Assessing the Readability and Quality of Web-Based Resources on Exercise Stress Testing: A Cross-Sectional Readability and Quality Analysis

Munir Rahbe, Dhrumi Mistry, Ali M Khawaja, Natalie A Sous, Alan Y Tso

Submitted to: Journal of Medical Internet Research
on: October 25, 2024

Table of Contents

Assessing the Readability and Quality of Web-Based Resources on Exercise Stress Testing: A Cross-Sectional Readability and Quality Analysis

Munir Rahbe[1] MBS, BA; Dhrumi Mistry[1] BS; Ali M Khawaja[2] BA; Natalie A Sous[1, 3] MD; Alan Y Tso[1, 3] MD, MA

[1]Rutgers New Jersey Medical School Newark US
[2]School of Arts and Sciences Rutgers University Newark US
[3]Department of Medicine Rutgers New Jersey Medical School Newark US

Corresponding Author:
Munir Rahbe MBS, BA
Rutgers New Jersey Medical School
185 S Orange Ave,
Newark
US

Abstract

Background: Health literacy is a crucial determinant and independent predictor of patient outcomes. To enhance patient understanding across diverse populations, the American Medical Association (AMA) advises that health information be written at a 6th-grade reading level. The exercise stress test is a widely used diagnostic tool, frequently ordered in outpatient settings to assess cardiovascular health. Although numerous studies have evaluated the readability of various cardiovascular health topics, there is a noticeable gap in the literature regarding the readability and quality of health information pertaining to exercise stress tests.

Objective: The goal of this study was to evaluate the readability and quality of publicly available information regarding exercise stress tests on the Google and Bing search engines.

Methods: Four search terms were entered into both the Google and Bing search engines, and websites containing information pertaining to exercise stress tests were assessed for their readability and quality. Web pages were classified as “academic” or “non-academic” based on whether the publishing website was affiliated with an academic institution. Readability was assessed using the Flesch-Kincade Grade Level (FKGL) and Flesch Reading Ease (FRE) scores, while quality was assessed using the DISCERN criteria. The resulting scores were calculated, and independent sample t-tests were performed to compare the individual scores of the academic and non-academic sources.

Results: This study assessed 18 websites in total, and the mean grade level (mean ± SD, 8.36 ± 1.32) and reading ease (62.31 ± 6.65) exceeded the current AMA recommendations. Furthermore, the quality of the web pages was determined to be “fair” as per the DISCERN grading criteria. Academic sources had greater FKGL and FRE scores that were additionally found to be statistically significant in comparison to the scores of the non-academic sources with p-values of 0.032, and 0.006, respectively.

Conclusions: Websites containing information pertaining to exercise stress tests are written at a grade level that exceeds current recommendations, with academic sources being significantly more difficult to read when compared to non-academic sources. The quality of the web pages evaluated in this study was deemed to be fair. However, quality can be improved by including references to sources as well as additional resources for readers seeking more information. Improving readability can be challenging as medical terminology is often intrinsically complex. While substituting or omitting these terms may improve readability, it risks delivering incomplete or oversimplified information to patients.

(JMIR Preprints 25/10/2024:68000)
DOI: https://doi.org/10.2196/preprints.68000

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Assessing the Readability and Quality of Web-Based Resources on Exercise Stress Testing: A Cross-Sectional Readability and Quality Analysis

Introduction

As of 2023, nearly 95% of adults in the United States use the internet[1]. In today's information age, the challenge has shifted from simply accessing the internet to finding information that is both reliable and accessible. This shift is especially significant as more individuals rely on the internet for crucial health information. In 2022, 58.5% of adults reported using the internet at least once in the past year[2], and 70% turn to the internet as their primary source for health-related inquiries[3]. However, medical information is often nuanced, complex, and difficult to comprehend. What happens when patients encounter content that is difficult to read? Research shows that low health literacy is a significant independent predictor of all-cause mortality[4,5]. This emphasizes the vital role the internet plays in modern healthcare and highlights the need for accessible, understandable health information to improve patient outcomes.

As early as the year 2000, professionals began publishing articles urging physicians to guide their patients using objective scales to "grade" the quality of online health information. The suggestion was that physicians should take responsibility for advising patients on how to "surf the net safely" when seeking health-related information[6]. Since then, the readability and quality of online health content accessed by patients have remained a key concern for clinicians, leading to numerous studies on the topic[7-10]. These studies examine various factors, but one of the most consistent is readability—specifically, the grade level at which the material is written. According to the American Medical Association (AMA), health materials should be written at or below a 6th-grade reading level to ensure appropriate health literacy for today's patients[11].

The exercise stress test is a widely used diagnostic tool in the outpatient setting for individuals with potential coronary artery disease. It is also a valuable tool in the risk stratification of patients undergoing surgical procedures. The results of these tests have good prognostic value and essential information for clinicians[12]. Numerous studies evaluating the readability and quality of cardiovascular health topics have found that much of the available information is written at a reading

level that exceeds recommended standards[13-15]. However, no studies to date have explicitly focused on exercise stress testing. This study will be the first to assess the readability and quality of online resources related to exercise stress testing.

The goals of this study are to 1) evaluate the readability and quality of health information on exercise stress tests and 2) compare these metrics between academic and non-academic online resources.

Methods

Data Collection

To objectively assess the quality of available information for patients regarding cardiovascular exercise stress testing, the Google and Bing search engines were accessed. These search engines are ranked the top two respective search engines globally, accounting for approximately 94% of all search engine traffic[16]. The engines were interrogated using the four terms: “exercise stress test,” “exercise ECG,” “stress ECG,” and “treadmill ECG” on August 28, 2024. Searches were conducted on a non-university affiliated internet connection in the respective engine’s incognito mode to simulate the experience of a default user. Analysis of relevant sites was limited to the first page of produced results in accordance with previous findings indicating that approximately 92% of all search engine traffic fails to extend beyond the first page[17].

As of August 2024, the Google search engine accounts for approximately 90.5% of global search engine traffic[16]. Additionally, we sought to include the Bing search engine, accounting for an additional 3.9% of search engine traffic, to assess a broader scope of patient experiences. Searches were conducted, and relevant sites were compiled. Search results were excluded if they were published in academic journals, published in academic textbooks, were written in a non-English language, were redundant with previous search results, were outside of the scope of the cardiac stress testing assessment, redirected to a non-article link, or promoted the sale of a product or service. These criteria were established to analyze the most relevant information a patient may encounter in researching this assessment that they would likely evaluate. The final analysis of results included eighteen web pages. The compiled results were categorized as academic (n = 7) or non-academic (n = 11) and assessed for readability and quality. Academic results were operationally defined as sites affiliated with an academic institution, whereas non-academic results were defined as publicly available websites containing relevant information not affiliated with an academic institution.

Assessment of Readability and Quality

The Flesch-Kincaid Grade Level (FKGL) and Flesch Reading Ease (FRE) scores were utilized to objectively assess the readability of internet-based resources pertaining to cardiovascular exercise stress testing. A modified DISCERN criterion (mDISCERN) was used to characterize the objective quality and reliability of the internet-based resource.

The Flesch-Kincaid Reading Ease (FRE) and Flesch-Kincaid Grade Level (FKGL) are the most widely used readability tools in healthcare literature[18]. These tools were originally developed in the 1970s for the United States Navy and were found to correlate well with the reading levels of sailors[19]. Since their inception, they have been adopted by various government organizations, including the Social Security Administration[20]. While these readability tools are not perfect, they are considered reliable when compared to other psychological tests used to assess reading level[21]. The FRE and FKGL indexes have been utilized extensively to objectively quantify the readability of health-related materials and are widely used in the literature for assessing the accessibility of healthcare information[8,9]. The FRE and FKGL metrics use mathematical formulas to quantify an objective readability score of written material based on the number of words per sentence and syllables per word[22,23]. FRE scores range from 0-100, with greater scores indicating greater ease of reading material and lower scores indicating greater difficulty. The FKGL index utilizes the FRE score and extrapolates the produced value to a respective reading grade level. The FKGL scores utilize the same scale as the FRE index, with scores 80-90 indicating a 6th-grade reading level and scores 0-30 indicating a college reading level. Results within these indexes for the eighteen compiled internet-based resources were compared to the AMA guidelines, indicating that patient-targeted reading materials should be written at or below a 6th-grade level[11].

The DISCERN instrument, another widely used tool in the literature, has been validated for the evaluation of health-related information by both healthcare professionals and patients[24]. This instrument was used to objectively assess the quality of internet-based information regarding cardiac exercise stress tests[9,10]. The questionnaire consists of three sections with a total of sixteen questions, designed to evaluate both the reliability and overall quality of online health information. The first component of the instrument, questions one through eight, characterizes the reliability of a published resource. The second component, questions nine through fifteen, characterizes the quality of treatment-related information. The third component, question sixteen, uses the compiled scores from the previous fifteen questions to characterize how well the resource provides information regarding

specific treatment options[25]. All questions are scored on a scale ranging from one to five, with lower scores indicating poorer rankings and higher scores indicating more positive rankings. For the purposes of this study, a modified version of the DISCERN instrument (mDISCERN) was utilized, consisting of only the first component of the tool, as this study solely concentrated on information surrounding diagnostic assessment and not subsequent treatment (Table 1). A mDISCERN score below 40% classified the reliability of a web result as “poor,” scores between 40% and 79% were deemed “fair,” and scores above 80% were categorized as “good.” These thresholds align with scoring criteria established in prior research studies[26,27].

Table 1. Modified DISCERN (mDISCERN) scoring tool

#	Question	Score
1	Are the aims of the article clear?	1-5
2	Does the article achieve stated aims?	1-5
3	Are the topics covered relevant?	1-5
4	Is it clear what sources of information were used to compile the publication other than the author?	1-5
5	Is it clear when the information used or reported in the publication was produced?	1-5
6	Is it balanced and unbiased?	1-5
7	Does it provide details of additional sources of support and information?	1-5
8	Does it refer to areas of uncertainty?	1-5

Text from each of the eighteen compiled internet articles was copied into a Microsoft Word document. Image-derived text within the article was only included in the document if the text was present on a relevant graphic within the article and was clearly intended for the viewer to read. Two independent reviewers used the Microsoft grammar assist tool to calculate the FRE score and FKGL using their respective equations[22,23]. The reviewers also utilized the mDISCERN criteria to evaluate each source material. Scores for each of the questions of the mDISCERN criteria were independently calculated, and scores were averaged.

Statistical Analysis

Scores for the mDISCERN, FRE, and FKGL indexes for each of the analyzed web pages were independently calculated by both reviewers with the assistance of the Microsoft grammar assist tool. Scores for each criterion were averaged, and standard deviations for each variable were calculated.

Independent samples t-tests were performed between mDISCERN, individual mDISCERN questions, FRE, and FKGL scores between academic (n = 7) and non-academic (n = 11) resources to distinguish any significant differences in the quality and readability of information based on the publisher. An alpha value of p = 0.05 was used to determine statistical significance for all performed analyses. All analysis was conducted on version 29.0.2.0 of the SPSS platform.

Ethical Considerations

This study was exempt from review by an institutional review board as all information obtained over the course of the study is publicly available.

Results

Each search term was queried over two search engines, yielding a total of 58 websites. Of the 58 search results, 40 were textbook chapters, duplicates, advertisement links, journal publications, or outside the scope of exercise stress testing. These results were subsequently excluded, and the remaining 18 websites were evaluated. The websites were categorized into two groups: academic (n=7) and non-academic (n=11). Results for the analyzed websites are available in Table 2. The FKGL ranged from 5.15 to 9.85; the mean ± SD was 8.36 ± 1.32. The mean FRE score of all the websites included in this study was 62.31±6.65. The mDISCERN scores ranged from 22-36, with the mean being 29.44 ± 3.98.

Table 2. Calculated metrics for academic (1-7) and non-academic (8-18) sources. Abbreviations: FKGL (Flesch-Kincaid grade level), FRE (Flesch reading ease), mDISCERN (modified DISCERN)

#	Website	FKGL	FRE	mDISCERN
1	www.my.clevelandclinic.org	8.8	53.1	36.0
2	www.mayoclinic.org	7.5	61.7	36.0
3	www.health.harvard.edu	9.0	60.0	27.5
4	www.health.harvard.edu	9.8	56.4	27.0
5	www.health.harvard.edu	9.2	60.5	29.5
6	www.hopkinsmedicine.org	8.9	58.9	27.0
7	www.massgeneral.org	10.4	52.5	28.0
8	www.heart.org	8.1	62.4	22.0
9	www.svhhearthealth.com	8.6	61.3	26.5
10	www.medlineplus.gov	7.8	64.2	36.0
11	www.webmd.com	6.7	70.1	29.5

12	www.healthline.com	8.4	62.7	33.5
13	www.bhf.org.uk	6.7	73.3	29.5
14	www.heartandstroke.ca	9.0	58.4	27.5
15	www.myhealth.alberta.ca	5.2	77.7	29.5
16	www.heartwest.com.au	9.9	60.1	29.5
17	www.sozocardiology.com	9.5	57.5	28.0
18	www.ascotcardiologygroup.co.nz	7.0	70.8	27.5

When evaluating by website type, academic websites had a greater FKGL than non-academic websites (9.09 ± .90 vs. 7.90 ± 1.40). An Independent Samples t-test found this difference to be significantly different, with a p-value of 0.032 (Figure 1). Similarly, a statistically significant difference in the mean FRE for academic (57.6 ± 3.66) and non-academic (65.32 ± 6.6) sources was found with a p-value of 0.006 (Figure 2). Academic (30.14 ± 4.18) and non-academic (29 ± 3.88) sites had similar mDISCERN scores. For academic websites, mDISCERN scores ranged from 27-36, while non-academic websites mDISCERN scores ranged from 22-36.

Analyzing the 8 individual mDISCERN questions, question 1 had the greatest score amongst all the websites (4.89 ± 0.24). Academic sites (5 ± 0) had a greater score for question 1 compared to non-academic sites (4.81 ± 0.50). Question 4 had the lowest score amongst all websites (1.83 ± 1.23). Academic sites (2 ± 1.30) had a greater score for question 4 versus non-academic (1.73 ± 1.20) sites. Academic sites had a perfect score for question 2 as well, with a mean of 5 ± 0.

Figure 1. Mean FKGL of academic and non-academic sources. The dotted red line represents the AMA's recommended 6th-grade reading level.

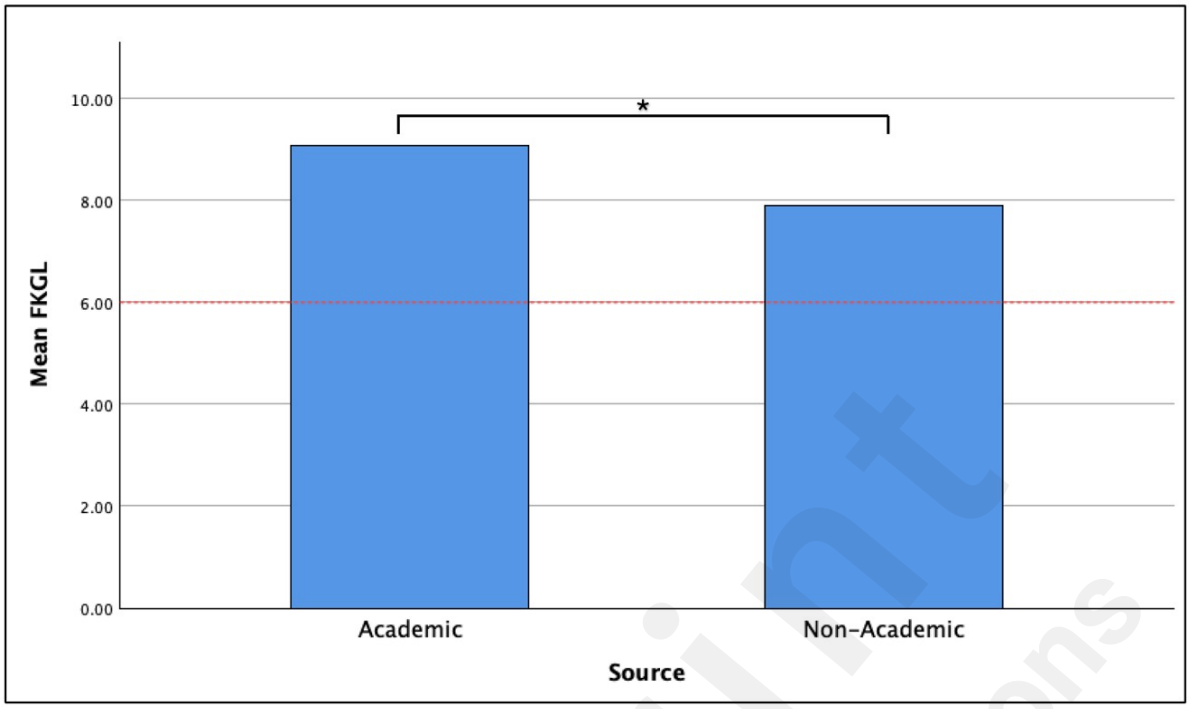
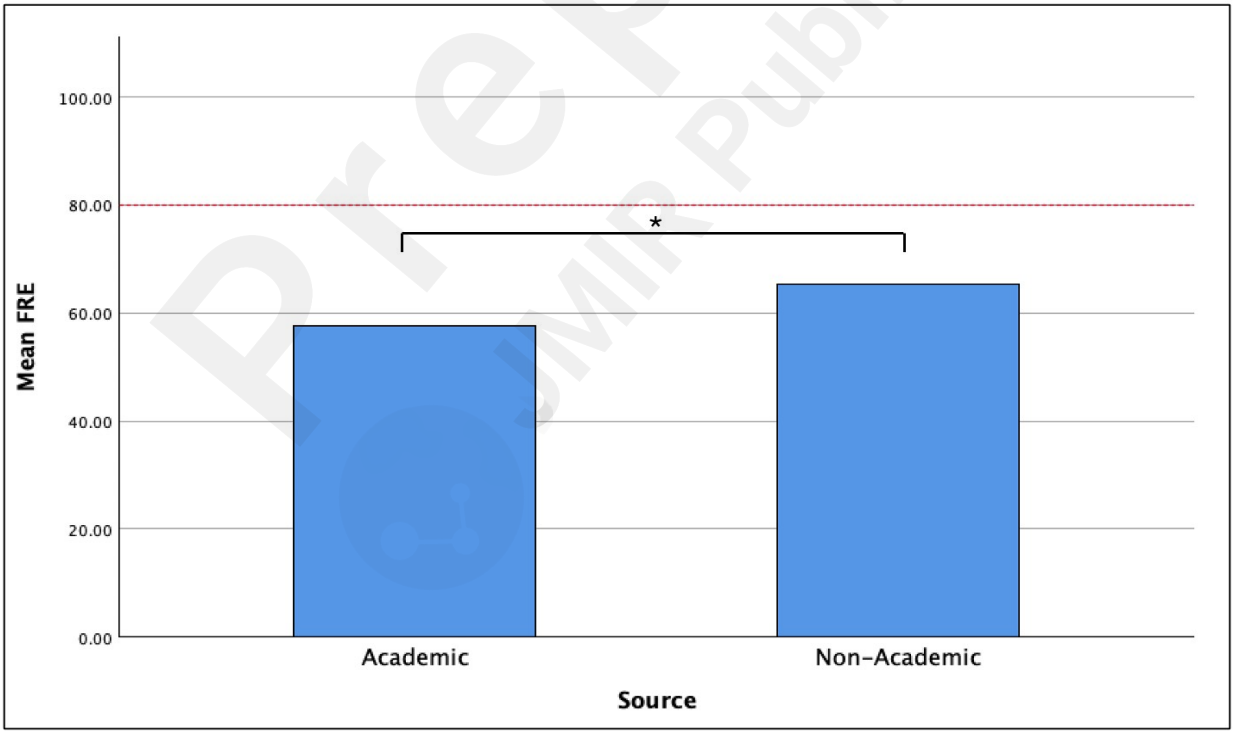


Figure 2. Mean FRE of academic and non-academic sources. The dotted red line represents the equivalent FRE score for a 6th-grade reading level.



Discussion

Our study is the first to assess the readability and quality of web-based sources on exercise stress

testing. The results of our study indicate that the average FKGL for the 18 web pages analyzed is 8.36, exceeding the AMA recommendation of health information being written at a 6th-grade level[11]. Our analysis also shows that academic sources for health information are written at a significantly higher grade level than non-academic sources, with average FKGLs of 9.1 and 7.9, respectively. While the grade level of academic sources is statistically greater, both categories exceed the AMA's recommendation for health information to be written at a 6th-grade level, indicating that even non-academic sources may still be too complex for the average patient to fully understand. Only one of the 18 websites accessed was at or below the recommended reading level with a FKGL of 5.2, with eleven above the 8th-grade reading level.

Being that the FRE formula is derived from the FKGL formula, it follows that the average FRE for academic sources was statistically lower than for non-academic sources. An FRE score of 80-90 is considered appropriate for a 6th-grade reading level[22,23]. In contrast, the average FRE scores for academic and non-academic sources were 57.6 and 65.3, respectively. Although this difference was statistically significant, both scores indicate reading difficulty levels that exceed the recommended ease of 80-90. These findings highlight that both academic and non-academic institutions are producing resources on exercise stress tests that fail to meet current readability recommendations, potentially hindering patients' ability to fully understand their health information. This lack of concordance with recommended readability standards may limit patients' capacity to make informed decisions about their care, thereby undermining their participation in shared decision-making and their ability to manage their health effectively.

Our analysis also focused on the quality of web resources using the mDISCERN criteria. As described in our methods, web resources with a mDISCERN score below 40% classified the reliability as "poor." Scores between 40% and 79% were deemed "fair," and scores above 80% were categorized as "good." The average mDISCERN score for all web resources included in our analysis was 29.44 out of 40 (74%). This classifies the average quality of web resources as "fair" and closer to "good" than "poor." 14 of the web resources were rated as "fair," and the remaining four were rated as "good." Every web resource analyzed lost points when evaluated on questions four, five, seven, and eight. Questions four and five pertain to clearly citing the sources used in the article and providing publication dates for both the article and its references. Questions seven and eight assess whether the author offers resources for further information and addresses potential uncertainties,

such as the risks associated with the test. We considered these areas crucial for assessment, as they were consistently lacking across all the web resources reviewed. By addressing these deficiencies, the overall quality of these resources can be significantly enhanced, providing patients with more reliable and comprehensive information.

The available web resources for patients seeking information on exercise stress tests are, on average, of fair quality but exceed the recommended reading level when assessed using the FKGL and FRE scores. These findings are consistent with other studies analyzing the readability of cardiovascular topics[13-15], suggesting that many online health resources related to cardiovascular medicine may be difficult for the average patient to understand, potentially limiting their accessibility and effectiveness in promoting patient education. Authors of both academic and non-academic resources could improve their content by consulting the AMA's guidelines on best practices for creating patient-friendly health information[11]. In particular, attention should be given to factors that reduce reading ease and elevate the reading grade level.

Limitations

As mentioned, the FKGL and FRE are the most widely used readability formulas in healthcare literature[18]. However, limitations do exist. The Flesch formulas (FKGL and FRE) focus on word count, sentence length, and syllable count[22,23]. These formulas are not capable of assessing the complexity or technical nature of words, which is particularly relevant in healthcare, where terms related to diseases, treatments, and tests are inherently complex. Alternative language can be used when appropriate, however, avoiding vital terminology to lower the reading grade level may compromise the accuracy and clarity of the information provided. Simplifying medical jargon might increase readability, but at the potential cost of misinforming patients.

Additional limitations arise when using the DISCERN tool for evaluating articles. Scoring is inherently subjective, requiring an evaluator to assess a web page and rate each question accordingly. While this non-formulaic approach allows for a more nuanced evaluation, individual evaluator bias and interpretation can affect the consistency and standardization of scores.

Conclusion

The web resources available to patients seeking information regarding exercise stress tests is written at a reading level that surpasses the current AMA recommendation. This is complicated by the inherently complex nature of medical terminology. Despite this, efforts should be made by authors to use alternative terms or omit terms altogether if the medical jargon being used does not contribute to the message attempting to be conveyed. The quality of the web resources evaluated were determined to be fair but could be improved by providing more comprehensive information. While improving health literacy across the population is crucial, it is equally important to provide patients with clear, accessible materials that empower them to make informed decisions about their health. These materials must also be high-quality, offering relevant, unbiased, and concise information, along with appropriate access to sources and additional resources for those seeking further guidance.

Conflict of Interest Statement: None declared

Abbreviations:

American Medical Association (AMA)

modified DISCERN (mDISCERN)

Flesch-Kincaid Grade Level (FKGL)

Flesch Reading Ease (FRE)

References:

1. Sidoti OG-W, R; Faverio, M; Atske, S; Radde, K; Park, E;. Internet, Broadband Fact Sheet. Updated February 2024. Accessed August, 2024. https://www.pewresearch.org/internet/fact-sheet/internet-broadband/

2. Wang XC, RA;. Health Information Technology Use Among Adults: United States, July - December 2022. *CDC Stacks Public Health Publications*. 2023;(482)doi:https://dx.doi.org/10.15620/cdc:133700

3. Prestin A, Vieux SN, Chou WY. Is Online Health Activity Alive and Well or Flatlining? Findings From 10 Years of the Health Information National Trends Survey. *J Health Commun*. 2015;20(7):790-8. doi:10.1080/10810730.2015.1018590

4. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. *Arch Intern Med*. Jul 23 2007;167(14):1503-9. doi:10.1001/archinte.167.14.1503

5. Bostock S, Steptoe A. Association between low functional health literacy and mortality in older adults: longitudinal cohort study. *BMJ: British Medical Journal*. 2012;344(7852):15-15.

6. Grandinetti DA. Doctors and the Web. Help your patients surf the Net safely. *Med Econ*. Mar 6 2000;77(5):186-8, 194-6, 201.

7. Sullivan B, Abed V, Joiner J, et al. The Quality of Online Information for the Treatment of Knee Osteoarthritis: A Google Study. *Cureus*. Oct 2022;14(10):e29995. doi:10.7759/cureus.29995

8. De La Chapa JS, Bellinger JR, Timm AR, Daniero JJ. "Quality, Readability, and Understandability of Online Posterior Glottic Stenosis Information". *J Voice*. Dec 28 2022;doi:10.1016/j.jvoice.2022.11.036

9. Arts H, Lemetyinen H, Edge D. Readability and quality of online eating disorder information-Are they sufficient? A systematic review evaluating websites on anorexia nervosa using DISCERN and Flesch Readability. *Int J Eat Disord*. Jan 2020;53(1):128-132. doi:10.1002/eat.23173

10. Oman SP, Zaver H, Waddle M, Corral JE. Quality and Readability of Web-Based Information for Patients With Pancreatic Cysts: DISCERN and Readability Test Analysis. *JMIR Cancer*. Mar 16 2021;7(1):e25602. doi:10.2196/25602

11. Barry DW. *Health literacy and patient safety: Help patients understand. Manual for clinicians*. 2nd ed. American Medical Association Foundation; 2007.

12. Garner KP, W; Arnold, JJ. Exercise Stress Testing: Indications and Common Questions.

American Family Physician. 2017;96(5):293-299.

13. Siddiqui E, Shah AM, Sambol J, Waller AH. Readability Assessment of Online Patient Education Materials on Atrial Fibrillation. *Cureus*. Sep 11 2020;12(9):e10397. doi:10.7759/cureus.10397

14. Wasir AS, Volgman AS, Jolly M. Assessing readability and comprehension of web-based patient education materials by American Heart Association (AHA) and CardioSmart online platform by American College of Cardiology (ACC): How useful are these websites for patient understanding? *Am Heart J Plus*. Aug 2023;32:100308. doi:10.1016/j.ahjo.2023.100308

15. Ayyaswami V, Padmanabhan D, Patel M, et al. A Readability Analysis of Online Cardiovascular Disease-Related Health Education Materials. *Health Lit Res Pract*. Apr 2019;3(2):e74-e80. doi:10.3928/24748307-20190306-03

16. GlobalStats S. Search Engine Market Share Worldwide. https://gs.statcounter.com/search-engine-market-share

17. Insights C. The Value of Google Result Positioning. Chitika, Inc. Accessed August, 2024. https://research.chitika.com/wp-content/uploads/2022/02/chitikainsights-valueofgoogleresultspositioning.pdf

18. Wang L-W, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: Application, results, and recommendations. *Research in Social and Administrative Pharmacy*. 2013/09/01/ 2013;9(5):503-516. doi:https://doi.org/10.1016/j.sapharm.2012.05.009

19. Kincaid JR, RL; Fishburne, RP; Chissom, BS. *Derivation of New Readability Formulas (Automated Readability Index, FOG count, and Flesch Reading Ease Formula) For Navy Enlisted Personnel*. 1975. https://apps.dtic.mil/sti/pdfs/ADA006655.pdf

20. Administration SS. What is the Flesch-Kincaid Readbaility Test? Social Security Administration of the United States of America. Accessed August, 2024. https://secure.ssa.gov/poms.nsf/lnx/0910605105

21. Stockmeyer N. Using Microsoft Word's Readability Program. *Michigan Bar Journal*. 2009:46-47.

22. Flesch R. A new readability yardstick. *J Appl Psychol*. Jun 1948;32(3):221-33. doi:10.1037/h0057532

23. Flesch R. *How to Write Plain English*. 1st ed. Barnes & Noble; 1981:144.

24. Griffiths KM, Christensen H. Website Quality Indicators for Consumers. *J Med Internet Res*. 2005/11/15 2005;7(5):e55. doi:10.2196/jmir.7.5.e55

25. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health*. 1999;53(2):105-111. doi:10.1136/jech.53.2.105

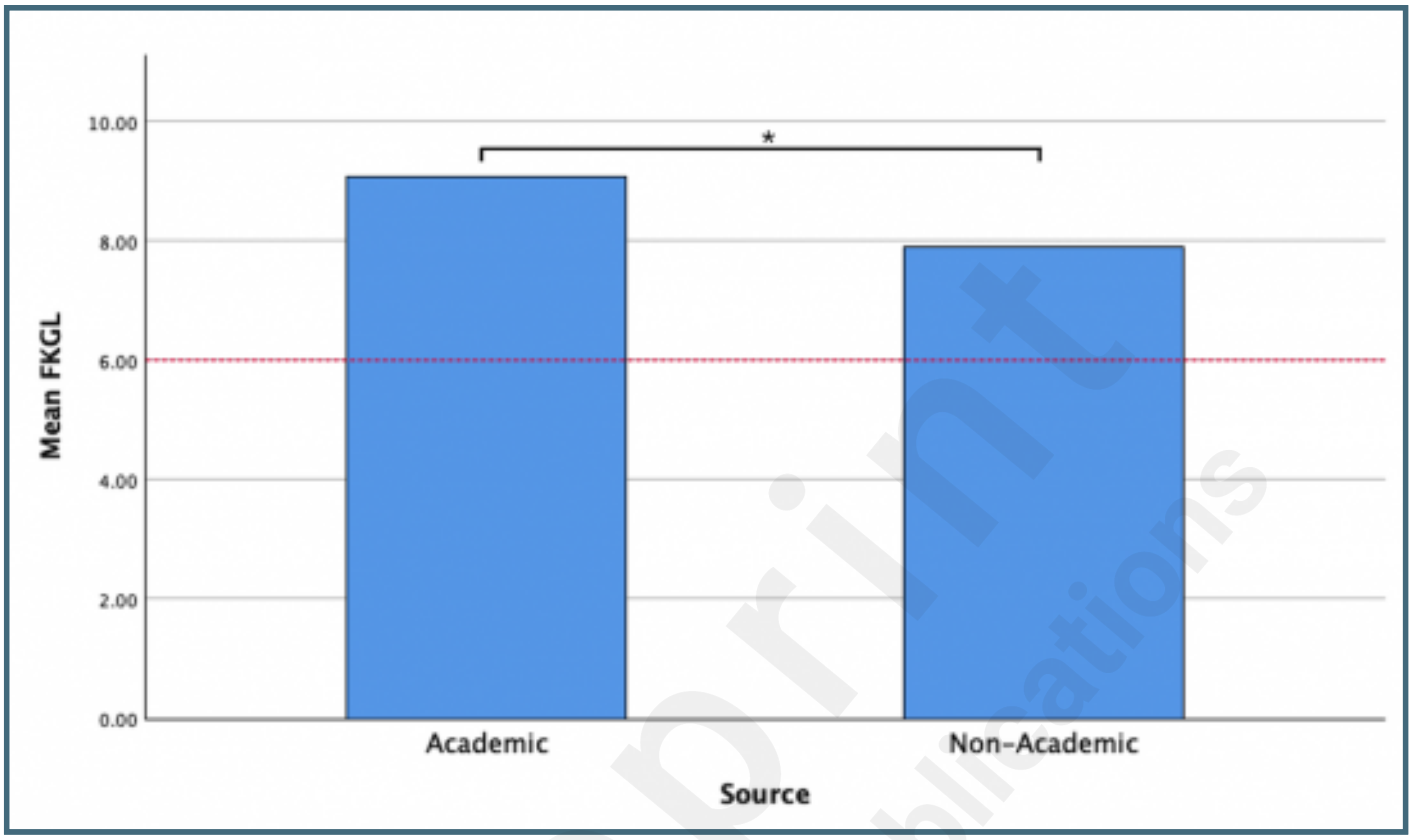
26. Kumar VS, Subramani S, Veerapan S, Khan SA. Evaluation of online health information on clubfoot using the DISCERN tool. *J Pediatr Orthop B*. Mar 2014;23(2):135-8. doi:10.1097/bpb.0000000000000000

27. Onder CE, Koc G, Gokbulut P, Taskaldiran I, Kuskonmaz SM. Evaluation of the reliability and readability of ChatGPT-4 responses regarding hypothyroidism during pregnancy. *Sci Rep*. Jan 2 2024;14(1):243. doi:10.1038/s41598-023-50884-w

Supplementary Files

Figures

Mean FKGL of academic and non-academic sources. The dotted red line represents the AMA's recommended 6th-grade reading level.



Mean FRE of academic and non-academic sources. The dotted red line represents the equivalent FRE score for a 6th-grade reading level.