

Usability test plan: a patient information page on a rare disease

Kelly Stowell, MSN, APRN, RNFA, FNP-C, AG/ACNP-BC

A public information page on a rare disease has two very different readers. The parent who heard the diagnosis last week has to find a clinical trial, a support group and help with the bills on the same page a hematologist opens for a quick refresher. If only the hematologist can use it, the page has failed the reader who needs it most. This plan tests whether one such page works for both.

The plan uses two methods. A reviewer scores the page text against the Federal Plain Language Guidelines [1], and five readers, two clinical and three lay, work through five practical tasks and then rate what they found. Five readers cannot tell a site owner how many people struggle with the page, but they will show where readers get stuck, which is what an editor needs to know.

Research question

Can a medical professional and a lay reader each use the page to find practical help and understand what they read? If the two groups differ, we also want to know which parts of the page cause the gap.

Participants

We will recruit five adults: two medical professionals and three lay readers. Each must know someone recently diagnosed with the condition the page covers, so every participant comes to the page with a real reason to use it.

The lay group is larger because lay readers carry the most risk. Norman and Skinner define eHealth literacy as the ability to seek, find, understand and appraise health information from electronic sources, then apply it to a health problem [2]. A page that assumes clinical vocabulary asks lay readers for skills they may not have, and the two professionals give us a group to compare them with. Each participant receives a role label (MP1 and MP2, LP1 to LP3), and no names appear in the data or the report.

Tasks

Each participant works through five tasks on the site.

1. Find a clinical trial for the condition.
2. Contact the site's sponsor to ask about using screen reader equipment with the website.
3. Find financial resources to help with medical care.
4. Find a caregiver support group.
5. Find nutrition support information.

The tasks mirror what a newly diagnosed family asks in the first weeks. None can be done from the disease summary alone, so the reader has to search the site and follow links, then decide whether the page they land on answers the question.

Session procedure

Each participant works alone at a computer in an office. The session opens with a standard script that explains the tasks, and the participant then uses the live site and fills out a paper survey. Once the survey begins, the proctor answers no questions. Participants may stop at any point if they become frustrated or confused, and we ask them not to discuss the session with other participants.

Instrument 1: readability checklist

One reviewer scores the page text with a yes or no checklist built on the 2011 Federal Plain Language Guidelines [1]. Participants do not complete it. The checklist has 22 items under 10 criteria, and a "yes" means the statement accurately describes the page. For each item the reviewer writes a comment that names the passage behind the answer.

Item	Criterion	Question
1.1	Understanding the audience	Does the content speak to the reader as "you" or "your" throughout?
1.2		Does the content suit the knowledge level of patients, caregivers and medical professionals?
2.1	Clarity and conciseness	Are sentences 20 words or fewer?
2.2		Is the content mainly in active voice, with the subject performing the action?
2.3		Are technical terms and jargon avoided or explained?
3.1	Organizational logic	Does the most important information come first?
3.2		Do headings and subheadings guide the reader?
4.1	Accessibility and layout	Does each paragraph hold a single idea and stay short?
4.2		Are font size and style easy to read for a wide range of readers?
4.3		Are complex points broken into lists?

5.1	Visual aids	Are there helpful images, charts or graphs?
5.2		Do visuals carry captions or descriptive text?
6.1	Consistency	Is the same term used for the same thing throughout?
6.2		Are formatting and style uniform across the site?
7.1	Tone and engagement	Do the style and mood of the writing suit the subject and the readers?
7.2		Does the style hold the reader's attention and make the content easy to move through?
8.1	Actionability	Are action steps and instructions clear?
8.2		Is it clear what readers can do after reading?
9.1	Link use	Are links well labeled, and do they lead to relevant information?
9.2		Is link text descriptive and specific, without generic phrases?
10.1	Feedback	Can readers give feedback on the content?
10.2		Is contact information for further questions easy to find?

Instrument 2: user survey

The paper survey comes after the tasks. Its first question records role: "I work in the medical field" or "I'm not a medical professional." Ten rated questions follow, each on a 5-point scale with space for a comment. The questions draw on Markel's chapter on research methods for technical communicators [3] and on the Nielsen Norman Group method for heuristic evaluation [4]. A final open box asks for any other comments or suggestions.

Survey question	Scale, most to least favorable
Finding information. How simple was it to locate the five topics on the website?	Extremely simple, simple, neutral, challenging, extremely challenging
Understanding the content. Was the information clear and easy to understand, with words and ideas you know?	Extremely clear, clear, neutral, challenging, extremely challenging
Your navigation experience. How easy was it to move around the website, and did you feel in control?	Extremely simple, simple, neutral, challenging, extremely challenging
Consistency and design. Was the website consistent in its words, actions and design?	Extremely consistent, consistent, neutral, inconsistent, highly inconsistent
Avoiding and handling errors. Were there clear ways to fix navigation mistakes, and were error messages easy to understand?	Always, often, sometimes, rarely, never
Ease of recall. Did the website show important information where you needed it, without making you remember it from another page?	Always, often, sometimes, rarely, never

Flexibility for new users. Did the website work well for someone using it for the first time?	Extremely well, well, neutral, not very well, not at all
Design and style. The website has a minimalist look. How appealing is it to you?	Extremely appealing, appealing, neutral, unappealing, extremely unappealing
Explaining complex terms and visual help. Were harder terms made easy to understand, and did pictures or graphs help?	Extremely clear and helpful, clear and helpful, neutral, unclear and unhelpful, very unclear and unhelpful
Your thoughts. How satisfied are you with what the website told you about the condition?	Extremely pleased, pleased, neutral, displeased, very displeased

Analysis plan

We report the checklist total, the number of "yes" answers out of 22, along with a score for each of the 10 criteria. A criterion that scores zero points to a problem across the whole text, and that matters more than any one missed item.

On the survey, either of the top two points on a scale counts as a favorable rating. We count favorable ratings for each question within each group and report each count over the size of that group. When professionals rate a question favorably and lay readers do not, that question goes first in the report. With five participants we compare counts only and will not compute means, percentages or tests of significance, since numbers that small would look more precise than they are.

We sort every comment by the problem it names (search, page structure, empty pages, language, finding help) and tie each problem to the checklist item or survey question it explains, and we quote participants word for word.

Common problems to test for

Pages written for mixed audiences tend to fail in the same places. The reviewer should look for each of these in the checklist review and the session notes.

Check	Where it shows up
Are medical terms and acronyms, including the disease's own abbreviation, defined at first use?	Checklist 2.3. Survey questions on understanding and on complex terms.
Do most sentences stay at 20 words or fewer, in active voice?	Checklist 2.1 and 2.2
Do link labels spell out acronyms and name their destination?	Checklist 9.1 and 9.2
Can a reader search within the page, or jump to a section from a list at the top?	Survey questions on finding information and navigation
Does the page scroll as one, or does a second scroll area sit inside it?	Survey question on navigation

Do any links lead to empty pages or pages marked as in development?	Survey question on errors
Are practical resources (trials, financial help, caregiver support, nutrition) grouped where a reader can find them?	Tasks 1, 3, 4 and 5
Do images carry captions that say what they show?	Checklist 5.2

Limits of the design

Five participants can surface problems but cannot estimate how often they occur, so the report will rank problems without claiming prevalence. Every participant knows someone with the condition. That makes them motivated readers, and they may be more patient than a first-time visitor.

One reviewer scores the checklist, although the Nielsen Norman Group recommends three to five evaluators working independently [4]. Both instruments were written for this test and have not been validated. Participants rate the tasks afterward on paper and no one observes them, so the data capture what readers remember rather than task success or time on task. Because the site can change at any point, the report will record the date of each session.

What we will report

The site owner gets a short report. It opens with the checklist total and the score for each criterion, followed by favorable survey ratings by group, with denominators, and a list of every question where the groups split. Participant comments appear word for word under the problem they describe. The report ends with the problems ranked. For each one the editor gets a fix to make and the checklist item or survey question that will show whether it worked.

Track how many lay readers rate the content clear. After the wording is fixed, retest with three new lay readers and compare that count with the first round.

References

1. Plain Language Action and Information Network. Federal plain language guidelines. March 2011, revision 1, May 2011. <https://digitalgovernmenthub.org/wp-content/uploads/2022/07/FederalPLGuidelines.pdf>
2. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res. 2006; 8(2):e9. doi:10.2196/jmir.8.2.e9. PMID: 16867972.
3. Markel M. Technical communication. 11th ed. Bedford/St. Martin's. 2015. Chapter 6, Researching your subject.
4. Moran K, Gordon K. How to conduct a heuristic evaluation. Nielsen Norman Group. June 25, 2023. <https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/>

Writing sample prepared for demonstration. Not peer reviewed or published.