Accessibility and Quality of Online Cancer-Related Clinical Trial Information for Naïve Searchers

Gregory A. Abel1,2,3, Angel M. Cronin1, Kristofer Earles4, and Stacy W. Gray1,3,5

Abstract

Although the Internet may help to increase cancer patients’ awareness of clinical trials, little is known about the accessibility and quality of online clinical trial information. We simulated the experience of a naïve cancer patient without clinical trial knowledge by searching three popular search engines for treatment information for breast, lung, and prostate cancer, and myelodysplastic syndromes (MDS). Two coders independently evaluated website content for accessibility and quality. We screened 120 websites and identified 40 unique sites for analysis. Overall, 85% (95% confidence interval [CI], 70%–94%) of sites mentioned clinical trials on the landing page and 68% (51%–81%) included links to specific trials. Overall readability was poor. Approximately half of websites (36%–68%) included information on the potential benefits and risks of clinical trials and 40% provided information about when the site had been updated (25%–57%). Among sites with links to specific clinical trials, only 44% (25%–65%) provided an interactive interface that would allow patients to customize search results; breast (100%) and prostate (50%) sites were more interactive than lung (25%) and MDS (14%; P = 0.007). Although cancer clinical trial information is widely available on the Internet, its quality is highly variable. Given the fact that many emerging cancer therapeutics are personalized based on disease or genomic characteristics, interactive web-based interfaces could serve as powerful vehicles to help patients locate appropriate clinical trials. Without enhanced efforts to ensure greater interactivity of cancer treatment websites, patient awareness of relevant clinical trials may remain low.

Introduction

Although advances in cancer care depend upon the rigorous evaluation of interventions through clinical trials, participation by cancer patients in the United States is known to be extremely low (5% or less; refs. 1, 2). Although some barriers involve oncologists, health systems (3), or patient eligibility (4), another important cause of low enrollment may be lack of clinical trial awareness among cancer patients (5, 6). Indeed, a survey of the general population found that 32% of adults would be “very willing” to participate in an oncologic clinical trial, and an additional 38% would be “inclined” to participate if provided more information (7). Increasing trial awareness and filling in knowledge gaps regarding participation seem like excellent roles for the Internet, especially given the advent of interactive websites (so-called “Web 2.0”). Moreover, there is evidence that cancer patients have increasingly been using the Internet to seek health knowledge (8), and Internet use has been associated with higher levels of trial awareness (9).

A study of web-based clinical trial content from NCI-designated comprehensive cancer centers (n = 39) published a decade ago (10) found that 74% of sites contained a link to clinical trial information on their homepage, but that the type and quality of information varied widely. Most presented data that would be important to researchers but not necessarily of interest to the public; for example, 94% gave protocol ID numbers, but only 38% presented the purpose of each study in straightforward language. Moreover, only 16% presented specific treatment details, readability was low, and while a majority had some search capability, it was often very limited.

A more recent study analyzed the websites of seven prominent international cancer-related institutions, including the NCI (11). Using a variety of tools, the researchers found that “no sites performed well in terms of all assessment criteria,” and every site “required an unacceptably high level of literacy.” In this context, we aimed to explore the accessibility of clinical trial information for naïve users searching for treatments, that is, those not specifically looking for clinical trial information and who would not be familiar with prominent cancer institutions or websites.

Materials and Methods

In January 2013, we simulated the experience of a naïve user without knowledge of clinical trials by searching for “breast cancer treatment”; “lung cancer treatment”; “MDS treatment”; and “prostate cancer treatment” using 3 engines (Google, Bing, and Yahoo). We chose the 4 conditions as examples, aiming to represent a range of malignancies (e.g., common and rare; gender and non-gender specific; more and less curable). The top 10 unsponsored websites for each were analyzed, after eliminating duplicates and news articles. We chose to analyze the top 10 results because in standard search engines such as Google, the first 10 results are...
typically displayed automatically without the searcher having to go to an additional screen. We felt this best simulated the initial efforts of a naïve searcher.

The resulting sites were evaluated independently by two coders for quality using measures adopted from standard guidelines (12) as well as Health On the Net (13) certification (Tables 1 and 2). We also evaluated the accessibility of clinical trial information in two ways. First, we determined the ease with which a naïve user might find relevant clinical trial information (i.e., the presence or absence of links to specific clinical trial information, an interactive interface, and number of clicks from landing page to definition of a clinical trial; ref. 10) as well as the readability of the landing page of the website (i.e., Flesch Reading Ease Score; refs. 14, 15). We operationalized accessibility as including both readability and/or ease of finding relevant information because naïve users can only gain access to information about trials if they are both able to find it and can understand what is written.

Websites were considered to have an interactive interface if there was an option to input patient-specific information into a web search engine to identify clinical trials for which a patient might be eligible. To test this feature, we created a scenario for each disease type with clinical information we felt that a naïve searcher would likely know. For example, for breast cancer, we assumed that the searcher was a 50-year-old female with stage 2, estrogen-receptor–positive disease who had had surgery but no other treatment, and lives in New Haven, CT. Importantly, websites needed to contain interactive content for just one of the clinical details (age, gender, cancer subtype, treatment status, and/or location) to be considered interactive. Finally, these criteria were applied equally when a website was linked to an external clinical trials database, such that if that database were interactive, the original site was also considered interactive.

When coding disputes arose, a coder worked with the two primary coders to resolve the dispute by consensus. Before consensus, interobserver agreement was determined for 10 key items. Fisher's exact tests (for categorical variables) and t tests (for continuous variables) were used to evaluate associations between disease group and website characteristics. After testing for a global difference between the four disease groups, cancers were categorized into two groups for further statistical comparisons [breast and prostate vs. lung and myelodysplastic syndrome (MDS)]; these categories were chosen to reflect disease prevalence, as we felt that naïve searchers would likely have a higher level of baseline health literacy with respect to more common cancers.

## Results

Searches using the same terms yielded highly duplicative results among the three engines, with only 61 unique addresses of the 120 resulting URLs. Another 21 were excluded because they yielded news articles (6) or because they directed to the same home page despite having different web addresses (15). This resulted in 40 unique websites (10 breast, 9 prostate, 10 lung, and 11 MDS); 18% were .gov., 0% .edu., 40% .com, and 42% .org. Interobserver agreement ranged from 67% to 100% for the key items (mean, 86%).

Overall, 85% of the websites [95% confidence interval (CI), 70%–94%] mentioned clinical trials on the landing page, and 73% (95% CI, 56%–85%) formally defined them; however, only 53% (95% CI, 36%–68%) listed benefits and risks of trials. A total of 68% (95% CI, 51%–81%) contained a link to a specific clinical trial. Other measures of accessibility and quality are detailed in Table 1; these did not differ significantly by disease type (data not shown).

Of the sites with a link to a specific trial (n = 27), only 44% (95% CI, 25%–65%) contained searchable content. Additional trial-specific quality and accessibility measures are detailed in Table 2. Websites for the four disease types had different proportions with interactive features (P = 0.008), and those for lung cancer and MDS were less likely to have these features compared with breast and prostate cancer sites (20% vs. 75%, P = 0.007).

Interestingly, of those sites with interactive features (n = 12), half linked to an outside rather than an internal source of trial data. In all external cases, this was the NCI clinical trials database.

## Discussion

Most of the websites in our analysis contained clinical trial information on their landing page, although the quality of that information varied from site to site. Moreover, we found that

Table 2. Quality and accessibility of links to specific cancer-related clinical trials

<table>
<thead>
<tr>
<th>Trial information</th>
<th>Overall N = 27</th>
<th>Percentage (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title of study</td>
<td>96 (81–100)</td>
<td></td>
</tr>
<tr>
<td>Treatment regimen</td>
<td>78 (58–91)</td>
<td></td>
</tr>
<tr>
<td>Purpose of study</td>
<td>96 (81–100)</td>
<td></td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>93 (76–99)</td>
<td></td>
</tr>
<tr>
<td>Exclusion criteria</td>
<td>74 (54–89)</td>
<td></td>
</tr>
<tr>
<td>Trial status</td>
<td>93 (76–99)</td>
<td></td>
</tr>
<tr>
<td>NIHCT/NCI trial number</td>
<td>89 (71–98)</td>
<td></td>
</tr>
<tr>
<td>Principal investigator</td>
<td>59 (39–76)</td>
<td></td>
</tr>
<tr>
<td>Contact email or phone</td>
<td>81 (62–94)</td>
<td></td>
</tr>
<tr>
<td>Date clinical trial posted</td>
<td>78 (58–91)</td>
<td></td>
</tr>
<tr>
<td>Date clinical trial updated</td>
<td>85 (66–96)</td>
<td></td>
</tr>
<tr>
<td>Interactive interfacea</td>
<td>44 (25–65)</td>
<td></td>
</tr>
</tbody>
</table>

* Websites were considered to have an interactive interface if there was an option to input patient-specific information (age, gender, cancer subtype, treatment status, and/or location) into a web search engine to identify clinical trials for which a patient might be eligible.
interactive features were less likely with less curable (lung) and rarer (MDS) cancers, where clinical trials are arguably most important. In addition, given that patients desire to be well-informed about the potential financial aspects of investigators and trial sponsors (16), while it was reassuring that sponsors were always listed, we were disappointed to find that the principal investigator was named on fewer than 2 out of 3 websites.

We were also disheartened to see that the readability of the websites we assessed was overall poor (grade level 11.5), and posit that this may be one reason that patients with higher education levels have been more aware of clinical trials (9). Indeed, the United States National Library of Medicine recommends that “easy-to-read” health material should be designed at the 6th or 7th grade reading levels. A study evaluating 62 prostate cancer websites in 2011 similarly found the median grade level to be 12.0 (range, 8.0–12.0). Our data suggest that such poor readability extends beyond general prostate cancer information to clinical trial information for prostate and other cancers, and provide a clear area for improvement (17).

Our study has limitations. First, we did not specifically assess the layout and visual presentation of clinical trial information, despite studies demonstrating that patients have specific font, color, style, navigation, and icon preferences which may affect comprehension (18). Next, our methods did not allow us to determine ultimate effects: it is possible that the high accessibility and variable quality of the websites we evaluated have no impact on actual trial accrual. Finally, our analysis would be difficult to precisely reproduce, as real-world websites such as the ones we analyzed are constantly changing.

In summary, our data suggest that basic clinical trial information is widely available on the Internet for naïve users searching for cancer treatment information. On the other hand, the quality of that information is highly variable, and for some cancers, few sites offer interactive features. More research is needed to determine the best way of harnessing the power of the Internet—and its progeny, social media—to communicate high-yield cancer-related clinical trial information. Finally, given that a significant portion of interactive sites found through naïve searches ultimately link to NCI’s databases, the quality of the information preceding that database link may merit careful review.

**Disclosure of Potential Conflicts of Interest**

No potential conflicts of interest were disclosed.

**Authors’ Contributions**

Conception and design: G.A. Abel, S.W. Gray

Development of methodology: G.A. Abel, A.M. Cronin, S.W. Gray

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): G.A. Abel, K. Earles, S.W. Gray

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G.A. Abel, A.M. Cronin, K. Earles, S.W. Gray

Writing, review, and/or revision of the manuscript: G.A. Abel, A.M. Cronin, K. Earles, S.W. Gray

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): K. Earles

Study supervision: G.A. Abel, S.W. Gray

**Grant Support**

S.W. Gray is supported by the American Cancer Society (120529-MRSS-11-006-01-CPPB) and the National Human Genome Research Institute (U01HG006492).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 1, 2015; revised June 26, 2015; accepted July 14, 2015; published OnlineFirst August 11, 2015.

**References**
