SCIENTIFIC ARTICLE

# Assessing the readability of patient-targeted online information on musculoskeletal radiology procedures

Phuong T. Duong[1] · Matthew P. Moy[1] · F. Joseph Simeone[2] · Connie Y. Chang[2] · Tony T. Wong[1]

## Abstract

**Objective** To assess the readability of patient-targeted online information on musculoskeletal radiology procedures.

**Methods** Eleven common musculoskeletal radiology procedures were queried in three online search engines (Google, Yahoo!, Bing). All unique patient-targeted websites were identified ($n = 384$) from the first three pages of search results. The reading grade level of each website was calculated using 6 separate validated metrics for readability assessment. Analysis of word and sentence complexity was also performed. Results were compared between academic vs. non-academic websites and between websites found on different pages of the search results. Statistics were performed using a $t$ test.

**Results** The mean reading grade level across all procedures was 10th–14th grade. Webpages for nerve block were written at a higher reading grade level on non-academic websites ($p = 0.025$). There was no difference in reading grade levels between academic and non-academic sources for all other procedures. There was no difference in reading grade levels between websites found on the first page of search results compared with the second and third pages. Across all websites, 16–22% of the words used had 3+ syllables and 31–43% of the words used had 6+ characters (complex words); 13–24% of the sentences used had 22+ words (complex sentences).

**Conclusion** Patient-targeted online information on musculoskeletal radiology procedures are written at the 10th–14th grade reading level, which is well beyond the AMA and NIH recommendation. Readability can be lowered by decreasing text complexity through limitation of high-syllable words and reduction in word and sentence length.

✉ Tony T. Wong
ttw2105@cumc.columbia.edu

Phuong T. Duong
ptd2107@cumc.columbia.edu

Matthew P. Moy
mm5054@cumc.columbia.edu

F. Joseph Simeone
fsimeone@mgh.harvard.edu

Connie Y. Chang
cychang@mgh.harvard.edu

[1]    Division of Musculoskeletal Radiology, New York Presbyterian Hospital–Columbia University Irving Medical Center, 622 W 168th Street, New York, NY 10032, USA

[2]    Department of Radiology, Division of Musculoskeletal Imaging and Intervention, Massachusetts General Hospital, 55 Fruit St., Boston, MA 02114, USA

## Introduction

Health literacy is a public health concern, the severity of which varies based on age, education level, and socioeconomic status. According to the US Department of Health and Human Services, 77 million adults have basic or below basic health literacy [1]. Those who fall within this group can experience medical complications such as increased hospitalizations and higher mortality rates, even when accounting for confounders such as age, sex, and insurance status [2–4].

Many patients are increasingly using Internet resources to supplement their knowledge about medicine [1]. Therefore, it is necessary for electronic health information to be conveyed in an effective manner to patients of all literacy levels. According to the American Medical Association (AMA), the majority of Americans read at or below the 8th grade level. Consequently, the AMA and the National Institutes of Health (NIH) recommend that patient educational resources be written at a 3rd–7th

grade level [5, 6]. Unfortunately, previous studies have shown that patient resources for radiology tests and procedures are written at a much higher reading level [7–10]. As one example, a study by Hansberry et al. found that articles on RadiologyInfo.org, a patient-targeted resource sponsored by the American College of Radiology (ACR) and the Radiological Society of North America (RSNA), were written at the 10th—14th grade level [7].

A wide range of interventional musculoskeletal procedures are performed for both diagnostic and treatment purposes. To our knowledge, there has been no prior comprehensive study evaluating the readability of information on musculoskeletal procedures. Given that patients have a general need for more information prior to undergoing procedures, it is important to ensure that accessible online resources are written appropriately [11]. The purpose of this study was to assess the readability of patient-targeted online information on musculoskeletal radiology procedures. We hypothesized that the reading grade level of online material is higher than the AMA and NIH recommendation.

## Materials and methods

### Internet data extraction

Several search terms were used to identify common musculoskeletal radiology procedures (Table 1). Each term was input into Google, Yahoo!, and Bing search engines, which facilitate 98% of the internet searches performed in the USA [12]. Uniform resource locators (URLs) from the first 3 pages of results for each search engine were recorded. Searches were performed on the same day to ensure that there were no day-to-day changes in the results.

### Website selection

The first 3 pages of search results for each term yielded (mean ± standard deviation) 29.5 ± 0.76 websites from Google, 29.6 ± 0.99 from Yahoo!, and 34.5 ± 2.66 from Bing. Each website was manually screened with application of exclusion criteria as shown in Fig. 1. Patient-targeted websites were often marked as such; for example, they had words like "for patients" in their header. They were also verified when content was specifically directed at patients. This included text with frequent use of "you" as the subject, phrases such as "your doctor," and question and answer formatting or subheadings using the pronoun "I," such as "how do I prepare for my procedure." Mainstream media articles were classified as patient-targeted since their audience is broad and they are not directed towards physicians. Some websites had multiple hyperlinks to separate webpages pertaining to the same topic. For these, all of the separate webpages were included as one text. One of the websites had videos but was included because associated transcripts were available. This yielded 384 unique patient-targeted websites for analysis, which were then classified as being from academic or non-academic sources. Academic websites included those produced by academic centers or organizations such as the American College of Radiology (ACR). A fellowship-trained musculoskeletal radiologist with 8 years of radiology experience independently verified the appropriateness of the websites for inclusion.

### Readability analysis

The text from each of the 384 websites was copied into individual documents (Microsoft Word, Office Professional Plus 2016). All references, headings, and formatting were removed. Punctuation was added to all bullet points. Each document was then evaluated using a readability text analysis

**Table 1** Search terms used for each procedure type

| Procedure type | Search terms |
| --- | --- |
| Arthrogram | "arthrogram injection" |
| Soft tissue and bone biopsy | "ultrasound guided soft tissue biopsy", "CT guided bone biopsy" |
| Bone marrow biopsy | "fluoroscopic guided bone marrow biopsy", "CT guided bone marrow biopsy" |
| Calcific tendinitis treatment | "calcium barbotage", "calcium aspiration" |
| Morton neuroma treatment | "Morton neuroma injection", "Morton neuroma ablation" |
| Needle fenestration | "needle fenestration", "dry needling" |
| Nerve block | "nerve injection", "perineural block" |
| Osteoid osteoma treatment | "osteoid osteoma ablation" |
| Platelet rich plasma treatment | "platelet rich plasma injection" |
| Prolotherapy | "prolotherapy" |
| Therapeutic injection and aspiration | "joint steroid injection", "tendon sheath steroid injection", "ganglion cyst aspiration", "joint aspiration" |
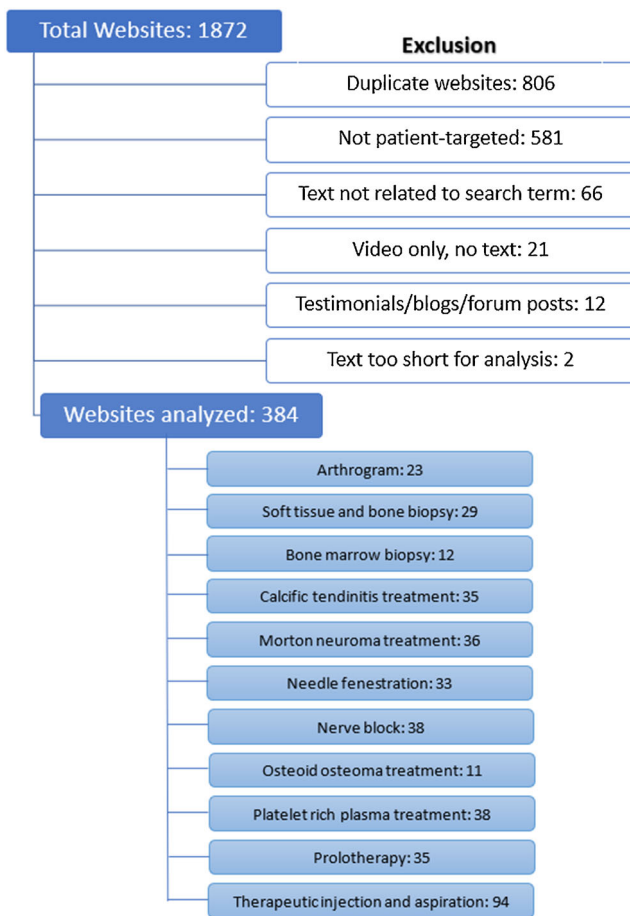
Fig. 1 Website selection criteria

program (Readability Studio 2015, Oleander Software). Readability scores were calculated using 6 validated metrics: the Flesch reading ease (FRE), Flesch-Kincaid Grade Level (FKGL), Gunning Fog index (GFI), Fry score, Raygor estimate, and simple measure of gobbledygook (SMOG) grading [13–18]. All of these metrics measure text readability, though each does so in a unique way.

The FRE is calculated with the number of syllables per 100 words (word length) and the average sentence length.

$$\text{FRE} = 206.835 - (846*\text{word length}) - (1.015*\text{sentence length})$$

The FRE scale falls between 0 (most difficult, lowest readability) and 100 (high readability, easy to read) [13]. The FKGL is a modification of the FRE formula which reports the grade level of the text rather than a 100-point scale [14].

$$\text{FKGL} = (0.39*\text{sentence length}) + (11.8*\text{word length})$$

The GFI includes both the average sentence length and the number of complex words (words with three or more syllables).

$$\text{GFI} = 0.4*(\text{average sentence length} + \text{percentage of hard words})$$

The resultant score shows a reading grade level similar to the FKGL [15]. The Fry score is based on a readability graph that takes into account the average number of sentences and average number of syllables per 100 words. Upon plotting these numbers on a readability graph, the grade level can be estimated [16]. The Raygor estimate is based on a graph derived from plotting the average number of sentences and average number of words with 6+ characters per 100 words [17]. Finally, the SMOG grading is calculated by selecting 30 sentences from the text, 10 each from beginning, middle, and end, then counting complex words of 3 or more syllables in these sentences [18].

$$\text{SMOG} = \sqrt{\text{complex words}} + 3.$$

As additional measures of complexity, automated analyses were performed for each webpage including the percentage of words with 3+ syllables, percentage of words with 6+ characters, percentage of sentences with 22+ words (difficult sentences), and average number of words per sentence (sentence length).

Statistics were performed using an independent $t$ test in Excel 2019 MSO (version 2003, Microsoft Corporation, Seattle, WA) with significance set to $p \leq 0.05$.
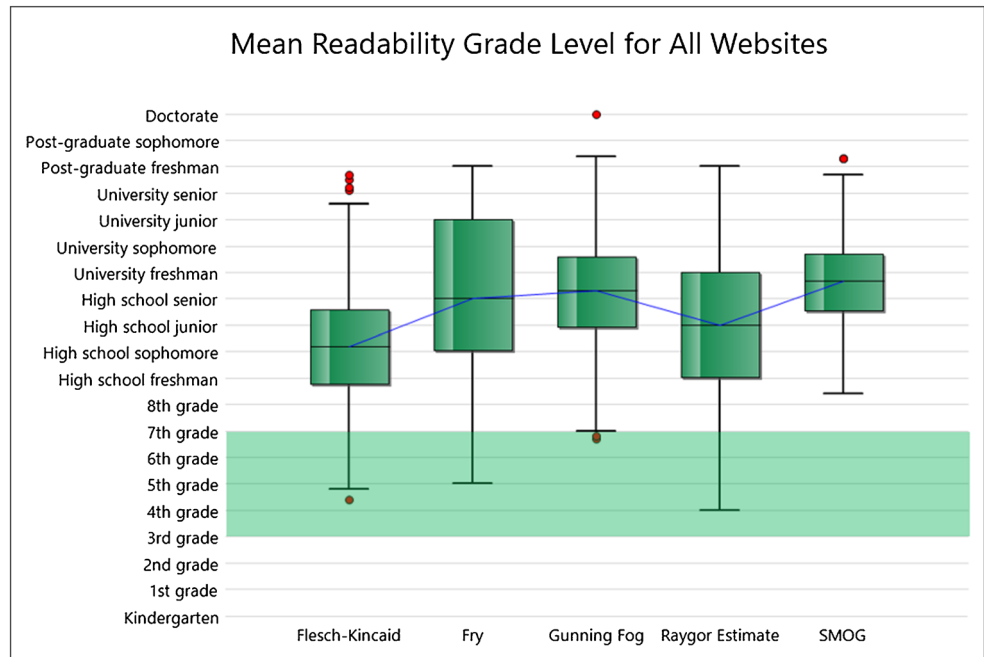
## Results

The reading grade level for all websites analyzed without regard for procedure type is displayed for 5 of the metrics (FKGL, Fry, GFI, Raygor estimate, and SMOG) in Fig. 2. The mean reading grade levels were all above the AMA and NIH recommendation for each of these metrics. Breakdowns of individual websites for the Fry and Raygor estimate metrics are shown in Figs. 3 and 4, respectively. These graphs plot each website as a single point with corresponding reading grade level. The majority of the websites had text written above the AMA and NIH recommendation on both of these metrics.

The 6th metric (FRE) is shown in Fig. 5. Unlike the other readability measures, the FRE produces scores ranging from 0 to 100 for each website analyzed rather than a reading grade level (lower values indicate more difficult text). A value of below 70 denotes a 7th grade or higher reading level. The majority of websites were therefore above the AMA and NIH recommendation on this metric as well.

Detailed information for all procedures with a single composite mean reading grade level for each is shown in Table 2. All of the procedures had websites with text written at or above the high school level (10th to 14th grade). Only 21 websites (5.5%) were written at the AMA- and NIH-recommended level (Supplement A). Prolotherapy was the procedure with the highest reading grade level (~ 14th grade).

Fig. 2 Readability scores for all websites. The green shaded area denotes AMA- and NIH-recommended grade level



Arthrogram was the procedure with the lowest reading grade level (~ 10th grade).

The mean reading grade level was compared between all the websites on the 1st page of search engine results to all the websites on the 2nd and 3rd pages with no difference identified across all procedures ($11.9 \pm 1.4$ vs. $11.6 \pm 1.2$, $p = 0.47$). For academic vs. non-academic websites, those on nerve block showed a significantly higher mean reading grade level on non-academic sites ($11.6 \pm 2.4$ vs. $10.0 \pm 1.2$, $p = 0.025$). There was otherwise no difference for the remainder of the procedures (Table 3).

When measuring reading complexity, prolotherapy had websites with text containing the highest percentage of words with 3+ syllables and 6+ characters (complex words) and sentences with 22+ words (complex sentences). Data for other procedures are shown in Table 4.

Fig. 3 Fry Readability Graph for all websites. Each website is plotted as a single point. The green shaded area denotes the AMA- and NIH-recommended grade level. The gray zones represent areas of two extremes: long words in the top right corner (words with more syllables) and long sentences in the bottom left corner (fewer number of sentences per 100 words). Plotted points in the gray zones are invalid for analysis (13)
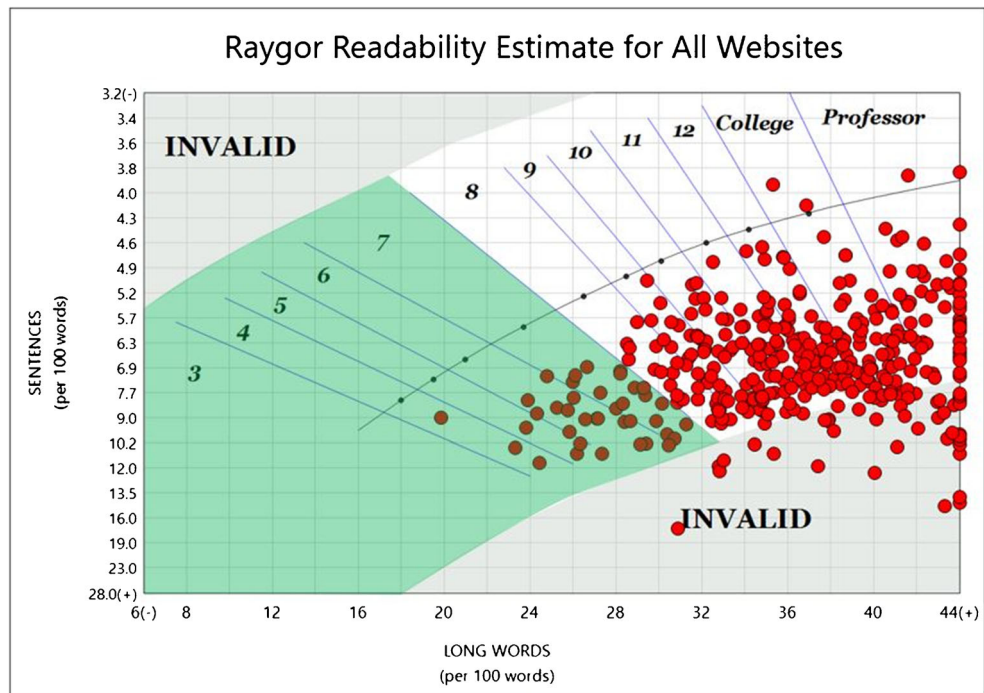
**Fig. 4** Raygor readability estimate for all websites. Each website is plotted as a single point. The green shaded area denotes the AMA- and NIH-recommended grade level. The gray zones represent areas of two extremes: long words in the bottom right corner (words with more syllables) and long sentences in the top left corner (fewer number of sentences per 100 words). Plotted points in the gray zones are invalid for analysis
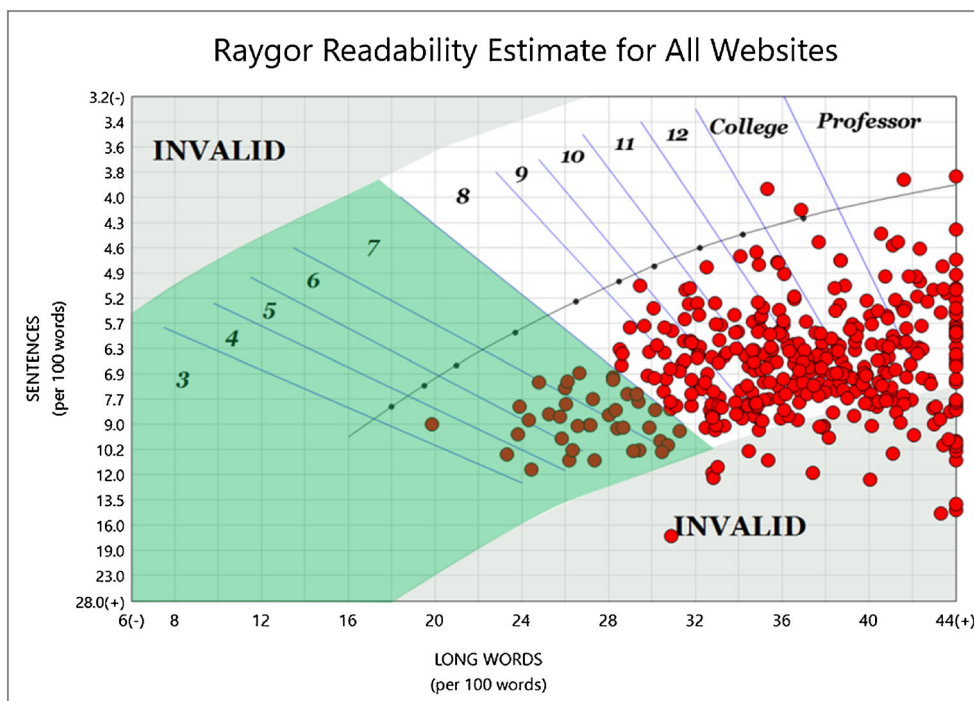


## Discussion

This investigation showed that patient-targeted online information on musculoskeletal radiology procedures is written at the 10th–14th grade level—much higher than the AMA- and NIH-recommended 3rd–7th grade level. There has been limited assessment of the readability of information on minimally invasive procedures, though this is not a problem exclusive to musculoskeletal radiology. Our results are comparable with one study which found that online procedure material in interventional radiology was written at a 10th–15th grade level [19]. They are also similar to another study which reported a 11th–12th grade reading level for information on breast lesions requiring biopsy or surgery [8]. Expanding the scope more widely within medicine, the same trends have also been found in other subspecialties including orthopedics, pediatrics, and ophthalmology [20–22].

A total of 36% of the US adult population has basic or below basic health literacy, which can compromise the ability of those people to obtain appropriate medical education [23]. This is further exacerbated when literature is written at too high of a reading level. Low health literacy has been associated with significant negative outcomes including increased mortality [2–4]. Given the minimally invasive nature of musculoskeletal procedures, more applicable negative outcomes to consider include an increase in post-procedural complications, requirement of longer recovery time, and increase in anxiety levels [24].

**Fig. 5** Flesch reading ease for all websites. The green shaded area denotes the AMA-and NIH-recommended grade level (values of 70 and below correlate to 7th grade and above reading level), i.e., the lower the index value, the more difficult the text
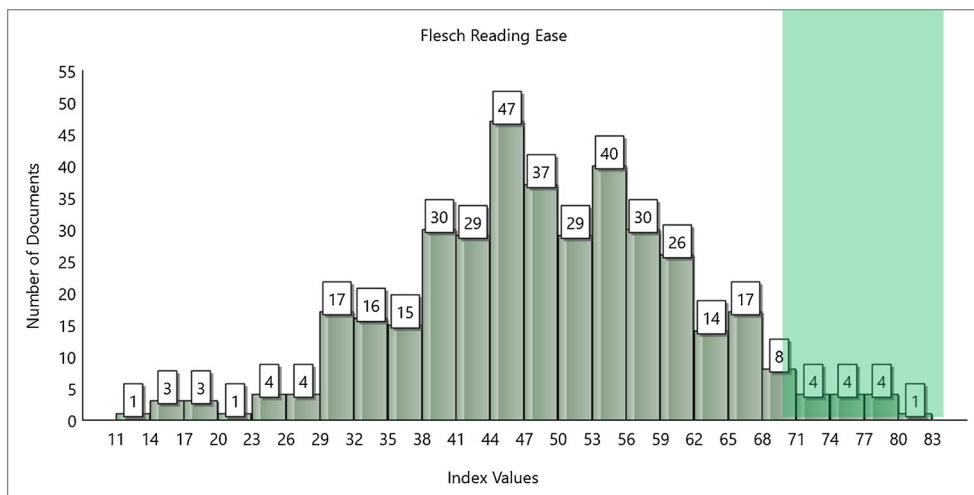
**Table 2** Mean readability for all metrics

| Procedure type | Flesch reading ease¥ | Flesch-Kincaid* | Fry* | Gunning fog* | Raygor estimate* | SMOG* | Composite mean grade level† |
|---|---|---|---|---|---|---|---|
| Arthrogram | 58 ± 12 (31–80) | 8.7 ± 2.1 (4.8–13.6) | 10.1 ± 3.2 (6–17) | 11.0 ± 2.3 (7.2–15.5) | 8.7 ± 2.5 (4.0–13.0) | 11.6 ± 1.7 (8.7–14.9) | 10.0 |
| Soft tissue and bone biopsy | 54 ± 11 (33–78) | 9.2 ± 2.1 (5.1–13.7) | 10.7 ± 3.3 (6–17) | 11.8 ± 2.1 (7.6–16.4) | 10.0 ± 3.3 (5–17) | 12.1 ± 1.6 (8.9–15.3) | 10.8 |
| Bone marrow biopsy | 56 ± 13 (35–76) | 8.9 ± 2.4 (5.6–12.7) | 10.6 ± 3.9 (6–17) | 11.5 ± 2.4 (7–15.2) | 9.4 ± 3.1 (6–13) | 11.7 ± 1.8 (8.7–14.6) | 10.4 |
| Calcific tendinitis treatment | 47 ± 9 (26–68) | 10.6 ± 1.9 5.3–15.5) | 13.7 ± 2.7 (9–17) | 13.3 ± 2.1 (8.2–17.7) | 12.3 ± 2.4 (9–17) | 13.1 ± 1.6 (8.6–16.9) | 12.6 |
| Morton neuroma treatment | 51 ± 13 (13–80) | 10 ± 2.3 (5.1–16.1) | 11.7 ± 3.2 (6–17) | 12.3 ± 2.4 (7.3–17.3) | 11.3 ± 2.8 (5–17) | 12.6 ± 1.8 (8.7–16.5) | 11.6 |
| Needle fenestration | 44 ± 12 (17–69) | 11.2 ± 2.3 (6.6–16.5) | 13.6 ± 2.8 (8–17) | 12.6 ± 2.4 (7.5–16.7) | 14.7 ± 2.9 (10–17) | 12.9 ± 2.0 (9–17.3) | 13.1 |
| Nerve block | 54 ± 10 (32–69) | 9.5 ± 2.1 (6.3–14.1) | 11.3 ± 2.9 (7–17) | 12.2 ± 2.0 (9.3–16.4) | 10.1 ± 2.8 (5–17) | 12.4 ± 1.6 (9.9–15.5) | 11.0 |
| Osteoid osteoma treatment | 43 ± 5 (32–50) | 11.4 ± 0.9 (9.4–12.9) | 14.7 ± 1.4 (13–17) | 12.9 ± 1.2 (11.1–15.1) | 12.2 ± 2.2 (9–17) | 13.2 ± 0.7 (11.6–14.4) | 12.9 |
| Platelet rich plasma treatment | 48 ± 6 (39–63) | 10.7 ± 1.2 (8.1–13.7) | 13.0 ± 2.0 (9–17) | 12.4 ± 1.5 (9–15.4) | 12.4 ± 1.5 (9–17) | 13 ± 1.0 (10.9–15.4) | 12.2 |
| Prolotherapy | 33 ± 8 (15–50) | 12.9 ± 1.8 (9.6–16.7) | 16.7 ± 0.9 (14–17) | 13.9 ± 1.6 (10.8–19) | 15.4 ± 2.3 (10–17) | 14.2 ± 1.5 (11.5–17.3) | 14.6 |
| Therapeutic injection and aspiration | 53 ± 11 (26–81) | 9.6 ± 2.0 (4.4–14.5) | 11.3 ± 3.1 (5–17) | 11.7 ± 2.0 (6.7–17.1) | 10.4 ± 2.6 (4–17) | 12.1 ± 1.6 (8.4–15.8) | 10.9 |

¥ Values on Flesch reading ease scoring scale from 0 to 100 (lower scores correspond to higher reading grade levels)

*Values in mean reading grade level ± standard deviation (range)

† Includes the grading scales that express results in grade level (i.e., all except for Flesch reading ease)

The reading level for nearly all procedures did not depend on whether the websites originated from an academic source. This is consistent with a prior study which found that readability for online materials pertaining to pediatric patients was similar across authors and institution groups, both academic and non-academic [22]. There are a couple of potential explanations for this. First, non-academic materials may fail to adapt language to the patient level because they reference academic sources. Second, there may be a general lack of awareness among content producers about the need to write material at a lower reading level. Apart from this, similar high reading levels were also found pervasively on the first three pages of search results. Taken together, these findings highlight the difficulty that patients may have in finding suitable education material, even when motivated to perform a comprehensive search.

Our results indicate that simple interventions can lower the readability of patient information to be closer to the AMA and NIH recommendation. The webpages we analyzed had a high proportion of complex words (3+ syllables or 6+ characters) and difficult sentences (22+ words). Readability would be enhanced if one or two-syllable words, shorter words, and shorter sentences were used more frequently. Similar interventions and other strategies for improving communication with patients have been advocated by the AMA and Center for Disease Control and Prevention (CDC). They suggest use of the active voice, limiting contents to only what is necessary, and avoidance of complex tables and graphs [5, 25]. Other techniques to improve comprehension also include the use of more colors and illustrations [26].

The following is an example of how these interventions can be helpful. The text below is a sentence excerpt taken from one of the websites in the study, which has a mean reading grade level of 14.7 based on the 6 metrics we used:

"Bone biopsies may be used to confirm the diagnosis of a bone disorder, investigate an abnormality, determine the cause of pain or infection, or distinguish bone tumor from other conditions." [27].

The same information can be rewritten with a reduction in the number of syllables and characters in words and decreased sentence length, resulting in a mean reading grade level of 6.7:

"A bone biopsy is done to look at bone diseases. It gives your doctor information to find out if there is something abnormal. It can help to find the cause of pain or test for cancer and other problems."

As another illustrative example, the following sentence excerpt from the description of an arthrogram has a mean reading grade level of 14.2:

"The procedure is often used to help diagnose persistent, unexplained joint pain or discomfort. In some cases, local anesthetic medications or steroids may be injected into the joint along with the contrast material." [28].

**Table 3** Comparison of mean readability grade level of academic vs. non-academic websites

| Procedure type | Academic[†] | Non-academic[†] | p value |
|---|---|---|---|
| Arthrogram | 10.7 ± 1.6 (8.7–13) | 9.7 ± 2.5 (6.2–14.8) | 0.32 |
| Soft tissue and bone biopsy | 11.0 ± 2.3 (6.7–15.9) | 10.6 ± 2.2 (6.7–13.8) | 0.58 |
| Bone marrow biopsy | 10.9 ± 2.4 (8–13.2) | 10.2 ± 2.7 (6.7–14.4) | 0.67 |
| Calcific tendinitis treatment | 12.7 ± 4.1 (9.8–15.6) | 12.3 ± 1.9 (7.4–16.8) | 0.81 |
| Morton neuroma treatment | 10.5 ± 1.6 (8.1–12.4) | 11.8 ± 2.5 (6.4–16.8) | 0.22 |
| Needle fenestration | 13.0 ± 3.0 (8.9–16.5) | 12.5 ± 2.3 (7.7–16.8) | 0.63 |
| Nerve block | 10.0 ± 1.2 (7.6–11.8) | 11.6 ± 2.4 (8–15.7) | 0.025* |
| Osteoid osteoma treatment | 12.5 ± 1.5 (10.3–14.9) | 12.8 ± 0.9 (12–14.2) | 0.66 |
| Platelet rich plasma treatment | 12.0 ± 1.4 (10.6–14.5) | 12.2 ± 1.4 (9.2–14.8) | 0.73 |
| Prolotherapy | 14.5 ± 2.2 (12–16.1) | 14.1 ± 1.8 (10.8–17.2) | 0.73 |
| Therapeutic injection and aspiration | 10.5 ± 1.3 (7.9–12.7) | 11.0 ± 2.2 (5.7–16.2) | 0.33 |

[†] Values in mean reading grade level ± standard deviation (range)

*Statistical significance at $p \leq 0.05$

Using the same principles, this can be rewritten to a mean reading grade level of 6.5:

"This test helps to find out what is causing the joint pain that bothers you. Sometimes, numbing or steroid medications are put into the joint with the dye fluid."

Implementation of some of these strategies in the revision of reading materials has resulted in significant improvement of readability levels [21]. More importantly, patients have demonstrated enhanced understanding of these revised materials [21, 26, 29]. Although similar initiatives have taken place in radiology, Bange et al. have recently shown that some reading materials still persist at too high of a level, even after a 5-year evolutionary period [7, 30].

When deciding to revise patient education materials, it is important to consider that not every patient may benefit from lowering the reading grade levels. Lower levels may not lead to increased comprehension and satisfaction for everyone [21, 31, 32]. Simplification of written texts may eliminate language details and nuances desired by some patients. It is conceivable that those with higher levels of education or health literacy may want more complex reading material. To account for this, some authors have advocated having two sets of documents available—one for easier comprehension and another that contains more extensive details [33].

Our study has clinical relevance and notable strengths. First, we evaluated literature from the Internet, which is where a large percentage of the population finds information [1, 34]. Second, although more than 90% of Internet users do not look past the very first page of results on any given search engine [34, 35], we were comprehensive with our methods by analyzing websites from the first three pages of results on separate search engines.

**Table 4** Word and sentence complexity for all websites

| Procedure type | % of words with 3+ syllables[†] | % of words with 6+ characters[†] | % of sentences with 22+ words (difficult sentences)[†] | Average number of words/sentence (sentence length)[†] |
|---|---|---|---|---|
| Arthrogram | 16.0 ± 4.2 | 31.6 ± 5.3 | 13.3 ± 7.7 | 13.8 ± 2.7 |
| Soft tissue and bone biopsy | 18.3 ± 4.2 | 35.1 ± 5.4 | 13.4 ± 12.8 | 13.6 ± 3.0 |
| Bone marrow biopsy | 16.3 ± 4.1 | 33.8 ± 5.4 | 13.8 ± 13.1 | 13.9 ± 3.2 |
| Calcific tendinitis treatment | 21.6 ± 3.0 | 37.5 ± 3.7 | 13.4 ± 10.6 | 14.8 ± 3.5 |
| Morton neuroma treatment | 18.4 ± 4.8 | 36.8 ± 5.5 | 18.3 ± 10.3 | 14.9 ± 2.7 |
| Needle fenestration | 18.7 ± 5.1 | 41.5 ± 5.2 | 22.7 ± 15.1 | 16.0 ± 4.5 |
| Nerve block | 17.7 ± 3.6 | 33.7 ± 4.7 | 22.6 ± 26.6 | 14.9 ± 3.2 |
| Osteoid osteoma treatment | 19.7 ± 2.2 | 37.9 ± 3.4 | 21.4 ± 12.4 | 15.5 ± 2.6 |
| Platelet rich plasma treatment | 19.2 ± 3.0 | 39.0 ± 4.5 | 21.6 ± 12.5 | 15.7 ± 3.2 |
| Prolotherapy | 23.7 ± 3.5 | 43.2 ± 3.9 | 24.0 ± 16.0 | 15.9 ± 4.1 |
| Therapeutic injection and aspiration | 16.9 ± 4.2 | 34.5 ± 4.8 | 16.9 ± 10.9 | 14.6 ± 3.0 |

[†] Values in mean ± standard deviation

We also acknowledge that there are some limitations. First, the webpages we evaluated were not representative of all reading material for each procedure and not all types of procedures were included. We did use several search terms to capture as many unique websites as possible, though variations in the keywords would have produced different results. Second, although there is evidence that visual aids are useful for improving recall, we were unable to assess most of the videos and pictures and their role in improving patient comprehension [36, 37]. Third, individual reading level metrics that were used have their own limitations. For example, there is variability of the emphasis on word and sentence length, as well as the assumption that increased word or sentence length is linearly correlated to reading difficulty [13–18, 38]. We did manage to mitigate some of this through the use of 6 separate validated metrics, each of which demonstrated similar results. Finally, and most importantly, we did not directly assess patient understanding, which is a critical outcome that deserves further study.

In conclusion, patient-targeted online websites on musculoskeletal radiological procedures are written at 10th–14th grade reading level, which is well beyond the AMA and NIH recommendation (3rd–7th grade). Readability can be lowered by decreasing text complexity through limitation of high-syllable words and reduction in word and sentence length.

## Compliance with ethical standards

**Conflict of interest**    The authors declare that they have no conflict of interest

**Ethical approval**    N/A. IRB exempt.

**Informed consent**    N/A. IRB exempt.

## References

1. America's Health Literacy: Why We Need Accessible Health Information. U.S. Department of Health and Human Services (HHS) 2008.
2. Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP. Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med. 2004;19(12):1228–39.
3. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med. 2011;155(2):97–107.
4. Sudore RL, Yaffe K, Satterfield S, Harris TB, Mehta KM, Simonsick EM, et al. Limited literacy and mortality in the elderly: the health, aging, and body composition study. J Gen Intern Med. 2006;21(8):806–12.
5. Weiss BD. Health literacy: a manual for clinicians. In: Foundation AMAaAM, ed. Chicago 2003.
6. How to write easy-to-read health materials. National Institutes of Health.
7. Hansberry DR, John A, John E, Agarwal N, Gonzales SF, Baker SR. A critical review of the readability of online patient education resources from RadiologyInfo.Org. AJR Am J Roentgenol. 2014;202(3):566–75.
8. Miles RC, Baird GL, Choi P, Falomo E, Dibble EH, Garg M. Readability of online patient educational materials related to breast lesions requiring surgery. Radiology. 2019;291(1):112–8.
9. Prabhu AV, Donovan AL, Crihalmeanu T, Hansberry DR, Agarwal N, Beriwal S, et al. Radiology online patient education materials provided by major university hospitals: do they conform to NIH and AMA guidelines? Curr Probl Diagn Radiol. 2018;47(2):75–9.
10. Trofimova A, Vey BL, Safdar NM, Duszak R Jr, Kadom N. Radiology report readability: an opportunity to improve patient communication. J Am Coll Radiol. 2018;15(8):1182–4.
11. Hoermann S, Doering S, Richter R, Walter MH, Schussler G. Patients' need for information before surgery. Psychother Psychosom Med Psychol. 2001;51(2):56–61.
12. Top 5 Search Engines in the United States on March 2020. StatCounter Global Stats: StatCounter 2020.
13. Flesch R. A new readability yardstick. J Appl Psychol. 1948;32:221–33.
14. Kincaid JP, R Flesch, Rogers RL, Chissom BS. Derivation of new readability formulas) automated readability index, fog count and Flesch reading ease formula) for navy enlisted personnel: Institute for Simulation and Training, 1975.
15. Gunning R. The Fog Index after twenty years. J Bus Commun. 1969;6:3–13.
16. Fry E. A readability formula that saves time. J Read. 1968;11(7):513–6 575-578.
17. Raygor AL. The Raygor readability estimate: a quick and easy way to determine difficulty. In: Pearson PD, ed. Reading: Theory Rap: National Reading Conference, Clemson; 1977:259–263.
18. Harry GLM. SMOG grading-a new readability formula. J Read. 1969;12:639–46.
19. McEnteggart GE, Naeem M, Skierkowski D, Baird GL, Ahn SH, Soares G. Readability of online patient education materials related to IR. J Vasc Interv Radiol. 2015;26(8):1164–8.
20. Roberts H, Zhang D, Dyer GS. The readability of AAOS patient education materials: evaluating the progress since 2008. J Bone Joint Surg Am. 2016;98(17):e70.
21. Williams AM, Muir KW, Rosdahl JA. Readability of patient education materials in ophthalmology: a single-institution study and systematic review. BMC Ophthalmol. 2016;16:133.
22. D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. Arch Pediatr Adolesc Med. 2001;155(7):807–12.
23. Kutner M. The Health Literacy of America's Adults: results from the 2003 National Assessment of Adult Literacy. Washington: U.S. Department of Education, National Center for Education Statistics; 2006.
24. Chang ME, Baker SJ, Dos Santos Marques IC, Liwo AN, Chung SK, Richman JS, et al. Health literacy in surgery. Health Lit Res Pract. 2020;4(1):e46–65.
25. Simply Put: A guide for creating easy-to-understand materials. In: Prevention CfDCa, ed. 3 ed 2010.
26. Davis TC, Bocchini JA Jr, Fredrickson D, Arnold C, Mayeaux EJ, Murphy PW, et al. Parent comprehension of polio vaccine information pamphlets. Pediatrics. 1996;97(6 Pt 1):804–10.
27. "Bone Biopsy ". RadiologyInfo.org for Patients (January 17, 2020). www.radiologyinfo.org/en/info.cfm?pg=bonebiopsy. Accessed on June 1, 2020.
28. "Direct Arthrography". RadiologyInfo.org for Patients (July 17, 2019). www.radiologyinfo.org/en/info.cfm?pg=arthrog. Accessed on July 12, 2020.
29. Eaton ML, Holloway RL. Patient comprehension of written drug information. Am J Hosp Pharm. 1980;37(2):240–3.

30. Bange M, Huh E, Novin SA, Hui FK, Yi PH. Readability of patient education materials from RadiologyInfo.org: has there been Progress over the past 5 years? AJR Am J Roentgenol. 2019;213(4):875–9.

31. Davis TC, Fredrickson DD, Arnold C, Murphy PW, Herbst M, Bocchini JA. A polio immunization pamphlet with increased appeal and simplified language does not improve comprehension to an acceptable level. Patient Educ Couns. 1998;33(1):25–37.

32. Davis TC, Holcombe RF, Berkel HJ, Pramanik S, Divers SG. Informed consent for clinical trials: a comparative study of standard versus simplified forms. J Natl Cancer Inst. 1998;90(9):668–74.

33. Mehta MP, Swindell HW, Westermann RW, Rosneck JT, Lynch TS. Assessing the readability of online information about hip arthroscopy. Arthroscopy. 2018;34(7):2142–9.

34. Eysenbach G, Kohler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ. 2002;324(7337):573–7.

35. van Deursen AJ. Internet skill-related problems in accessing online health information. Int J Med Inform. 2012;81(1):61–72.

36. Houts PS, Witmer JT, Egeth HE, Loscalzo MJ, Zabora JR. Using pictographs to enhance recall of spoken medical instructions II. Patient Educ Couns. 2001;43(3):231–42.

37. Schenker Y, Fernandez A, Sudore R, Schillinger D. Interventions to improve patient comprehension in informed consent for medical and surgical procedures: a systematic review. Med Decis Mak. 2011;31(1):151–73.

38. Friedman DB, Hoffman-Goetz L. A systematic review of readability and comprehension instruments used for print and web-based cancer information. Health Educ Behav. 2006;33(3):352–73.