ABSTRACT

Consumer reliance on Internet health information is proliferating, creating concern as to the quality and authoritativeness of this information. Given that access to information is a fundamental component of shared decision making, and improved medical, social and health outcomes, the challenge is to provide information which is relevant, timely, accurate and as far as possible meets the dynamic information needs of the user. The paper reports on the Breast Cancer Knowledge Online research project and its ‘intelligent portal’ designed to support medical and psycho-social decision making through its innovative matching of user profiles, metadata tagging and user-aware resource descriptions. The portal will provide a gateway to both breast cancer sources and resources which have been ‘value-added’ in terms of accuracy, provenance, and evidence-based best practice protocols.

KEYWORDS: breast neoplasms; information services; decision-making, computer assisted
INTRODUCTION

In March 2002, *The British Medical Journal* devoted an entire issue to the question of the quality of Internet health information. In its June 5th edition, *JAMA: Journal of the American Medical Association* followed suit with an editorial and feature article investigating the potential harm of unmediated internet health information provision (Croce, Villasis-Keever, Jadad, 2002). The reliance on the internet by the health consumer for medical, psycho-social, and health support needs has been well documented, giving rise to the issues examined by these two journals. Regardless of these concerns however, Internet health use by the general public continues to rise dramatically. In a March survey, Harris Interactive (2002) reported that in 1998 50 million Americans resorted to finding health information via the web; in 1999 this figure had climbed to 69 million; in 2001 it had reached 97 million and currently stands at 110 million. They concluded:

the Internet continues to be used by huge, and growing, numbers of the public interested in getting information about particular diseases or treatments or about staying healthy. The results also demonstrate the critical importance to health care websites of the need to be quickly and easily accessible through search engines and portals. (Taylor: 2002)

For the woman with breast cancer, her family, friends and/or carers, the ability to access timely, accurate and relevant information is a vital component in facilitating shared decision making, possibly better treatment compliance, and hence improved medical, societal, and general health outcomes. (see National Health Medicine Advisory Council: 1998). Government initiatives such as HealthInsite and Healthfinder, are clear indicators of the role of the Internet in fostering consumer access to a wider and more authoritative range of information resources. Through its portal structure and own search engine capacity, the objective is the delivery of quality health care information, which counteracts the dangers of quackery and commercialism often associated with Internet health provision. Yet, portal technology as it currently exists, largely fails to discriminate between users; the gateway to information sources and resources generally caters for the generic patient and/or user. To all intents and purposes, current health portals fail to distinguish between the highly individual and dynamic information needs of, for example, a 63 year old widowed grandmother with early breast cancer and a 35 year old mother of two experiencing advanced disease. In the health arena, factors such as disease stage, age, educational levels, ethnicity, urban or rural location and life-styles, are generally ignored within the rigid parameters of the generic patient. Even with a portal, the information recovered through this gateway by the woman with breast cancer tends to be of the ‘one-size-fits-all’ category.

Intelligent portal technology and knowledge management have the capacity to address issues of ‘personalisation’ and the decision support needs of the Internet health user. A critical challenge facing the designers of intelligent portals is how to provide differentiated access to resources to meet the diverse and changing needs of individual users, and to ‘add-value’ by providing information that also enables users to judge the quality and reliability of the resource. The focus of this paper is thus to outline the
innovative features of the envisaged Breast Cancer Knowledge Online (BCKOnline) intelligent portal and its major objective of delivering differentiated quality information which has been ‘value-added’ in terms of provenance, currency, authority, and possible accessibility costs.

THE BCKONLINE PROJECT


The BCKOnline project consists of the following phases (see Figure 1):

Phase 1  An extensive field study and user needs analysis leading to the development of context-rich user profiles.

Phase 2  Identification, assessment, and ‘quality rating’ of breast cancer information resources in accordance with evidence-based protocols and ‘best practice’ principles of web design.

Phase 3  Description of breast cancer information resources with reference to the outcomes of the user needs analysis, including requirements for value-added information about provenance, authoritativeness, quality, reliability and conditions and possible cost of access. This phase will involve the development of a descriptive metadata schema to support user-aware resource description (metadata schemas define standardised elements to be used in resource description).

Phase 4  Development of architecture for an electronic information resource that includes an intelligent portal that provides a gateway to existing breast cancer information resources; a metadata repository of resource descriptions, and a knowledge repository of additional resources.

Phase 5  Building a prototype of the information resource that demonstrates how an intelligent portal can provide differentiated access by matching user profiles to user-aware resource descriptions. This phase will also involve interpretive analysis of the field study data to refine the information search characteristics to be used in the intelligent interface design. The portal will utilise complex sets of metadata, resource descriptions and user profiles derived from the rich picture of user needs and patterns of use that emerge from Phase 1.

Phase 6  Intensive testing of the prototype with reference to user needs and expectations as identified in the first phase. The evaluation will be conducted from two perspectives: whether the intelligent portal meets differential user needs, and the level of satisfaction with the new resource, including perceptions of the quality of decision-making resulting from its use.
As outlined above, the BCKOnline intelligent portal serves multiple purposes, including the delivery of tailored, better-targeted and reliable medical, health care and related information to users who are in the process of making critical health and life-style decisions. The intelligent portal will improve the structure and manage the quantity of information presented. It will also provide the user with data about the authority, quality, and provenance of the information accessed, so that users can make their own judgments as to its reliability. It will more precisely meet the information needs of the users and provide information in a timelier manner, which can potentially help solve critical life problems. Support for decision-making and related quality and reliability issues are critical aspects of the current project. The intelligent portal will be designed to deal with this requirement, as well as with requirements relating to relevance, accessibility, and useability.

**USER NEEDS ANALYSIS**

The following section reports on the outcomes of the user-needs analysis, which is the pivotal phase in determining the information and support needs of the breast cancer community.

Numerous studies conducted over the past decade have reported on the perceived information needs of women with breast cancer. (see Gray et al.: 1998; Meredith et al.: 1996; McGrath et al.: 1999) Yet, whilst materials on breast cancer, both in print and online, appear to be proliferating, studies continue to document dissatisfaction with the quality, format, accessibility and relevance of existing information. In fact, information provision remains one of the two highest unmet needs of the breast cancer community.
The results of the BCKOnline user needs analysis, which involved the in-depth interviewing of a purposively selected sample of 53 women with breast cancer and 11 breast care nurses, supports such findings. Summarised below are some of the key points relating to the information requirements which an intelligent portal will need to address. Participant comments are included throughout and anonymity is maintained through the random assignment of ‘interviewee numbers’.

**Information ‘overload’**

For many women, searching for breast cancer information on the Internet proved to be a time-consuming and often fruitless experience, exacerbated by the sheer volume of materials retrieved. Interviewee 18 noted that it felt like an ‘avalanche’, whilst other women commented that:

> you have to read, you have to read through too many things. ... it can be ...a bit awesome,. You think “Oh boy - there’s too much, too much information here”, but you want to... know... which way to find it. Have I got to read through all that to find what I want? (Interviewee 3)

**Relevance and timing**

The need for relevant information was widely discussed by participants. Comments revealed a preference for ‘identification’ with the materials provided and an acknowledgement of the individual’s personal circumstances and/or required ‘depth’ of information. Again, the following comments reaffirm the inadequacies of the ‘one-size-fits-all’ approach to health information provision:

> When I was first diagnosed they said I’d have to have a mastectomy and I was just agog. I just couldn’t cope with that and the book...talked about ‘your partner will still love you’. But yeah, what if you haven’t got one? [Interviewee 22]

> I wanted the ‘hard’ information; I didn’t want the ‘fluffy, there, there’ stuff at all. I wanted ‘What does this mean? What’s the pathology? What’s the description of the medical condition? What’s the background? (Interviewee 15)

Other themes that emerged included the need for information to be presented in a way that caters for a range of reading levels and format preferences. One woman found the visual element crucial to her decision making process:

> Look, the most important thing for me was to go and speak to someone else and see someone else’s reconstruction. As soon as I saw that reconstruction I was completely at peace about having it. ..... The visual thing really helped me. I think that was very important in the reconstruction. Or knowing that I was going to have a mastectomy, what was I going to look like down the line? (Interviewee 16)

**Information quality**

For many of the participants, issues of quality and authority were crucial. Contentious medical ‘controversies’, such as the purported links between breast cancer and abortion make awareness of the source of the information essential in order for the individual to
decide whether such material is ‘objective’ or simply biased. The following comment captures this concern:

Yes I looked at the internet, and again, ... you felt that you had brought it on yourself. Like some of the sites that I saw; any breast problems are problems that are emotional and sort of things like that, and you just sort of think “Oh”.
(Interviewee 19)

Nor is this an isolated issue. A recent study investigating the quality of breast cancer web sites, found that “less than one third of the sites identified references or sources for information cited at the site” (Goetz and Clarke, 2000: 283). This concern with accuracy and validity amongst the women in the sample is further reflected in the following comment:

I mean if you’re going to go into the Internet and try to find a site that is going to tell you something, how do you know that it’s valid information? Just by reading it? I think that’s a lot of hogwash. (Interviewee 13)

The data collected and analysed in Phase 1 of the project is being used to develop a rich picture of the patterns of user needs and profiles of ‘classes’ of users. This will feed into the identification of relevant resources and their description in user-aware ways, and into the design of search strategies for the intelligent portal.

USER AWARE RESOURCE DESCRIPTION

In Phase 2 of the project, breast cancer information resources are being identified, assessed, and quality rated. Gaps in information provision are also being identified. Relevant resources will then be described with reference to the outcomes of the user needs analysis and profiling described above. Criteria established by organisations such as DISCERN (http://www.discern.org.uk/), HEALTH ON THE NET FOUNDATION (http://www.hon.ch/), CANCERINDEX (http://www.cancerindex.org/clinks18.htm); MITRETEK (http://hitiweb.mitretek.org/iq/), BIOME (http://biome.ac.uk/guidelines/eval/factors.html), and the JAMA guide to ‘Quality on the Net’ are being used to assess and quality rate the resources, taking into account factors such as accuracy, source, authorship, commercial affiliations and extent of evidence-based protocols. Information resources are also being evaluated according to defined criteria, which address issues such as reading levels, intended audience, use of medical terminology, and suitability for the differing needs of the target audience for the portal. These criteria are being specified using the results of the user needs analysis, and include factors such as educational levels, socio-economic backgrounds, stage of disease, ethnicity, age group and residential location in city or rural areas.

In Phase 3, the breast cancer information resources identified, assessed and quality rated in Phase 2 will be described in ways which enable the portal to match user needs to highly relevant information resources, and to provide value-added information to the user about the resource being accessed, e.g. its provenance, authoritativeness, quality, conditions and costs associated with access – if applicable. This will be supported by the extension of existing resource discovery metadata schema to meet the requirements of user-aware resource description.
Use of metadata

The breast cancer information resources will be described using a set of standardised descriptive elements. Specification of an appropriate metadata schema, which identifies in a standardised way the elements needed to describe a resource, will be a critical part of this component. Evaluation of the potential of existing resource discovery metadata schemas to provide user-aware resource descriptions and information about the quality and reliability of resource is being undertaken. The international de facto standard for describing documents on the web, the Dublin Core (http://purl.oclc.org/metadata/dublin_core), does not address the issue of the target audience for the resource described, and has a limited capacity to provide information that would enable users to judge quality and reliability.

The Australian Government Locator Service (AGLS) is based on and extends the Dublin Core set of descriptive metadata elements (Office of Government Online and National Archives of Australia, 1999). It was originally developed for application in Australian government contexts, and is being redeveloped as an Australian national standard. As the AGLS is about to become the national standard for resource discovery metadata in Australia, and provides a more extensive set of metadata elements than the Dublin Core, including an Audience Element, the AGLS metadata schema is being used as the basis for resource description in this project. However, as it currently has a limited capacity to describe the target audiences for the resource, the project will extend and customise this aspect of AGLS to enable the description of resources in ways that are more sensitive to the needs of target audiences or classes of individual users. New metadata elements will be defined with reference to the user profiles and patterns of use emerging from Phase 1. The AGLS may also need to be extended to deal with the requirements relating to reliability and quality. Schema analysis, metadata modelling, concept mapping of metadata schema and standards, and empirical instantiation (which populates models with examples) will then be used to develop the semantic structure for the metadata repository (McKemmish 1999; http://www.sims.monash.edu.au/research/rcrg/).

New methods will be developed to translate the outcomes of the user needs analysis, user profiling and resource assessment into specifications for metadata schema requirements for user-sensitive resource descriptions. The metadata repository will be designed, built and populated with resource descriptions as a component of the intelligent portal.

DESIGNING AND PROTOTYPING AN INTELLIGENT PORTAL

In Phases 4, 5 and 6 the project will develop specifications for, prototype and evaluate an intelligent portal designed to:

- enable the user to control the flow and 'content' (hence relevance) of the information accessed, e.g. if 'scientific medical articles' which deal with treatments relevant to a particular stage of the disease are desired, then the portal will provide access to this type of information resource only;
- offer the user alternatives as to format, e.g. aural, visual, textual by matching descriptors and profiles, so that the user may decide to refer to visual material while reading related text (as in www.dipex.org);
• present a 'hierarchy' of subject matter which runs simultaneously along user profiles and descriptors, i.e. the continuum of disease may be the basis, but then this is overlaid with specific user profile material;
• optimise precision rather than recall;
• assist the user in making judgments about the quality and authoritativeness of the information resource based on information in the related metadata record.

Consideration will also be given to providing
• 'deconstructed' print materials, allowing users to focus on specific subject domains rather than an entire publication (with the permission of the copyright owners);
• information about support groups – it is envisaged that these groups will see the BCKOnline portal as a key avenue for the dissemination of information about their activities;
• provide a channel through integrated bulletin-boards and ‘chat-rooms’ for communication which enables women with shared interests to transcend the physical barriers of time and place.

CONCLUSION

I think somehow your pathway is a bit self-directed in the sense that... I don’t think any of us would have sought identical information because we all do it in different ways and our personalities handle it in different ways and we progress at a different speed with the level of information that we want (Interviewee 8)

The above comment encapsulates the vision and the challenge of the BCKOnline research project. In order for the portal to assist the consumer in critical decision making, to empower her to participate in her own health care, or to provide avenues for psycho-social support and communication, the format, content, and delivery mechanisms must, as far as possible, recognize the diversity and dynamic nature of both the information seeking process itself and the individual predilections for the presentation of that information. Moreover, given the proliferation of health sites and the sheer volume of material available, a system which would clearly and simply inform the user as to the ‘quality’, currency and authoritativeness of the retrieved information, would obviate many of the concerns listed in the introduction to this paper. As a ‘one-stop-shop’, the BCKOnline intelligent portal has the potential to provide a ‘window’ onto ‘tiered’ information content and sources which take into account patterns of information use derived from the user-profiling and descriptive metadata tagging of Phases 1-3, as well as the benefits of ‘value-added’ components derived from existing best practice. In place of a deluge of material, much of which is of dubious quality and/or originating from spurious sources, the structure and design of the envisaged portal seeks to meet the primary medical, psycho-social and decision support needs of the women with breast cancer by directing her to sources or materials which are relevant, timely, accurate, and ‘personalised’.
REFERENCES


