Are English-language online patient education materials related to breast cancer risk assessment understandable, readable, and actionable?

Leslie R. Lamb a, Grayson L. Baird b, Ishita T. Roy a, Paul H.S. Choi c, Constance D. Lehman a, Randy C. Miles a, *

a Massachusetts General Hospital, Department of Radiology, 55 Fruit Street Boston, MA, 02114-2696, USA
b Rhode Island Hospital, Warren Alpert School of Medicine at Brown University, Department of Diagnostic Imaging, 593 Eddy Street, Providence, RI, 02903, USA
c Tufts Medical Center, 800 Washington St Boston, MA, 02111, USA

Article info

Article history:
Received 16 September 2021
Received in revised form 16 November 2021
Accepted 16 November 2021
Available online 25 November 2021

Keywords:
Readability
Patient education
Risk assessment
Breast cancer

Abstract

Purpose: To evaluate the readability, understandability, and actionability of online patient education materials (OPEM) related to breast cancer risk assessment.

Material and methods: We queried seven English-language search terms related to breast cancer risk assessment: breast cancer high-risk, breast cancer risk factors, breast cancer family history, BRCA, breast cancer risk assessment, Tyrer-Cuzick, and Gail model. Websites were categorized as: academic/hospital-based, commercial, government, non-profit or academic based on the organization hosting the site. Grade-level readability of qualifying websites and categories was determined using readability metrics and generalized estimating equations based on written content only. Readability scores were compared to the recommended parameters set by the American Medical Association (AMA). Understandability and actionability of OPEM related to breast cancer high-risk were evaluated using the Patient Education Materials Assessment Tool (PEMAT) and compared to criteria set at ≥70%. Descriptive statistics and inter-rater reliability analysis were utilized.

Results: 343 websites were identified, of which 162 met study inclusion criteria. The average grade readability score was 12.1 across all websites (range 10.8–13.4). No website met the AMA recommendation. Commercial websites demonstrated the highest overall average readability of 13.1. Of the 26 websites related to the search term breast cancer high-risk, the average understandability and actionability scores were 62% and 34% respectively, both below criteria.

Conclusions: OPEM on breast cancer risk assessment available to the general public do not meet criteria for readability, understandability, or actionability. To ensure patient comprehension of medical information online, future information should be published in simpler, more appropriate terms.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

In the United States, approximately one in eight women will develop breast cancer, which is the second leading cause of cancer death in women [1,2]. There are cohorts of women considered at increased risk for developing breast cancer, yet the majority of patients are unaware of their risk status, despite demonstrating interest in learning about their personal risk [3]. This suggests that patients at increased risk may not always be counselled appropriately for screening and risk reduction strategies. As internet usage has increased, patients often rely on the internet for health information [4]. The majority of patients indicate that online patient educational material (OPEM) influences their medical decision making [5]. It is therefore important to assess the readability, understandability, and actionability levels of OPEM available related...
to risk, as this material offers an important opportunity for patient education.

For patients at average risk of developing breast cancer, the American College of Radiology recommends annual screening mammography [6]. Patients with a calculated lifetime risk of 20% or higher, or a history of chest/mantle radiation therapy at a young age, may benefit from supplemental screening with contrast-enhanced breast magnetic resonance imaging (MRI) and other risk reduction strategies such as chemoprevention [7–12]. Several commercial statistical models are available to predict a patient’s lifetime risk of developing breast cancer and/or the risk of carrying a high-risk mutation, including the Tyrer-Cuzick and Gail models [13]. These are typically performed in sub-specialty breast clinics and imaging facilities; however, they are also available online for patients to access. Some breast imaging facilities report patients risk status at the time of screening mammography. As of November 2020, with the introduction of the 21st Century Cures Act, all patients now have immediate access to clinical notes, including radiology reports [14]. For patients identified as high-risk, particularly if this was previously unknown, this can cause substantial anxiety and vulnerability.

Health literacy, or the capacity to obtain, communicate, process, and understand health information, is a significant component of effective shared decision making [15]. More than one-third of the United States adult population has difficulty comprehending health-related information and nearly 50% demonstrate health literacy equivalent to or below fifth-grade reading level [16]. The current recommendation by the American Medical Association (AMA) states that all health materials should be written at or below a sixth-grade reading level, or equivalent to six years of schooling, to be more comprehensible to the general public [17]. Despite these recommendations, prior studies have shown that the readability of OPEM does not meet the suggested guidelines [18,19]. As such, the development of accurate, accessible, and actionable health information has been recognized as a national priority [20].

To our knowledge, there have been no prior multimetric investigations including understandability and actionability related to breast cancer risk assessment. With the fast-changing internet landscape and recent implementation of the Cures Act, we sought to evaluate readability, understandability, and actionability of OPEM related to breast cancer risk assessment. This information may be used to inform clinicians and those involved with online information dissemination, to ensure online material is appropriate for patients’ level of understanding.

2. Material and methods

2.1. Study design

Our study did not involve human subjects. The need to obtain Institutional Review Board (IRB) approval was waived. OPEM was defined as any online educational material on breast cancer risk assessment that was directed towards the general public. Sponsored hits, research journal articles, and websites intended for healthcare professionals were excluded from this evaluation.

2.2. Data acquisition and refinement

On January 24, 2020, the English-language terms: breast cancer high-risk, breast cancer risk factors, breast cancer family history, BRCA, breast cancer risk assessment, Tyrer-Cuzick, and Gail model were queried using Google (Google Inc., Mountain View, CA). Tyrer-Cuzick was included as it is the most consistently accurate commercial model for breast cancer risk, according to limited prospective data [21,22]. It incorporates a wide variety of personal risk factors and extensive family history [13]. The Gail model, developed in 1989 and also known as the National Cancer Institute’s Breast Cancer Risk Assessment Tool (BCRAT), was included as it is the most widely used model in clinical practice [23]. It focuses mainly on a woman’s medical and reproductive history with limited data on family history and is designed to predict the absolute risk of developing breast cancer [13]. All information was publicly available. Location, cookies, and user account information were disabled prior to our queries. The first five pages of generated results were evaluated.

Websites were categorized as either academic, academic/hospital-based, commercial, government or non-profit, depending on the organization hosting the sites. Academic websites included those affiliated with a university or an academic medical center. Academic/hospital-based websites included those which were associated with a medical center (ex. private clinic), but not affiliated with a university. Commercial websites included those that displayed advertisements or sold products. Government websites included those that operated with governmental funding. Non-profit websites included those that were operated by non-profit organizations. Information on the presence of diagrams, links to resources/other websites, reference citations, and videos for further information was then collected for each website.

2.3. Text editing

Patient-directed information from each website was downloaded, and their uniform resource locators (URLs) were noted. For the readability assessment, content was formatted into plain text in a Microsoft Word document (Microsoft Office 2020; Microsoft, Redmond, Washington), excluding images, videos, copyright notices, citations, hyperlinks, figures, captions, advertisements, references, links, disclaimers, and acknowledgments. Within the documents, all punctuation, symbols and non-numerical/alphabetical bullet points were removed, with the exception of periods.

2.4. Data analysis readability assessment

Formatted Word documents were fed into a text analysis program (https://readable.io/, Sussex, England) to calculate their readability scores based on written content only. Scores were based on Flesch-Kincaid reading ease index [24], Flesch-Kincaid grade level [25], Gunning-Fog score [26], Coleman-Liau index [27], simple measure of gobbledygook (SMOG) index [28], and automated readability index (ARI) [29]. Readability scores calculated by these algorithms, account for letters, words and complexity of sentence structures. All tests report a score or range of scores to provide an overall readability analysis metric.

2.5. Understandability and actionability assessment

Websites identified related to breast cancer high-risk were further evaluated using the Patient Education Materials Assessment Tool (PEMAT); [https://www.ahrq.gov/neckcr/tools/self-mgmt/pemat.html](https://www.ahrq.gov/neckcr/tools/self-mgmt/pemat.html). The PEMAT is a systematic method to evaluate and compare understandability and actionability of patient education materials. The PEMAT was designed to assess patients’ ability to both comprehend and act on health information [30]. It evaluates...
both website written content and visual aids (diagrams, photographs, graphs, etc.) It consists of 19 questions to determine understandability and seven questions to determine actionability. Answer options include: “Agree” (1 point), “Disagree” (0 points), and “Not Applicable.” “Not Applicable” items are excluded from analysis when calculating overall score (Supplementary Tables 2 and 3).

Two independent physician evaluators (RM, LL) scored each website (26 questions total). The score for each question was divided by the total possible points, excluding “Not Applicable” items and multiplied by 100 to determine each score as a percentage. A higher score represents a greater understandability or actionability. Based on prior literature, OPEM was defined as understandability and actionability score of ≥70% on either section [31].

2.6. Statistical analysis

All analyses were conducted using SAS Software (9.4; SAS Inc., Cary, NC) using the GLIMMIX procedure. Readability, understandability, and actionability were evaluated using generalized estimating equations (GEE) with sandwich estimation assuming normal and binomial distributions, where appropriate. Interval estimates were calculated for 95% confidence. Inter-rater reliability between physicians’ understandability and actionability scores was estimated using the intraclass correlation coefficient (ICC) assuming random effects.

3. Results

Our initial search identified a total of 343 websites. We excluded 167 peer-reviewed journals and 14 websites that did not qualify as patient-directed education resources (ex. Cancer Network, ASCO Post, American Academy of Family Physicians). A total of 162 websites met study inclusion criteria.

The highest number of websites were categorized as commercial (30.9%, 50/162), followed by non-profit (25.9%, 42/162), academic (21.0%, 34/162), academic/hospital-based (12.3%, 20/162), and government (9.9%, 16/162). Commercial websites had the highest average grade reading level at 13.1, followed by academic (12.3), government (12.0), non-profit (11.7), and academic/hospital-based (11.3).

Of the seven search terms, Tyrer-Cuzick had the highest grade level readability score of 13.3 across industries, followed by breast cancer risk assessment (13.0), Gail model (12.7), breast cancer family history (12.0), BRCA (11.8), breast cancer risk factors (11.6), and breast cancer high-risk (11.6). Table 1 summarizes the overall readability of OPEMs across all search terms.

The overall grade level readability across all websites was 12.1 [95% CI (10.8, 13.4)]. No website met the AMA-suggested limit of a sixth-grade reading level or lower. The readability score distribution according to grade reading is illustrated in Fig. 1. The website with the highest average grade level readability score was Duke University–Precision Medicine at 19.3 (search term: breast cancer risk assessment), and the website with the lowest readability score was American Cancer Society at 7.6 (search term: breast cancer family history). Supplementary Fig. 1 illustrates the average grade level readability scores for each website, in contrast to AMA-recommended parameters. Supplementary Table 3 illustrates examples of website passages written at higher and lower grade reading levels.

Of the 26 websites for the search term breast cancer high-risk, the average understandability across websites was 62.4% [95% CI (0.48, 0.75)] and actionability was 33.9% [95% CI (0.27, 0.41)] based on PEMAT scoring. 23.1% (6/26) met criteria (≥70%) for understandability and 7.7% (2/26) met criteria for actionability. The industry with the highest average understandability was government (84.6%) and lowest was commercial (56.7%). The industries with the highest average actionability were government and non-profit (50.0%) and lowest was academic (20.0%). Inter-rater reliability of understandability and actionability between raters was high at ICC of 0.91 and 0.85 respectively. Table 2 summarizes the understandability and actionability scores, Supplementary Table 4 illustrates examples of website passages written at higher and lower understandability and actionability levels, and Supplementary Fig. 2 illustrates the average understandability and actionability scores for each website.

### Table 1
#### Website readability across seven search terms by industry.

| Search Term                      | Readability | Average | Lower Limit | Upper Limit |
|----------------------------------|-------------|---------|-------------|-------------|
| Breast cancer high-risk (26)     | 11.6        | 10.4    | 12.9        |
| Academic/hospital-based (4)      | 11.7        | 10.5    | 12.9        |
| Commercial (8)                   | 12.8        | 11.6    | 14.1        |
| Government (1)                   | 9.7         | 8.7     | 10.7        |
| Non-profit (8)                   | 10.5        | 9.2     | 11.8        |
| Academic (5)                     | 12.0        | 10.9    | 13.0        |
| Breast cancer risk factors (37)  | 11.6        | 10.5    | 12.7        |
| Academic/hospital-based (5)      | 10.9        | 9.6     | 12.2        |
| Commercial (11)                  | 12.1        | 11.0    | 13.2        |
| Government (4)                   | 11.5        | 10.3    | 12.7        |
| Non-profit (10)                  | 10.7        | 9.7     | 11.8        |
| Academic (7)                     | 12.5        | 11.5    | 13.6        |
| Breast cancer family history (33)| 12.0        | 10.5    | 13.5        |
| Academic/hospital-based (1)      | 10.2        | 8.4     | 12.0        |
| Commercial (9)                   | 12.2        | 10.7    | 13.8        |
| Government (3)                   | 12.9        | 11.4    | 14.3        |
| Non-profit (12)                  | 11.4        | 10.0    | 12.8        |
| Academic (8)                     | 12.6        | 11.1    | 14.1        |
| BRCA (27)                        | 11.8        | 10.4    | 13.3        |
| Academic/hospital-based (3)      | 10.8        | 9.2     | 12.5        |
| Commercial (8)                   | 13.2        | 11.7    | 14.6        |
| Government (4)                   | 10.8        | 9.3     | 12.4        |
| Non-profit (3)                   | 12.0        | 10.5    | 13.5        |
| Academic (9)                     | 11.4        | 10.0    | 12.8        |
| Breast cancer risk assessment (17)| 13.0      | 11.6    | 14.5        |
| Academic/hospital-based (2)      | 12.2        | 10.7    | 13.7        |
| Commercial (4)                   | 12.7        | 11.3    | 14.0        |
| Government (2)                   | 13.5        | 12.4    | 14.6        |
| Non-profit (4)                   | 14.0        | 12.8    | 15.3        |
| Academic (5)                     | 13.2        | 11.6    | 14.8        |
| Tyrer-Cuzick (12)                | 13.3        | 12.1    | 14.5        |
| Academic/hospital-based (4)      | 14.4        | 13.2    | 15.5        |
| Commercial (5)                   | 12.9        | 11.7    | 14.0        |
| Government (1)                   | 14.3        | 13.1    | 15.6        |
| Non-profit (2)                   | 11.7        | 10.5    | 13.0        |
| Academic (0)                     | -           | -       | -           |
| Gail model (10)                  | 12.7        | 11.4    | 14.0        |
| Academic/hospital-based (1)      | 9.3         | 7.9     | 10.7        |
| Commercial (5)                   | 12.4        | 11.2    | 13.6        |
| Government (1)                   | 15.7        | 14.4    | 17.1        |
| Non-profit (3)                   | 13.4        | 12.0    | 14.7        |
| Academic (0)                     | -           | -       | -           |
| Overall (162)                    | 12.1        | 10.8    | 13.4        |
Online patient education is important, as patients increasingly rely on the internet as a source for medical information. Our study shows that the average readability of OPEM on breast cancer risk assessment terms, exceeded the AMA-recommended limit set at a sixth-grade reading level, with an overall grade level readability of 12.1, across all websites; the website closest to the AMA-recommendation had a readability score of 7.6. The average understandability and actionability scores of breast cancer high-risk were 62% and 34% respectively, exceeding the set criteria. These results highlight a need for patient-directed resources written at a lower reading level that are more understandable and actionable, considering nearly half of the United States adult population demonstrate health literacy equivalent to or below a fifth-grade level [16].

Our findings are in concordance with prior studies on breast cancer prevention and treatment, which indicate that current OPEM are well beyond national AMA-recommended parameters.

**4. Discussion**

Online patient education is important, as patients increasingly rely on the internet as a source for medical information. Our study shows that the average readability of OPEM on breast cancer risk assessment terms, exceeded the AMA-recommended limit set at a sixth-grade reading level, with an overall grade level readability of 12.1, across all websites; the website closest to the AMA-recommendation had a readability score of 7.6. The average understandability and actionability scores of breast cancer high-risk were 62% and 34% respectively, exceeding the set criteria. These results highlight a need for patient-directed resources written at a lower reading level that are more understandable and actionable, considering nearly half of the United States adult population demonstrate health literacy equivalent to or below a fifth-grade level [16].

Our findings are in concordance with prior studies on breast cancer prevention and treatment, which indicate that current OPEM are well beyond national AMA-recommended parameters.

Readability scores assessed for each individual website in our study exceeded a sixth-grade reading level. Cortez et al. found similar findings when they evaluated breast cancer risk assessment tools (search terms: calculate breast cancer risk, breast cancer risk calculator, estimate breast cancer risk, assess breast cancer risk, and breast cancer risk assessment), in a study performed more than five years previously. They found the overall readability of websites was 12.1, the same as our study [32]. This suggests there has been no change in the readability of OPEM over the last several years, despite national attention. A readability study on OPEMs for operative treatment of breast cancer found an overall readability level across sites of 12.9 (range 11.2–16.5), also similar to our study [21].

A readability study on OPEM for breast cancer screening, found that websites are written significantly above the sixth-grade reading level [22].

There are no published studies, to our knowledge, on the understandability and actionability of OPEM related to breast cancer risk assessment. In our analysis of breast cancer high-risk, the average understandability and actionability scores were 62% and 34% respectively and did not meet criteria based on previous literature [31]. Understandable OPEM is "when consumers of different backgrounds and varying levels of health literacy can process and explain key messages" and actionable OPEM is "the ability of patients with diverse backgrounds and varying levels of health literacy to identify what they can do based on the information presented" [20,28]. Thus, understandable and actionable OPEM help provide a more comprehensive picture of analysis, when combined with readability, particularly because they account for when complex terms are defined or when words with complex syntax are actually common terms to the lay public. Our results are concerning. Although a foundational understanding of health information is crucial, acting is equally, or perhaps more important, as actionability is critical for breast cancer screening and prevention. Since OPEM is a primary source for self-education, websites need to make significant improvements in targeting realistic approaches that are actionability-focused to support change in behavior.

In our analysis, government websites published the most understandable sources of information. While this is reassuring,
healthcare system distrust has risen over the last several decades and is playing an increasing role in breast cancer care [33,34]. Fewer than one third of Americans believe government officials are credible and only a quarter have expressed a great deal or quite a lot of confidence in the healthcare system [33,35]. Thus, the public may be reluctant to trust the accuracy of information provided by the government and rely on other websites, such as commercial sources. In our study, commercial websites had the highest readability, lowest understandability, and second lowest actionability scores, highlighting the need for these OPEM to be simplified to layman’s reading levels. Patients may also turn to academic sources for information which we found, not surprisingly, were geared to more educated patients. Academic sources had the second highest readability, second lowest understandability, and lowest actionability scores. There needs to be a better attempt at adapting academic based OPEM to showcase improved readability, understandability, and actionability for patients, particularly given they are viewed as certified sources, second only to healthcare providers [36]. Alternatively, publishers may consider two tiers of OPEM - a collection targeted to less than a sixth-grade reading level (including appropriate referral strategies and the availability of high-risk clinics prior to disseminating risk assessment results, it also highlights the need to have appropriate OPEM, as high-risk referral strategies do not always exist. Patients will inevitably search the internet for information related to risk upon reading a mammographic report detailing their risk, and possibly before they have a chance to discuss with their provider. Nearly 77% of adults are known to browse online for health-related topics [6] and approximately 88% of patients are known to browse online directly following a cancer diagnosis [39].

While OPEM certainly need to be amended to meet the general public’s reading, understandability and actionability levels, clinicians should also be aware of the educational level of available material to patients to proactively inform their discussions. They can be mindful to address questions and clarify information that patients receive from OPEM. Effective communication is important to reduce unnecessary confusion and anxiety. Studies have shown that information alleviates patient anxiety [40-43]. If patients are informed and engaged, they are also more likely to be involved with health literacy through education while also ensuring patients are aware that OPEM does not replace consultation with healthcare professionals.

There are important limitations to our study. First, our search included only websites written in English, therefore results cannot be generalized for OPEM written in other languages. Future studies should incorporate search terms in other languages, as English is the primary language in only 78.1% of Americans [46]. Our search was also limited to Google as it is the most widely used online search engine [47]. Finally, our list is not comprehensive of all possible breast cancer risk assessment terms. Since it is not possible to include every term in our search, we aimed to include those with commonly used terms in breast imaging practices.

5. Conclusions

In our study of 162 English-language websites related to breast cancer risk assessment OPEM, the average readability score was 12.1, much higher than the sixth-grade readability score recommended by the AMA. The average understandability and actionability scores were low, at 62% and 34%. OPEM serves as a valuable resource for patients to educate themselves; however, the unfiltered nature of information can lead to misinformation, and unnecessary anxiety and confusion, if the readability of these resources is not catered to the readers’ level of understanding. To ensure patient comprehension of medical information online, future information should be published in simpler and more appropriate terms.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethics approval

Our study did not involve human subjects. The need to obtain Institutional Review Board (IRB) approval was waived.

Declaration of competing interest

Randy Miles has an early career award funded by General Electric (GE) through the Association of University Radiologists (AUR), GE Radiology Research Academic Fellowship (GERRAF) Award (https://www.aour.org/en/awards/aur-gerraf-award).

Constance Lehman receives institutional research support from GE Healthcare Inc., Hologic, Inc., Breast Cancer Research Foundation and is a cofounder of Clarity, Inc.

Leslie Lamb receives research support from the Breast Cancer Research Foundation.

Acknowledgements

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.breast.2021.11.012.

References

[1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA A Cancer J Clin 2015;65(2):87–108.
[2] American Cancer Society. How common is breast cancer?. Available via, https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html. [Accessed 31 January 2021].
[3] Amornsiripanitch N, Mangano M, Neill BL. Screening mammography: patient perceptions and preferences regarding communication of estimated breast cancer risk. AJR Am J Roentgenol 2017;208(5):1163–70.
[4] Fox S, Duggan M. Pew research center. Health online 2013. Available via, http://www.pewinternet.org/2013/01/15/health-online-2013/. [Accessed 31 January 2021].
[5] Berland GK, Elliott MN, Morales LS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. J Am Med Assoc 2001;285(20):2612–21.
