Published in final edited form as:

J Health Commun. 2013 August ; 18(8): 974–990. doi:10.1080/10810730.2012.757395.

Iterative Evaluation of a Web-Based Health Information Resource

LINDSAY ROSENFELD,

Brandeis University, The Heller School for Social Policy and Management, The Institute for Child, Youth, and Family Policy, Waltham, Massachusetts, USA

AMY SHEPHERD,

Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA

AMENAH A. AGUNWAMBA, and

Department of Society, Human Development, and Health, Harvard School of Public Health, Boston, Massachusetts, USA

ALEXA T. McCRAY

Center for Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA

Abstract

This paper presents the research process and methods used to evaluate and improve a web-based health information resource intended for the public. The resource is called Community Connect to Research (CC2R) (www.connecttoresearch.org). The research process was an iterative one that involved collaboration with many partners. Two formal evaluations were conducted in 2009 and 2010 using key informant interviews, usability interviews, focus groups, an online survey, and readability and suitability assessment tools. These methods provided us with users' perspective on the overall design, content, and literacy demands of the web site as well as valuable feedback on their interaction with the web site. We subsequently redesigned CC2R, making significant improvements based on what we learned from the evaluation. The second evaluation revealed that the redesign addressed many issues found in the first evaluation and identified additional areas of possible improvement. Overall, both evaluations suggested that participants believed that the web site was useful and valuable, indicating that CC2R is indeed a health information resource that provides patients and families with accessible, relevant, and high-quality information. Our experience suggests that regular formal evaluation is an essential tool for effective ongoing enhancement of health information resources meant for the public.

Community Connect to Research (CC2R), an online health resource, was first developed as a project of the Community Engagement Program of the Harvard Catalyst Clinical Translational Science Center (CTSC) (www.catalyst.harvard.edu). Since then, it has evolved into a partnership, called the Collaborative for Community Engagement and Research, among all of the community engagement programs of the four Clinical and Translational Science groups in Massachusetts: Boston University, Harvard University, Tufts University, and the University of Massachusetts. Experts in health, medicine, public health, community-engaged research, and health informatics have all been involved in developing and improving CC2R. The resource is novel in bringing together in one place, local, targeted, scientifically sound, and easily understandable information about the clinical research enterprise. It provides the public with information about a wide range of health topics,

clinical research, and community-engaged research. We have worked to make the resource accessible by keeping our primary audience -patients and their families - in mind at all stages of development, and we have been dedicated to making continual improvements to the resource such that it best serves our intended audience. (Koch-Weser, et. al., 2010; Bernhardt, et. al., 2002; Kaphingst, et. al., 2010; Goldberg, et. al., 2011)

CC2R provides clear information to the public on what health research is and what it means to be a health research participant. CC2R is regularly updated with new content, and all research trials displayed on the site are those currently being conducted in Massachusetts. Trials are updated daily through a search of ClinicalTrials.gov, <http://clinicaltrials.gov> (McCray & Ide, 2000) Importantly, like ClinicalTrials.gov, CC2R does not recruit participants; it is a resource where the public can access expert health information, learn about health research, and search for available Massachusetts clinical trials.

In this paper, we present the methods we used to evaluate Community Connect to Research. The web site was made available to the public in the spring of 2009, and we conducted the first evaluation shortly thereafter, in the summer of 2009. Paying close attention to the findings of this evaluation, we made significant improvements to the resource over approximately one year. To assess whether the changes we made were effective, we conducted a second evaluation in the late summer of 2010. Our experience suggests that regular formal evaluation is an essential tool for effective ongoing enhancement of health information resources meant for the public.

A variety of metrics has previously been used to evaluate online health information resources. These include usability interviews (Choi & Bakken, 2006; Fleisher, et al, 2008; Hinchliffe & Mummery, 2008; Moore, Bias, Prentice, Fletcher, & Vaughn, 2009; Street, et al, 2007), surveys and questionnaires (Fleisher, et al., 2008; Madle, Berger, Cognat, Menna, & Kostkova, 2009; Street et al., 2007), focus groups (Kukafka, Khan, McFarlane, Ancker, & Cohall, 2007; Mackert, Kahlor, Tyler, & Gustafson, 2009; Street et al., 2007), key informant interviews (Street, et al., 2007), and assessment of the literacy demands and readability levels of online content (Bluman, Foley, & Chiodo, 2009; Friedman, & Hoffman-Goetz, 2006). Such evaluation methods have revealed important issues: inconsistent, incorrect or outdated content, high literacy demands of the presented content, confusion about the intended audience, lack of interactivity problems with layout and navigation, and overall design flaws.

In this paper, we will describe the methods we used to evaluate CC2R. We also present our results and subsequently discuss what we learned by conducting both evaluations. We assess the evaluations from 2009 and 2010, and, in particular, examine the impact of the changes we made from the 2009 to the 2010 evaluation. We also discuss certain limitations of our approach, and, finally, we discuss the implications of our research, not only for our own resource, but also for designing and evaluating health web sites in general.

Methods

CC2R Evaluation

The CC2R project and resulting evaluations have been, and will continue to be, an iterative process (Figure 1). The project has been highly collaborative since its inception, working towards the National Institutes of Health Clinical and Translational Science goal of engaging communities in clinical research efforts. Specifically, we strive to overcome traditional barriers to translational research –such as a lack of suitable information and inappropriate health literacy¹ demands—by involving many stakeholders and engaging collaborators across disciplines and institutions (Berglund & Tarantal, 2009). During the

creation of CC2R, we elicited expert guided development and input from our collaborators. Our initial collaborators included Harvard Health Publications (HHP) (authors of the “Health Topics” and “Health Articles” sections of CC2R) and the Institute for Community Health (ICH) (authors of the “Community-Engaged Research Projects”). Later, collaborators included colleagues from all four Massachusetts CTSCs. This is particularly important, since CC2R is designed to serve the citizens of Massachusetts. Immediately following the launch of CC2R, we conducted our first formal evaluation. Based on the results of the evaluation, we redesigned and subsequently relaunched the web site.

The 2009 and 2010 CC2R evaluation processes provided us with the user’s perspective -via demonstrated interactions with the site and described perceptions of the content. We also assessed the site’s complexity - using standard measures to assess the literacy demands placed on users. The goal in conducting these assessments was to improve the quality and accessibility of the site.

We used the following qualitative and quantitative methods² to assess the health literacy demands of the CC2R web site. The evaluation comprised four major components:

1. Key informant interviews with professionals who work in health communication;
2. Usability testing with members of the general public;
3. Focus groups with members of the general public;
4. Literacy demand assessments: SAM (Suitability Assessment of Materials) (Doak, Doak, & Root, 1996), which included SMOG (Simple Measure of Gobbledygook for readability) (McLaughlin, 1969).

In addition, for the 2010 evaluation we also conducted an online survey.

Key informant interviews—We conducted key informant interviews with experts in public communication in their offices, or at Harvard Medical School’s Center for Biomedical Informatics. Participants were shown the main sections of the CC2R web site and asked to comment from their perspective, as someone who develops health-related materials for the public. Questions about layout, information presented, and purpose were asked. After each section was viewed the interviewer asked participants for any recommendations for improvement.

Usability interviews—Usability interviews assessed participants’ ability to navigate through the web site successfully (e.g., participants were asked to click through a series of links and to discuss their response to the web site layout), and provided an opportunity for the participants to give feedback on visual elements of the web site. Usability interviews were held at the ECHO (Enhancing Communications for Health Outcomes) lab of the Dana Farber Cancer Institute, which is a facility designed to observe and understand how participants navigate web sites and react to messages. One CC2R project staff member posed a series of questions and asked participants to complete a series of tasks while the other staff member, located behind a two-way mirror looking into the room where the interviews took place, recorded their comments.

¹The US Department of Health and Human Services Healthy People 2010 defines health literacy as “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” (US Department of Human Services Healthy People 2010, 2000).

²The Harvard School of Public Health: Health Literacy Studies Web Site (<http://www.hsph.harvard.edu/healthliteracy>) and the U.S. Department of Health and Human Services’ (HHS) Research-Based Web Design and Usability Guidelines (http://www.usability.gov/guidelines/guidelines_book.pdf) were resources utilized by project staff to assess the CC2R website.

Focus groups—All focus group sessions were conducted at the Harvard Medical School Center for Biomedical Informatics. Each session was led by a CC2R project staff moderator, and was digitally recorded by another staff member. During each 1.5 hour session, the moderator showed the group the CC2R web site and asked questions about content, layout, and design. Participants were encouraged to give their honest opinions to help improve the content and design of the web site.

Online survey (2010 evaluation only)—We conducted the online survey during the 2010 evaluation to provide insight on how users, exploring the web site on their own, located and understood the CC2R content. The survey was meant to determine the acceptability and applicability of content and how the content might affect self-efficacy.³ These factors were established by asking participants to complete a series of tasks (on the web site), and asking questions regarding their opinions and experience. Modeled after the Office of Disease Prevention and Health Promotion “Remote Clickable” studies (US Department of Health and Human Services, 2010), the survey mimicked a “real life” exposure and experience with the web site, since participants were asked to review the web site independently of a moderator. This method is intended to reduce bias that may result from moderator-participant expectations and interactions.

To measure the health literacy demands on users, we used a validated assessment tool called the Suitability Assessment of Materials (SAM) (Carden, Jefford, & Rosenthal, 2007). SAM provides a numerical score for health-related materials to help determine literacy demand (Doak, Doak, & Root, 1996). SAM is calculated by applying a set of criteria to the material and then scoring the material based on attributes that fit the SAM criteria; the results are a “suitability score.” Health material is categorized by the score it receives as *superior* (70–100% score), *adequate* (40–69% score), or *not suitable* (0–39% score). We used SAM to determine how appropriate our material was for the intended audience; this allowed us to examine the layout and purpose as well as whether the cultural and linguistic needs of diverse populations were met (Neuhauser, & Kreps, 2008).

Included in the SAM assessment is the readability level. This is an important part of literacy demand given that ninth grade or above is not suitable for the average abilities of a United States citizen (National Assessment of Adult Literacy, 2003). To assess readability, we used the Simple Measure of Gobbledygook (SMOG), a validated measure (Bailin & Grafstein, 2001; Friedman, & Hoffman-Goetz, 2006) determined to deliver superior accuracy when compared with other readability assessments (Friedman & Hoffman-Goetz, 2006). SMOG finds the estimated level of schooling needed to read and understand a piece of writing (readability level) which is calculated through a formula that analyzes sentence length and the number of polysyllabic words (McLaughlin, 1969). We determined SMOG categories, as suggested by literacy experts (Doak, Doak, & Root, 1996): *superior*: 5th grade level or lower (5 years of schooling), *adequate*: 6th, 7th, or 8th grade level (6–8 years schooling) or, *not suitable*: 9th grade level and above (9 years or more of schooling).

Our goal was to recruit at least five participants for each set of key informant and usability interviews, as well as for the focus groups. It has been shown that at least that number of key informants is needed for reliability in the responses received (Seidler, 1974). Similarly, usability testing with five users reveals 85% of usability problems while typically producing repetitive information after that (Nielsen & Landauer, 2000). Participants were recruited

³Acceptability is the extent to which the intended users like the content and find it easy to use. Applicability is the extent to which the content is relevant to the needs of the intended user (i.e., whether it helps the user in everyday life). Self-efficacy measures participants’ level of confidence in a health related action, given the information provided on the web site. (US Department of Health and Human Services, 2010)

only if they were 18 years of age or older and had experience with computers. For both evaluations, participants from the Greater Boston area were recruited for usability interviews and focus group discussions via flyers posted at YMCAs, community centers, community health centers, and libraries as well as through craigslist.com. For the online survey, participants were recruited exclusively through craigslist.com.

Key informants were all experts in health literacy and public communication and were referred to study staff with help from our collaborators. During the intake process, all potential participants were given a written and oral description of the study. While we employed a convenience sample, the participant pool (aside from key informants) was chosen from all eligible participants with particular attention to recruiting a diverse sample with regard to race and ethnicity, education level, and gender (Table 1). At the time of participation, all participants completed a demographic survey, provided their contact and demographic information, and signed a Harvard Medical School Institutional Review Board approved informed consent. All sessions (excluding the online survey) took approximately 1.5 hours. Usability and focus group participants received a \$25 Stop & Shop gift card for participating. Online survey participants received a \$5 Dunkin Donuts gift card. Key informants received no compensation.

All focus groups and usability interviews were audio-recorded and then transcribed. During key informant interviews, project staff took detailed notes but did not audio record the conversations. Usability participants' actions, particularly what links were clicked on when prompted, were also visually recorded. Two project staff separately identified key information from the transcriptions and notes; thematic coding was formulated with a consensus coding process. After coding, we analyzed the data to find themes and to calculate descriptive statistics.

To ensure reliability between the two independent reviewers, two different, trained project staff assessed a sample of the pages being assessed by SAM and SMOG. The scores produced by project staff and the independent reviewers were not statistically significantly different; a score difference never resulted in assigning different SAM and SMOG categories to assessed pages.

Results

2009 and 2010 CC2R Evaluation Overview

The results of the 2009 evaluation suggested a series of needed changes for the CC2R web site. Participants in 2009 noted that the web site is a useful contribution for linking community members to clinical research. Most said that they would return to the web site to find health information, information on medical research, and to learn how communities participate in research. Many said that they would return for clinical trials information. Overall, participants liked the information presented, though they indicated several areas of possible improvement: content, audience, structure, navigation, and graphics, and health literacy demands (Table 2). In Table 2, the details for suggestions and recommendations for improvement from the 2009 evaluation are first listed. Then, changes implemented resulting from the 2009 evaluation are displayed. This same information is also provided for the 2010 evaluation. For example, 2009 suggestions to improve content included 1) making the purpose of CC2R clearer and 2) improving cohesion between the web site sections. In the 2010 evaluation, content was not identified as a theme in need of improvement. Prior to conducting the 2010 evaluation, the project team worked together to implement all of the suggested changes raised in the 2009 evaluation.

The 2010 evaluation provided us with information on how successful we were in implementing the 2009 suggested changes, together with valuable information on how to further improve the CC2R site. In 2010, the key informant, usability interview, and focus group participants all said that they would return to the web site to find health information, information on medical research, and to learn how communities participate in research. All participants moved through the web site with ease and most understood the purpose and target audience. The areas in need of improvement in 2010 included: structure and navigation, and readability and health literacy demands. The issues needing improvement regarding structure and navigation were different in 2010 from those noted in 2009. For example, in 2009 participants wanted the search function expanded so it could search the entire web site. We subsequently implemented this feature, and it was used extensively by 2010 evaluation participants. In 2010, participants wanted features that allowed them to get a preview of sections before clicking on them. Content, audience, and graphics were not noted as areas for improvement in the 2010 evaluation.

Within each evaluation (2009 and 2010), themes were consistent across key informant, usability interview, and focus group participants. However, as noted, there were different themes between the 2009 and 2010 evaluations. In particular, in 2009, some focus group participants were uncomfortable with clinical trials and expressed that clinical trials made them suspicious of the purpose of the website. We put forth a great deal of effort to clarify what clinical trials are by including additional historical information and a further discussion of risks and benefits. In 2010, this uneasiness was no longer voiced as an issue by any study participants.

Readability and Suitability of Materials

Table 3 shows an overview of the 2009 and 2010 readability and suitability of materials assessments. The web site sections that were assessed are listed, along with grade levels (SMOG), suitability scores (SAM), the number of pages assessed, and the number of pages that received unsuitable ratings.

SMOG—In 2009, seventy-eight pages (of 120: 65%) were randomly chosen and assessed; the average grade level score across these pages was 13.46. None of these pages scored 9th grade or below. In the 2010 evaluation, ninety-two pages (of 150: 61%) were assessed; the average grade level across these pages was 10.62. Nine pages scored as 9th grade or below. The “Introduction to Health Resources” section was not assessed in the 2009 evaluation because the section consisted of only one landing page with links to external resources. We completely restructured this section following the 2009 evaluation to include more content, and it now has 15 landing pages. Eighteen “Health Topic Introductions” were assessed in the 2009 evaluation. This section was not evaluated in the 2010 evaluation because our collaborators had made no changes.

SAM—Twenty-nine pages (of 120: 24%) were randomly chosen and assessed with SAM in the 2009 evaluation; eleven received an “*unsuitable*” rating with an average overall score of 44.52%. Forty-five pages (of 150: 30%) were assessed in the 2010 evaluation; all pages received a “*suitable*” or above rating with an overall of 68.8%.

Online survey (2010 only)

For self-efficacy, we found that approximately 33% percent of participants “strongly agreed” and 49% “agreed,” (3% “disagreed”) that the information made them more confident to do something about their health. Approximately 28% percent of the participants “strongly agreed,” and 62% “agreed” (no one disagreed) that the information made them more prepared to do something about their health. Approximately, 23% of the participants

“strongly agreed” and 36% “agreed,” (10% “disagreed”) that the information made them feel more prepared to do something within the next month. For acceptability, 41% percent of the users “strongly agreed” and almost 54% “agreed” (5% “disagreed”) that the information on CC2R was easy to use. Almost 44% of the participants “strongly agreed” and 54% “agreed” (no one “disagreed”) that the information was presented in ways that the participant could easily understand. For applicability, about 31% of the users “strongly agreed”, and 51% “agreed” (5% “disagreed” or “strongly disagreed”) that the information was useful. About 26% of the users strongly agreed, and 59% “agreed” (2.6% “disagreed”) that the information provided gave them specific ideas about what to do for their health and those around them. (Table 4)

Discussion

The two-step evaluation practice we employed – of examining literacy demands and talking with end users – has become critical for evaluating and improving web sites. Studies that employ this method report that it is an integral part of designing effective web sites (Bluman, Foley, & Chiodo, 2009; Choi & Bakken, 2006; Street et al., 2007). Therefore, to design and implement the CC2R evaluations, we followed the best available practices for assessing web sites (Choi & Bakken, 2006; Fleisher et al., 2008; Street et al., 2007). The first evaluation of CC2R occurred after the initial April 2009 launch. Our objective was to explore whether the site was performing as planned –and we intended from the outset to use the gathered data in a redesign effort. Upon analyzing the data and applying our findings to the redesign, we performed another evaluation to examine if the redesign appropriately addressed the issues raised during the first evaluation. We found that they generally did address the issues. The additional feedback we received (2010 evaluation) on the re-developed site will be incorporated into future changes and improvements.

Both evaluations provided valuable feedback on how users interact with CC2R. Our goals were to examine whether CC2R was accessible to our target audience, with regard to health literacy demands, usability, and usefulness. Using the results of the first evaluation, the project team redesigned the web site to improve the site’s content and accessibility, and created more plain language text and graphics to lower the health literacy demands. We have now begun to implement further changes, using the results of the 2010 evaluation. Both of the evaluations suggested that participants felt that the web site was a useful and valuable resource. The evaluations also provided a user voice for many important suggestions for improvement.

In the 2009 evaluation, participants made suggestions in each of the following areas: a) content, b) audience, c) structure, navigation, and graphics, and d) health literacy demands. We addressed each of these areas respectively: a) we restructured and redesigned the site to better reflect the purpose of the site and to create greater cohesion across the site; b) we added a mission statement to the CC2R homepage, including a statement about the intended audience for the web site; c) we redesigned and renamed sections as needed, expanded the search function, and added relevant graphics; and d) we created more plain language on the site, began development of an interactive glossary, including more in-text definitions, added informative graphics and visual aids, made the purpose of pages more explicit, increased interactive features, and added culturally specific elements where needed (Table 2).

The positive impact of these changes was evident in the second evaluation. Participants did not raise any of the same concerns in 2010, and their comments were more tightly focused. We have already addressed the majority of these new suggestions and are well on our way to addressing others, including adding an extensive glossary to ease the demands of medical terminology.

We saw clear improvements in the SMOG and SAM assessments. Although SMOG readability levels are still too high, they did improve between evaluations. For example, the average SMOG readability level for the “Health Articles” section was 13.4 in the first evaluation. In the 2010 evaluation, the average SMOG readability level was 11.9. This is a difference of 1.5 grade levels, indicating that while “Health Articles” were previously accessible to those who had received an education equivalent to that of the first year of post-high school (e.g., college), the “Health Articles” are now accessible to those who have received an education equivalent to that of high school. While this is still above the average abilities of a United States citizen (National Assessment of Adult Literacy, 2003), the readability level of CC2R “Health Articles” is becoming more accessible. The average SAM suitability score of the “Health Articles” was 47.3% (lower end of the “adequate” category) in 2009, while in the 2010 evaluation the average SAM score of the “Health Articles” was 68% (the upward range of the “adequate” category). While further progress can surely be made, this is a clear step forward (Table 3).

In all, we found SAM to be a much more comprehensive assessment than SMOG alone, of how appropriate the site or material is for the end-user. Since SAM incorporates a readability assessment (for which we chose SMOG), SAM can determine readability and suitability in one assessment. Before talking with key informants and the public, researchers and practitioners can assess the organization, writing style, appearance, and appeal of a web site using SAM. Notably, we should employ the SMOG or any other readability assessment with caution, because reducing readability level only without regard to context can result in important information being lost. For example, it may or may not be appropriate to eliminate a technical term that is consistently repeated (e.g., nephrology), only to reduce the literacy level of a document; it may be that if the term is defined (e.g., nephrology is the study of the kidney), it could be useful to have the term included so that readers become familiar with it and are able to use it appropriately. Therefore, being cognizant of the need to retain quality and accuracy is as important as the need to maintain appropriate literacy levels. A useful solution is incorporating plain language wherever possible, including defining terms and providing an easy-to-use glossary. Key informant and usability interviews, focus groups, and an online survey additionally provided essential information that cannot be garnered from the SMOG and SAM alone.

The online survey, fielded only during the 2010 evaluation, revealed that such a format provides useful information for gauging public impressions of the site, which can help us efficiently and effectively create continuing improvements. The survey additionally explored new issues that were not included in the previous evaluations, such as self-efficacy, acceptability and applicability. After visiting the site, a high percentage of users strongly agreed or agreed that they were more confident to do something about their health, more prepared to do something about their health, and more prepared to do something within the next month. Acceptability was rated highly, with most users saying the information was easy-to-use and that the information was presented in ways that were easily understood. Applicability was also highly rated with a high percentage of users saying the information was useful and that the information gave them specific ideas about what to do for their health, and that of those around them. This survey helps us better understand what users find useful about the site – and provides us with additional ways to continue to improve the site. Additionally, this feedback suggests that CC2R is fulfilling its purpose of being a trusted communication tool.

It is not enough that the online survey reported “high marks” for the site, and that we have made important gains with each of our improvements. We expect each page to represent readability scores that match the skills of the average U.S. citizen (~8th grade level) (Kirsch, Jungeblut, Jenkins, & Kolstad, 1993; National Assessment of Adult Literacy, 2003).

Likewise, we expect each page to attain a “*superior*” SAM score. We plan to further develop the online survey that was administered in 2010 to help us understand more about what users do and do not understand or like about the site. Our development team, as well as our community and university partners, will continue to work together to improve CC2R, continuing rigorous evaluation to produce the best product for the public.

Limitations

Despite the important information garnered from these evaluations, we must recognize that these results were obtained from a small, cross-sectional sample through mainly qualitative measures, and therefore we may not understand the “whole picture”. Even though guidelines for usability suggest 5–8 individuals, given the diversity of the participants, the numbers are small. We also cannot examine whether this web-based health communication tool had a direct effect on people’s health outcomes. Last, as this is a collaborative process, we are not always able to control the pace of change or the complete adoption of evaluation suggestions; however, we will continue to evaluate *Community Connect to Research*, working with our partners to deliver a high quality health resource to the public.

Practice Implications

The CC2R web site is a useful resource with many dynamic parts, resulting from a truly interdisciplinary collaborative process. Such iterative evaluations can be relatively straightforward to conduct. In addition, this work with CC2R reinforced several points concerning web site development: 1) evaluation is important to conduct when launching a new site to understand what the intended audience does or does not prefer or understand, 2) the provided suggestions, in consultation with experts, should be implemented, 3) a subsequent evaluation should be performed to gauge whether the implemented suggestions were effective in addressing the issues proposed, and 4) continued evaluations should be performed and analyzed, and used to re-evaluate and improve the site.

Web sites should be accessible and acceptable to the target population; a “universal precautions” and “plain language” approach are important. Universal precautions were incorporated from the start of the project (Health Literacy Universal Precautions Toolkit, 2010). In particular, we aimed to use “plain language” throughout the site, defined as “communication your audience can understand the first time they read or hear it (Plain Language, <http://www.plainlanguage.gov/>).” Successfully doing this means your audience can find what they need, understand what they find, and use what they find to meet their needs. Accomplishing this is not an easy task, and we are constantly learning and improving on these objectives. We have found several techniques to help us accomplish these goals, including the evaluation work described in this manuscript.

To accomplish the tenets of “universal precautions” and “plain language,” we found that it was essential to examine web site components with both experts and the public. Even small evaluations that incorporate both user feedback and assessments can provide important insight into a web site’s design, usability, and literacy demand. In addition to readability, we found suitability to be an essential issue to assess. While readability is crucial for accomplishing “plain language,” suitability is also an important objective (making sure the content, layout, and graphics are appropriate); taking “universal precautions” means that our goal is to accomplish both readability and suitability. Many easy-to-use tools allow practitioners and researchers to assess readability and suitability; we found that the SMOG online tool and the SAM checklist were useful formats of such measures, respectively. Additional suggested methods, which we also used, include digital recording technology, and providing standard interview questions for key informants, usability interviewees, and focus groups. While we used a lab especially equipped to handle usability interviews – it is

not necessary to have a dedicated set-up. Usability interviews can be conducted without visual recording of mouse movements, for example. Following prepared interview guides and simply audio recording the interviews conducted can help assure that no information is missed. If audio recording is not possible two interviewers can observe and capture information.

Comprehensive assessment of web sites, including their literacy demands, is a crucial part of web site development for those practitioners and researchers who want their web sites to be accessible and understood by their intended audience.

References

- Bailin A, Grafstein A. The linguistic assumptions underlying readability formulae: A critique. *Language Community*. 2001; 21:285–301.
- Berglund LM, Tarantal A. Strategies for innovation and interdisciplinary translational research: Removal of barriers through the CTSA mechanism. *Journal of Investigative Medicine*. 2009; 57(2): 474–476. [PubMed: 19174699]
- Bernhardt JM, Lariscy RAW, Parrott RL, Silk KJ, Felter EM. Perceived barriers to Internet-based health communication on human genetics. *Journal of Health Communication*. 2002; 7(4):325–340. [PubMed: 12356290]
- Bluman EM, Foley RP, Chiodo CP. Readability of the patient education section of the AOFAS web site. *Journal of Foot and Ankle Surgery*. 2009; 30(4):287–291.
- Carden CP, Jefford M, Rosenthal MA. Information about cancer clinical trials: An analysis of Internet resources. *European Journal of Cancer*. 2007; 43(10):1574–1580. [PubMed: 17544266]
- Choi J, Bakken S. Heuristic evaluation of a web-based educational resource for low literacy NICU parents. *Studies in Health Technology and Informatics: Consumer-Centered Computer-Supported Care for Healthy People -Proceedings of NI2006*. 2006; 122:194–199.
- Doak, CC.; Doak, LG.; Root, JH. *Teaching patients with low literacy skills*. 2. J.B. Lippincott; Philadelphia: 1996.
- Fleisher L, Buzaglo J, Collins M, Millard J, Miller SM, Egleston BL, et al. Using health communication best practices to develop a web-based provider-patient communication aid: The CONNECT(TM) study. *Patient Education and Counseling*. 2008; 71(3):378–387. [PubMed: 18417312]
- Friedman DB, Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education Behavior*. 2006; 33(3):352–373. [PubMed: 16699125]
- Goldberg L, Lide B, Lowry S, Massett HA, O'Connell T, Preece J, Quesenberry W, Shneiderman B. Usability and accessibility in consumer health informatics current trends and future challenges. *American Journal of Preventive Medicine*. 2011; 40(5 Suppl 2):S187–S197. [PubMed: 21521594]
- Health Literacy Universal Precautions Toolkit. AHRQ Publication No 10–0046-EF. Rockville, MD: Agency for Healthcare Research and Quality; Apr. 2010 Available at: <http://www.ahrq.gov/qual/literacy/index.html>
- Hinchliffe A, Mummery K. Applying usability testing techniques to improve a health promotion web site. *Health Promotion Journal of Australia*. 2008; 19(1):29–35. [PubMed: 18481929]
- Huang, GJ.; Penson, DF. *Cancer Investigation*. Taylor & Francis Ltd; 2008. Internet health resources and the cancer patient.
- Kaphingst KA, McBride CM, Wade C, Alford SH, Brody LC, Baxevanis AD. Consumers' use of Web-Based information and their decisions about multiplex genetic susceptibility testing. *Journal of Medical Internet Research*. 2010; 12(3):e41. [PubMed: 20884465]
- Kirsch, IS.; Jungeblut, A.; Jenkins, L.; Kolstad, A. *Adult literacy in America: National Center for Education Statistics*. U. S. Department of Education; Washington, D.C.: 1993.
- Koch-Weser S, Bradshaw YS, Gualtieri L, Gallagher SS. The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. *Journal of Health Communication*. 2010; 15(S3):279–293. [PubMed: 21154099]

- Kukafka, R.; Khan, SA.; Hutchinson, C.; McFarlane, DJ.; Li, J.; Ancker, JS.; Cohall, A. Annual Symposium Proceedings. American Medical Informatics Association; 2007. Digital partnerships for health: Steps to develop a community-specific health portal aimed at promoting health and well-being; p. 428-32.
- Mackert M, Kahlor L, Tyler D, Gustafson J. Designing e-health interventions for low-health-literate culturally diverse parents: Addressing the obesity epidemic. *Telemedicine and Health*. 2009;10.1089/tmj.2009.0012
- Madle G, Berger A, Cognat S, Menna S, Kostkova P. User information seeking behavior: Perceptions and reality. An evaluation of the WHO Labresources Internet portal. *Informatics for Health & Social Care*. 2009; 34(1):30–38. [PubMed: 19306197]
- McCray AT, Ide NC. Design and implementation of a national clinical trials registry. *Journal of the American Medical Informatics Association*. 2000; 7(3):313–323. [PubMed: 10833169]
- McLaughlin GH. SMOG grading: A new reading formula. *Journal of Reading*. 1969; 12:639–640.
- Moore M, Bias R, Prentice K, Fletcher R, Vaughn T. Web usability testing with a Hispanic medically underserved population. *Journal of Medical Library Association*. 2009; 97(2):114–121.
- National Assessment of Adult Literacy. 2003. Available at: <http://nces.ed.gov/>
- Neuhauser L, Kreps GL. Online cancer communication: Meeting the literacy, cultural and linguistic needs of diverse audiences. *Patient Education and Counseling*. 2008; 71(3):365–377. [PubMed: 18424046]
- Nielsen, J.; Landauer, TK. Why you only need to test with 5 users. Fremont (CA): useit.com: Jakob Nielsen's; 2000. Web site. Retrieved from: <http://www.useit.com/alertbox/20000319.html>
- Plain Language: Improving Communication from the Federal Government. Available at: <http://www.plainlanguage.gov/>
- Seidler J. On using informants: A technique for collecting quantitative data and controlling measurement error in organization analysis. *American Sociological Review*. 1974; 39(6):816–831.
- Street A, Swift K, Annells M, Woodruff R, Gliddon T, Oakley A, et al. Developing a web-based information resource for palliative care: an action research inspired approach. *BMC Medical Informatics and Decision Making*. 2007; 7(1):26. [PubMed: 17854509]
- The Harvard School of Public Health: Health Literacy Studies. Web Site. Available at: <http://www.hsph.harvard.edu/healthliteracy>
- US Dept of Health and Human Services, Office of Disease Prevention and Health Promotion. Health literacy online: A guide to writing and designing easy-to-use health web sites. 2010. Available at: <http://www.health.gov/healthliteracyonline/>

FIGURE 1.

CC2R iterative research process: Site development and evaluation. HHP = Harvard Health Publications; ICH = Institute for Community Health; CCER = Collaborative for Community Engagement and Research.

Table 1

Participant Demographics

Variable	Key informant interviews, Usability interviews, and Focus groups		Online Survey
	2009 Evaluation (n=33)	2010 Evaluation (n= 23)	2010 Evaluation (n= 23)
	n=number (%)	n=number (%)	n=number (%)
Median age	43	50	37
Language			
English	0 (90.91)	23 (100)	37 (95)
Other	3 (9.09)	-	2 (5)
Gender			
Male	13 (39.39)	12 (57.89)	17 (43.6)
Female	20 (60.61)	11 (63.16)	22 (56)
Race/Ethnicity			
White	14 (42.2)	11 (57.89)	30(77)
Black/African American	10 (30.30)	7 (30.4)	4(10)
Hispanic/Latino origin	2 (6.06)	3 (13)	10(26)
Asian	5 (15.15)	3 (13)	3(8)
Other	2 (6.06)	1 (4.3)	3(8)
Household Income			
\$49,000 or less	17 (51.52)	14 (73.68)	26 (45)
\$50,000 or more	13 (39.39)	5 (26.32)	13 (33)
Chose not to answer	3 (9.09)	-	6(15)
Education			
High school graduate	1 (3.03)	2 (10.53)	2(5)
Some college education	7 (21.21)	8 (42.11)	5(13)
College grad/some postgrad education	19 (57.57)	8 (42.11)	26(67)
Other	6 (18.18)	5 (26.31)	5(13)
Employment			
Employed	24 (72.73)	8 (42.11)	22 (56)
Other	9 (27.27)	15 (78.95)	18 (42)
Self-rated health			
Poor/Fair	1 (3.03)	8 (42.1)	3(7.6)
Good/Very good/Excellent	32 (96.96)	15 (57.9)	36 (92)
Use computer for			
Send emails	33 (100)	11 (57.89)	39 (100)
Search for information	33 (100)	23 (100)	39(100)
Pay bills	20 (60.61)	17 (73.9)	29(74)
Social networking	19 (57.58)	16 (69.6)	31(80)
Read/Watch news	27 (81.82)	19 (82.6)	35(90)
Use the Internet to find health information			
Never	3 (9.09)	3 (9.1)	1(2.6)
Once a month	7 (21.12)	5 (21.2)	14(36)

Variable	Key informant interviews, Usability interviews, and Focus groups		Online Survey
	2009 Evaluation (n=33)	2010 Evaluation (n= 23)	2010 Evaluation (n= 23)
	<i>n=number (%)</i>	<i>n=number (%)</i>	<i>n=number (%)</i>
Once a week	4 (12.12)	5 (21.1)	11(28)
Couple times a week	15 (45.45)	6 (45.5)	26(10)
Everyday	4 (12.12)	3 (21.1)	3(8)

Table 2

2009 and 2010 CC2R Evaluation Overview

Themes	2009 Evaluation Details	Changes Implemented	2010 Evaluation Details	
			<i>Issues Addressed</i>	<i>New Topics</i>
Content	Participants indicated the purpose and cohesion between each section needed to be improved	We restructured and redesigned the web site to better reflect the purpose and create greater cohesion	Participants demonstrated an understanding of the purpose in oral discussions	Content was not identified as a theme in need of improvement
Audience	Participants expressed some confusion about the intended audience	We added a mission statement to the homepage to make the intended audience clearer	When asked, participants correctly identified the intended audience	Audience was not identified as a theme in need of improvement
Structure, Navigation, and Graphics	Structure: Through participant interaction, several pages/sections were identified in need of restructuring	We redesigned both the Homepage and the Research Centers section; the Research Centers section was renamed "Health Resources"	Participants easily found and understood information when asked	The sections redesigned were not identified as in need of improvement
	Navigation: Participants wanted to use the search function (which was limited to searching clinical trials) to search the entire web site	We expanded the search function to include the entire web site	Participants' expectations were met	The search function was not identified as a theme in need of improvement
	Graphics: Participants expressed that many pages lacked engaging, descriptive, interactive graphics, and explanatory captions	We added more photos/diagrams/graphs on the web site as well as captions	Graphics elicited positive feedback	Graphics were not identified as a theme in need of improvement
Health Literacy Demands	Terminology: Some language needed simplifying; many words were considered technical and/or missing definitions	We replaced many words that could be simplified and a plain language glossary is under development	Some words were still found to be technical and/or missing definitions	We will implement a plain language glossary developed by Harvard Health Publications
	Readability/SMOG: The readability levels were higher than recommended (Table 2)	When possible, text was re-written and definitions were provided within the text	Readability levels were lowered, but some text remains higher than recommended	We will continue to strive towards a lower literacy level
	Suitability/SAM: Some SAM scores were lower than recommended (Table 2)	We made improvements in all SAM categories that had scored low	No pages that were assessed received an "unsuitable" score	No changes necessary

Table 3

Readability and Suitability of Materials

Web site Sections*	SMOG***					SAM***						
	Average grade level		# of pages assessed		# of pages that were grade level	Average Score (%)		# of pages assessed		# of pages receiving "Unsuitable" rating		
	2009	2010	2009	2010		2009	2010	2009	2010	2009	2010	
Understanding Health Research	10.8	10.5	4	8	4	7	43.3	76	4	8	1	0
Health Research in MA	15.6	12.3	5	13	5	13	46.6	66.5	5	13	1	0
Introductions to Health Resources	-	8.7	-	14	-	4	-	65	-	7	-	0
Home and About pages	14.8	9.7	2	2	2	2	55	68.9	2	2	1	0
Health Topic Introductions**	12.5	-	18	-	18	-	30.1	-	5	-	5	-
Health Topic Articles	13.6	11.9	49	55	49	55	47.6	68	13	15	3	0
Average (Total)	13.46	10.62	(78)	(92)	(78)	(81)	44.52	68.88	(29)	(45)	(11)	(0)

Note.

* In the 2010 evaluation, the website sections were redesigned to reflect changes that were determined necessary in the 2009 evaluation. In the 2009 evaluation, Understanding Health Research was called Understanding Research, Health Research in MA was called Community Participation, Health Resources was called Research Centers.

** Health Topic Introductions were not assessed in the 2010 evaluation, as there were no changes made to the pages by our collaborators.

*** SMOG categories: *superior*: 5th grade level or lower (5 years of schooling level), *adequate*: 6th, 7th, or 8th grade level (6–8 years schooling) or, *not suitable*: 9th grade level and above (9 years or more of schooling level). SAM categories: *superior* (70–100% score), *adequate* (40–69% score), or *not suitable* (0–39% score).

Table 4

Online survey results: User self-efficacy, acceptability, and applicability (N=39)

Factor	Strongly Agreed	Agreed	Strongly Disagreed	Disagreed	Neutral
	%	%	%	%	%
Self-efficacy					
Information made participants more <i>confident</i> to do something about their health	33.3	48.7	0	2.6	15.4
Information made participants feel more <i>prepared</i> to do something about their health	28.2	61.5	0	0	10.3
Information made participants feel more <i>prepared to do something within the next month</i>	23.1	35.9	0	10.3	30.8
Acceptability					
Information was easy to use	41.0	53.8	0	5.1	0
Information was presented in ways that participants could easily understand	43.6	53.8	0	0	2.6
Applicability					
Information was useful to participants	30.8	51.3	2.6	2.6	12.8
Information gave participants specific ideas about what to do for their health and those around them	25.6	59.0	0	2.6	12.8