifiers or questions.

Issue 4: What Can Be Done to Improve Existing Search Technologies?

There is general consensus among e-health leaders that technology companies should continue to develop interactive features on search engines and sites to customize and personalize health searches. These could be visible to consumers (such as creating a demographic profile) or invisible, through search engines that learn consumer preferences for types of websites. Search engines could also offer a pop-up query to clarify the intent of the searcher. The technology would have to be compatible with acceptable standards for ensuring the privacy of consumers. Recognizing that commercial search providers may be hesitant to incorporate such changes, it was believed that interactive features could be implemented through technology innovations adopted by the search engines themselves, or through the development of marker technology embedded in health websites.

The potential exists to establish compatibility of the NHII and the EHR, which would enable search engines to mediate selected health queries by offering additional relevant information. For example, search engines could offer relevant links for general health searches, such as the Food and Drug Administration website, for EHRs that contain certain drug names.

Another issue related to both Web information and linked patient records is timeliness. A centralized coordination of technology will be needed to indicate when information supersedes previous information. The replacement function will update information in search listings and archive outdated sources. Also, development of technological markers or indicators across web-linked records will be needed to capture the accuracy and comprehensiveness of health websites as performance elements that would enhance a site ranking. Accreditation may be one such quality indicator. There was an identified need for search engines to collaborate with health quality experts to identify validated quality proxies to embed in search algorithms. This would increase search return rankings of higher-rated health websites. Expert raters might include medical specialty societies, accrediting bodies, and the federal government. Ratings would ultimately enhance the rankings of higher rated sites.

Data fusion programs currently used in other industries could be used to filter health information according to different measures of quality for dissemination to consumers, again either linked to their EHR or prompted by the patient record content. Search engines could also develop health push technology like Amazon.com to direct consumers to information they say they need, but also update and customize delivered content based on what they actually view and buy. For example, development of a function to clarify and enhance queries (e.g., people who search for xx also often search for xyz). Ultimately, personalized searches can be readily developed by building in the capacity for search engines to remember customer search strategies and what they are looking for, compatible with appropriate privacy protections.

Issue 5: What are the Most Pressing Research Needs or Unanswered Questions Related to Online Health?

Findings from this analysis suggest that there is a great variation in the ways in which consumers seek information via the internet, and in how successful they are in searching for health information. As there is significant consumer-level variation in how searches for health information are performed, search algorithms that support variation and still return expected results will be necessary. The literature suggests a need for strategies to support consumers who are looking for information, particularly those who are not familiar with the internet or for whom language or literacy is a concern.

As the consumer-centric NHII takes shape, additional research will be urgently needed on the information needs of different consumer segments and strategies to educate and improve the results of their health searches. Further, given the wide discrepancy in how consumers and experts evaluate the credibility of websites and problems with the completeness and accuracy of website content, additional research will be needed to validate efficiently the quality of websites and communicate this information to consumers. Stakeholder groups recommended a number of key research topics that would help to illuminate consumer behavior. This information would drive content management and site construction techniques and technologies in health websites, search engines, and educational approaches that could improve the results for consumers. We summarize here research priorities that are consistent with the likely NHII needs, including:

- Consumer segmentation research to identify search needs and capabilities of demographically different types of searchers, and searchers with differing health needs. Such segmentation has been shown to drive education and outreach initiatives for targeted subgroups.
- Research related to how consumers use health terms and the impact of using specific terminology. How do people search for health topics? How does the search term and search starting point influence the end results? How do semantics influence search behaviors and results?
- Identification of specific contents that information consumers are actually reading and responding to on a website. Also, developments of strategies to better understand what consumers really want from health information, and which elements of information drive behavior change. This will help identify markers to guide consumers and search engine information retrieval.
- Examination of how people define a successful search experience. How is success defined both for the search process itself, and for retrieval of content? Research could be designed to better understand the cognitive processing that goes into a health search and to correlate satisfaction with the quality of search results.

- Examination of stepwise processes within the search strategy. For example, searches starting at a health plan site might take a very different course than those beginning with a search site like Google. This information will help prioritize outreach to different types of search developers and design support systems to facilitate consumer searches using information from the EHR.
- Evaluation of differences between U.S. and international information seeking, particularly in terms of how different groups search and how they assess the credibility of health information on the Web. This is critical to an ethnically and racially neutral patient information resource and enhances the likelihood that patients will embrace their role as informed consumers.
- Initiation of more in-depth research on content factors influencing physicians' perceptions of web-based information quality, its relation to the EHR, the effect of physician recommendations to patients about using the Web, physician responses to patient queries about internet information, and facilitation of shared decision making.

Study Limitations

Alhough not contemplated at the initiation of this study, the announcement of plans for a NHII created a new focus on health consumer information. As outlined in the strategic plan, the goals of the NHII are to promote a more effective marketplace, create greater competition, and provide increased choice for consumers through wider availability of health information. The initiative makes clear the importance of consumer involvement in health care, as well as the need for expanded deployment of time-critical and integrated information, improved patient-provider interactions, public health, and national security.[10] The need becomes obvious, then, for expanded efforts to support better-informed consumers and shared decision-making environments. This study was not designed to address the NHII specifically, however, the findings take on added significance given the consumer-centric focus within the NHII. Although it is not ideal methodologically, a posthoc review of the focus group findings was used to include the evolving NHII plan.

There have been many cases in which the Delphi method has produced poor results. Some attribute this to poor application of the method and not to the weaknesses of the method itself. In areas such as science and technology, forecasting the degree of uncertainty is sometimes so great that exact and always correct predictions may be impossible. As a result, a high degree of error is to be expected. Another potential weakness of the Delphi method is that future developments are not always predicted correctly by iterative consensus of experts, but instead by unconventional thinking of amateur outsiders.

One of the initial problems found in applying Delphi techniques was its inability to make complex forecasts with multiple factors. Potential future outcomes were usually considered as if they had no effect on each other. Later on, several extensions to the Delphi method were developed to address this problem, such as cross-impact analysis, which takes into consideration the possibility that the occurrence of 1 event may change probabilities of other events covered in the survey. Overall, there is a clear need for more precise measures of e-health.

CONCLUSION

There is tremendous interest in online health information, from consumers, information providers, search engine scientists, and a range of other stakeholders, including participants is the design and implementation of the NHII. There is considerable opportunity to improve the online health experience for many consumers, to achieve consensus on quality criteria and methods, and for consumers to take a more active role in enhancing their own health experiences. The purpose of future national stakeholder summits, and ongoing stakeholder meetings, will be to craft a specific strategy for achieving these goals in an expected consumer-centric, shared decision-making environment.

This project was funded in part by grant number 046902 from The Robert Wood Johnson Foundation with additional support from the Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services, and Consumers' Union.

REFERENCES

1. **Fox S, Fallows D.** Internet health resources: health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Available at: http://www.pewinternet.org/pdfs/PIP_Health_Report_July_2003.pdf. Accessed April 16, 2005.
2. **Marable L.** False oracles: consumer reaction to learning the truth about how search engines work. Available at: www.consumerwebwatch.org/news/searchengines/index.html. Accessed April 16, 2005.
3. **RAND Corporation.** Proceed with Caution: A Report on the Quality of Health Information on the Internet. Oakland, CA: California HealthCare Foundation; 2001. Available at: www.chcf.org. Accessed April 16, 2005.
4. **Portoni L, Combi C, Pinciroli F.** User-oriented views in health care information systems. IEEE Trans Biomed Eng. 2002;49:1387–98.
5. **Khopkar Y, Spink A., Giles C, et al.** Search engine personalization: an exploratory study. First Monday. Available at: www.firstmonday.org/issues/issue8_7/khopkar/index.html. Accessed April 16, 2005.
6. **Bollacker K, Lawrence S, Giles C.** Discovering personally relevant scientific literature on the web. IEEE Intelligent Systems. 2000;15:42.
7. **Berland G, Elliott M, Morales L, et al.** Evaluation of English and Spanish health information on the internet. Rand Foundation Report (2001). Available at: http://www.rand.org/publications/documents/interneteval/. Accessed April 16, 2005.
8. **Berland G, Elliott M, Morales L, et al.** Health information on the internet: accessibility, quality, and readability in English and Spanish. JAMA. 2001;285:2612–21.
9. **Dykes P, Bakken S.** National and regional health information infrastructures: making use of information technology to promote access to evidence. Medinfo. 2004;2004:1187–91.
10. **Helmer O.** Analysis of the Future: The Delphi Method, Report No. P-3558. Santa Monica, CA: RAND Corporation; 1967.
11. **Brown B, Cochran S, Dalkey N.** The Delphi Method II: Structure of Experiments, Report No. RM-5957-PR. Santa Monica, CA: RAND Corporation; 1969.
12. **Seidman J, Steinwachs D, Rubin H.** The mysterious maze of the World Wide Web: what makes internet health information high quality? Available at: www.informationtherapy.org. Accessed April 16, 2005.