Systematic Review

Patient Education Materials Found via Google Search for Shoulder Arthroscopy Are Written at Too-High of a Reading Level

Youssef Abdullah, B.S., Aaron Alokozai, M.D., Abraham J. Mathew, B.S.A., Michaela A. Stamm, M.S., and Mary K. Mulcahey, M.D., F.A.A.O.S., F.A.O.A.

Purpose: To evaluate the quality and correlation of readability on actionability and understandability of shoulder arthroscopy-related patient education materials (PEMs) found via a routine Google search. **Methods:** Two independent authors performed an online Google search with the term "shoulder arthroscopy." The first 5 pages of search results were then screened for PEMs. Journal articles, news articles, nontext materials, and unrelated websites were excluded. The readability of included resources was calculated using objective metrics: Flesch–Kincaid Grade Score, Simple Measure of Gobbledygook index, Coleman–Liau Index, and the Gunning Fog Index. Patient Education Material Assessment Tool for Printed Materials assessed for understandability and actionability. Associations between readability and actionability and understandability were determined using Spearman correlation and linear regression. **Results:** The searches returned 53 websites related to shoulder arthroscopy. A total of 34 (64%) met inclusion criteria. A high school reading level or greater was required to read the average PEM according to all scales used. The average PEM received a Patient Education Material Assessment Tool for Printed Materials score of 61.33 in understandability (range 18.75-89.47) and 55.59 points in actionability (range 16.67-83.33). An easily understood or actionable article would score at least 70 points. A moderate correlation was observed between readability and actionability on three of the scales used ($r = 0.5$, $r = 0.59$, $r = 0.61$). **Conclusions:** Most shoulder arthroscopy PEMs identified on Google are not written at a level that the average patient can read, understand, or act on (actionability). **Clinical Relevance:** Orthopaedic surgeons should be aware of the resources that patients use to obtain medical information. More accessible PEMs should be developed for patients undergoing shoulder arthroscopy to enhance comprehension of their condition and improve shared decision-making.

Shoulder arthroscopy is a commonly performed procedure used for diagnosing and treating a wide variety of shoulder conditions, including shoulder instability, shoulder stiffness, and rotator cuff tears. Patient education is integral to understanding the risks, benefits, and attributes of care associated with shoulder arthroscopy. Patients seeking additional information may turn to the internet for patient education materials (PEMs).[1]

PEMs have been the focus of studies within various specialties, and Roberts et al. introduced the evaluation of PEMs in the context of general orthopaedics.[2,3] Recent studies have focused solely on PEMs located on the American Academy of Orthopaedic Surgeons website.[4] For example, pediatric orthopaedic PEMs that are produced by the American Academy of Orthopaedic Surgeons have readability scores that exceed recommendations.[5] However, many patients acquire health information online using search engines. The overall quality of these PEMs remains unclear. The purpose of this study was to evaluate the quality and correlation of

From the Department of Orthopaedic Surgery, Tulane University School of Medicine, New Orleans, Louisiana (Y.A., A.A., M.A.S., M.K.M.); and Baylor College of Medicine, Houston, Texas (A.J.M), U.S.A.

The authors report the following potential conflicts of interest or sources of funding: M.K.M reports personal fees and nonfinancial support from Arthrex and Alon Medical Technology, outside the submitted work. She is a board or committee member of AAOS, American Orthopaedic Association, and the American Orthopaedic Society for Sports Medicine. She is on the editorial or governing board of the American Journal of Sports Medicine, Electronic Media, *and* Arthroscopy. *Full ICMJE author disclosure forms are available for this article online, as supplementary material.*

Study performed at the Tulane University School of Medicine, New Orleans, Louisiana, U.S.A.

Received January 24, 2022; accepted April 27, 2022.

Address correspondence to Mary K. Mulcahey, M.D., F.A.A.O.S., F.A.O.A., Department of Orthopaedic Surgery, Tulane University School of Medicine, 1430 Tulane Ave., New Orleans, LA, 70112. E-mail: mary.mulcahey.md@gmail.com

2666-061X/22131

https://doi.org/10.1016/j.asmr.2022.04.034

readability on actionability and understandability of shoulder arthroscopy-related PEMs found via a routine Google search. We hypothesized that the quality of shoulder arthroscopy PEMs identified on Google would be inadequate for the average patient.

Methods

Two investigators (Y.A. and A.M.) searched Google in January 2022 using the search term "shoulder arthroscopy." The first 5 pages of Google search results were included in the study. Duplicates and inaccessible websites were excluded. This strategy was designed to mimic a typical patient's Google search for shoulder arthroscopy information. No other search engines were used in our study. Heap et al.[6] reported that readability and objective quality of information obtained online did not differ among search engines.

Assessment of Readability

The Flesch–Kincaid Grade (FKG) score, the Simple Measure of Gobbledygook (SMOG) index, the Gunning Fog Index (GFI), and the Coleman–Liau Index (CLI) were used to calculate the readability of the websites, which contained PEMs. Each website was independently scored by 2 coauthors and, to avoid human error and promote consistency, an open-source readability software was used as a secondary check for the calculated human results. Author's names, abbreviations, nonstandard texts, and hyperlinks were not included in the analysis to prevent skewing of the results.

FKG Scores

The FKG score is a commonly used test that determines readability based on the weighting of syllables and can be calculated using the FKG formula: $0.39 \times$ (average number of words per sentence) $+ 11.8 \times$ (average number of syllables per word) $- 15.59$.[7] This formula has been widely used for determining grade-level readability in similar studies, and it has even been incorporated into many programs, such as Microsoft Word.[7,8] In this study, 2 coauthors determined the FKG reading level of each website by transferring its text into a Microsoft Word document (Microsoft, Redmond, WA). From there, the text was highlighted and under the "Tools" tab, the "Spelling and Grammar" function was selected before choosing the "Options" bar at the bottom of the window and enabling the "Show Readability Statistics" function. In this way, Microsoft Word automatically calculated the FKG reading level of the 34 websites. The mean, standard deviation, and 95% confidence intervals for FKG reading level were calculated using Microsoft Excel. The FKG score was also calculated using the online readability tool, with the resulting number providing an approximate U.S. grade-level equivalent. A score of less than eight is considered acceptable for universal readability and accessibility.[9]

Simple Measure of Gobbledygook

The SMOG readability formula uses a hand-scored method that allows an evaluator to determine each patient education website's grade level by counting ten sentences at the beginning, the middle, and the end of each website. SMOG is only used for the English language and has been validated for studies regarding healthcare information.[10] The evaluator counts every word of 3 or more syllables in those 30 sentences. Afterwards, the total number of words counted is added before being used in the SMOG conversion table to find the grade level.[11] SMOG scores were calculated by averaging the 2 coauthors' independent scores and cross-checking with an online readability tool.

Coleman–Liau Index

The CLI is a readability metric that relies on characters per word rather than syllables per word. An online readability tool (readable.com) determined CLI for each PEM after it was copied and pasted on the website. The online readability tool was used due to the complexity of manually calculating CLI.[5]

Gunning Fog Index

The GFI measures readability by estimating the number of years of formal education that a reader would need to comprehend the materials being presented. The GFI uses the lengths of words and sentences to determine a text's difficulty and is calculated using the following formula: $0.4 \times$ [(words/sentences) $+ 100 \times$ (complex words/ total words)]. A lower GFI value indicates that a text is easier to read. The GFI was calculated and cross-checked with an online readability tool. A GFI of less than eight is considered appropriate for universal readability.[12]

Assessment of Actionability and Understandability

Patient Education Material Assessment Tool

The Patient Education Material Assessment Tool for Printed Materials (PEMAT-P) was developed by the Agency for Healthcare Research and Quality to ensure that PEMs satisfy the informational requirements of learners with a wide range of literacy abilities and challenges.[13,14] The 2 metrics that the PEMAT measures are actionability and understandability. The former is defined as a learners' ability to identify what actions can be taken based on the educational materials available and is determined by seven items with the following responses: "yes," "no," or "not applicable."[15] Understandability is the ability of individuals from diverse backgrounds and health-related knowledge to comprehend educational materials and draw critical conclusions from them. It is determined based on 17

items organized under the following categories: content area, word style and choice, number usage, design and layout, organization, and visual aids.[13] An easily understood or actionable article would score at least 70 points.

In this study, PEMAT-P was assessed individually by 2 coauthors with an interobserver reliability of 0.96, following the Agency for Healthcare Research and Quality guidelines via Microsoft Excel. This scale was used to ascertain whether the 34 websites out of the original 53 were both understandable and promoted patient actionability.

Statistical Analysis

To better describe the reading levels observed in this study, the average reading levels within the FKG, CLI, SMOG, and GFI were calculated. In addition, the standard deviation (SD), minimum, and maximum values for reading levels across each of the 4 scales were determined. When applicable, the average reading level values were recalculated after outliers were removed from the data set to provide a more accurate measure of a patient's experience. The analyses performed on the reading level scales also were conducted on both understandability and actionability, leading to the calculation of average values, SDs, minimums, and maximums for each as well as the removal of outliers when applicable with recalculation of the mean afterwards. Finally, Pearson correlations were calculated to assess the association between readability and understandability and the association between readability and actionability.

Results

Of the original 53 search results identified related to shoulder arthroscopy, 34 (64%) were determined to be PEMs (Appendix Table 1, available at www.arthroscopyjournal.org).

Readability

The readability of each was assessed using four validated scales (i.e., FKG score, SMOG index, GFI, and CLI). These scales reflected the U.S. school grade level required to read the materials, with values greater than 12 indicating college-level reading materials (Appendix Table 2, available at www.arthroscopyjournal.org).

According to FKG, on average, each site required a 10th-grade reading level or higher (score: 10.12), with an SD of 1.95. The highest reading level required of a patient by any of these sites was college-level (score: 13.6), whereas the least-demanding required a nearly fifth-grade reading level (score: 4.9). However, as seen in Table 1, the latter appeared to be an outlier, with the second lowest FKG score being 7.4, indicating a reading level slightly above seventh grade. When the outlier was removed from the data set, the mean FKG score increased to 10.27.

According to the SMOG index, the average reading level required was between the ninth and 10th grade (score: 9.58, SD: 1.92). In this case, the highest reading level needed was college-level (score: 13.60), with a seventh-grade reading level being the lowest required from a patient (score: 7.00). Using the CLI, the average requirement was a college reading level (score: 12.97, SD: 2.35). The highest required reading level exceeded the college level (score: 18.00), while a seventh-grade reading level was the lowest reading level required by a site. Finally, evaluation using the GFI returned an average reading level of approximately 12th grade (score: 12.14, SD: 2.71). The highest reading level observed was beyond college-level (score: 16.30), and the lowest was around a sixth-grade reading level (score: 6.3).

Understandability

The 34 articles included in this study were also assessed for understandability using the PEMAT-P tool, for which the score can range from 0 to 100. On average, the websites scored 61.33 for understandability, with the most understandable site scoring 89.47 and the least understandable site scoring 18.75 (SD 14.25 points). Scores less than 70 indicate unacceptable understandability (Appendix Table 3, available at www.arthroscopyjournal.org).

Actionability

Actionability was the final measure determined for the included websites. Like understandability, actionability was also assessed using the PEMAT-P tool. On average, the score was 55.59 for actionability (SD 28.29 points, range 0-100). Scores less than 70 indicate unacceptable actionability, meaning it is difficult to identify what can be done based on the information presented (Appendix Table 4, available at www.arthroscopyjournal.org).

Table 1. Readability Results as Obtained on the Different Scales Used

Scale	Average Reading Levels	Minimum	Maximum	Standard Deviation
Flesch–Kincaid Grade Scale	10.12	4.90	13.50	1.95
Simple Measure of Gobbledygook	9.58	5.90	13.60	1.92
Coleman–Liau Index	12.97	7.00	18.00	2.36
Gunning Fog Index	12.14	6.30	16.30	2.71

Table 2. The Correlation Coefficients Between Readability and PEMAT-P Scores (Understandability and Actionability)

Reading Level	Understandability	Actionability
Flesch–Kincaid Grade Score	−0.28	0.50
Coleman–Liau Index	−0.23	0.24
Gunning–Fog Index	−0.24	0.59
Simple Measure of Gobbledygook (SMOG)	−0.31	0.61

PEMAT-P, Patient Education Materials Assessment Tool for Printable Materials.

Correlation of Readability on Actionability and Understandability

A weak negative correlation was observed between understandability and all reading level scales. In addition, a moderate correlation between reading level and actionability was observed across all the reading level scales except for the CLI, which held a weak correlation with actionability. The correlations between reading level and understandability and between reading level and actionability are summarized in Table 2.

Discussion

Our analysis found that the average PEM related to shoulder arthroscopy within the first 53 search results, as analyzed by 4 validated readability scales, required a high school reading level. Further, the results reviewed in this study found that these 53 articles are not easily understood or actionable.

According to the American Medical Association and National Institutes of Health guidelines, a sixth-grade reading level is required for PEMs. This suggests that most shoulder arthroscopy PEMs found on Google are not written at a level that the average patient can easily read, which is supported by findings in other medical specialties. For instance, in a study conducted on a random sample of PEMs provided by the National Library of Medicine, which included a wide array of medical specialties, the authors found that the average reading level for Medline, Elton B. Stephens Company, and Micromedex were all greater than an eighth-grade reading level on each of the reading level metrics used.[16] Future PEMs related to shoulder arthroscopy should be developed at an appropriate reading level to enhance patient comprehension of their condition and improve shared decision-making.

The average site providing PEMs related to shoulder arthroscopy is not easily understandable or actionable for the typical patient, meaning it is challenging for patients of diverse backgrounds and health literacy levels to understand the article and identify what they can do based on the information presented. In terms of understandability, most of the sites fared similarly, indicated by data that clusters around the mean. In contrast, the SD in the case of actionability (28.29) was nearly twice that of understandability, indicating a greater spread between studies. Our findings corroborate a prior study, which found low understandability and actionability in the American Association of Neurological Surgeons database and in Medline.[17] Unacceptable quality of the articles was not confined to PEMs related to shoulder arthroscopy, but also indicate that improvement is required in multiple medical fields.[15,17] Patients need to be able to understand the information they are reading to select the treatment (actionability) that aligns with their goals, values, and preferences. If patients have trouble comprehending risks, benefits, and attributes associated with different treatment options, expectations may not be met. Creation of PEMs that are easy to understand may help guide some patients to nonoperative management if they don't think they will gain meaningful improvement from surgery. This can only be achieved if patients understand the attributes associated with surgery and if it aligns with their personal health goals.

Actionability was moderately, positively correlated with reading levels on all of the scales except for the CLI ($r > 0.5$, Table 2). Three of 4 reading level scales used in this study correlated with actionability, indicating that less readable PEMs are easier to act on. This implies that patients with a lower reading level are at a disadvantage utilizing these tools. PEMs for shoulder arthroscopy, although unintentionally, favor those with higher education levels. This subsequently impacts patient care because low health literacy has been shown to correlate with poorer health outcomes.[2] This potentially acts as a barrier for patient involvement with their medical care. These findings should be investigated more broadly in orthopaedics.

Limitations

There are several limitations to this study. First, the present results are based only on the included websites and do not reflect PEMs that were not captured in this study. In addition, findings are limited by the scales (FKG, CLI, GFI, SMOG, and PEMAT) used to measure them. Human error may have affected the application of these scales during data collection.

Conclusions

Most shoulder arthroscopy PEMs identified within the first 53 search results on Google are not written at a level that the average patient can read, understand, or act on (actionability).

References

1. Delaney FT, Doinn T, Broderick JM, Stanley E. Readability of patient education materials related to radiation safety: What are the implications for patient-centred radiology care? *Insights Imaging* 2021;12:1-9.

2. Roberts H, Zhang D, Dyer GSM. The readability of AAOS patient education materials: Evaluating the progress since 2008. *J Bone Joint Surg Am* 2016;98:e70.
3. Fortuna J, Riddering A, Shuster L, Lopez-Jeng C. Assessment of online patient education materials designed for people with age-related macular degeneration. *BMC Ophthalmol* 2020;20:1-7.
4. Horner NS, de Sa D, Heaven S, et al. Indications and outcomes of shoulder arthroscopy after shoulder arthroplasty. *J Shoulder Elbow Surg* 2016;25:510-518.
5. Ó Doinn T, Broderick JM, Abdelhalim MM, Quinlan JF. Readability of patient educational materials in pediatric orthopaedics. *J Bone Joint Surg Am* 2021;103:e47.
6. Heap JC, Dezfuli B, Bennett DM, Chapman E, DeSilva GL. The internet as a source of information for De Quervain's tendinitis. *Hand (N Y)* 2015;10:131-136.
7. Cooley ME, Moriarty H, Berger MS, Selm-Orr D, Coyle B, Short T. Patient literacy and the readability of written cancer educational materials. *Oncol Nurs Forum* 1995;22:1345-1351.
8. Albright J, De Guzman C, Acebo P, Paiva D, Faulkner M, Swanson J. Readability of patient education materials: Implications for clinical practice. *Appl Nurs Res* 1996;9:139-143.
9. Cotugna N, Vickery CE, Carpenter-Haefele KM. Evaluation of literacy level of patient education pages in health-related journals. *J Community Health* 2005;30:213-219.
10. McLaughlin GH. SMOG grading: A new readability formula. *J Read* 1969;12:639-646.
11. Fitzsimmons PR, Michael BD, Hulley JL, Scott GO. A readability assessment of online Parkinson's disease information. *J R Coll Physicians Edinb* 2010;40:292-296.
12. Keogh CJ, McHugh SM, Clarke Moloney M, et al. Assessing the quality of online information for patients with carotid disease. *Int J Surg* 2014;12:205-208.
13. Shoemaker SJ, Wolf M, Brach C. The Patient Education Materials Assessment Tool (PEMAT) and User's Guide. Agency for Healthcare Research and Quality Publications, 2014. Accessed November 7, 2021. https://www.ahrq.gov/health-literacy/patient-education/pemat.html
14. Schorr C, Hunter K, Zuzelo PR. Understandability and actionability of the CDC's printable sepsis patient education material. *Am J Crit Care* 2018;27:418-427.
15. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): A new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ Couns* 2014;96:395-403.
16. Stossel LM, Segar N, Gliatto P, Fallar R, Karani R. Readability of patient education materials available at the point of care. *J Gen Intern Med* 2012;27:1165-1170.
17. Lopez Ramos C, Williams JE, Bababekov YJ, Chang DC, Carter BS, Jones PS. Assessing the understandability and actionability of online neurosurgical patient education materials. *World Neurosurg* 2019;130:e588-e597.