Online Resources for Shoulder Instability: What Are Patients Reading?

Grant H. Garcia, MD, Samuel A. Taylor, MD, Christopher J. Dy, MD, MPH, Alexander Christ, MD, Ronak M. Patel, BS, and Joshua S. Dines, MD

**Background:** Evaluations of the medical literature suggest that many online sites provide poor-quality information. The purpose of our study was to investigate the value of online resources for patient education about shoulder instability.

**Methods:** Three search terms (‘‘shoulder instability,’’ ‘‘loose shoulder,’’ and ‘‘shoulder dislocation’’) were entered into three Internet search engines. Three orthopaedic residents independently gauged the quality and accuracy of the information with use of a set of predetermined scoring criteria, in addition to noting whether or not four potential surgery options were mentioned. The readability of the web sites was evaluated with use of the Flesch-Kincaid score.

**Results:** Eighty-two unique web sites were evaluated. Quality and accuracy were significantly higher with use of the term ‘‘shoulder instability’’ compared with the term ‘‘loose shoulder’’ (quality, \( p < 0.001 \); accuracy, \( p = 0.001 \)). However, the reading level was significantly more advanced for the ‘‘shoulder instability’’ web sites (\( p < 0.001 \)). Quality was significantly higher on web sites with reading levels above the eighth grade level (\( p = 0.001 \)) (88% of web sites). Only twenty-three sites (28%) mentioned surgical options for shoulder instability, and of these, only eight mentioned thermal capsulorrhaphy as a primary treatment.

**Conclusions:** Online information regarding shoulder instability is often inaccurate and/or at an inappropriately high reading level. The quality of information is highly dependent on the specific search term used. Clinicians need to be aware of the information that is available online and should help direct patients to proper sites and guide Internet search terms.

Over 220 million Americans use the Internet daily, and this number is growing rapidly. Recent studies demonstrate that over 61% of Americans search for medically related information on the Internet. In addition to the obvious educational advantage for patients seeking to learn about their condition, it is important to consider the risks of misinformation. In

**Disclosure:** None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.
particular, there is a paucity of online peer-reviewed material readily available to patients. Additionally, industry-driven information may be interpreted as legitimate by the average reader without recognition of inherent biases. Because there is no standard of quality or standard method of review, patients may be misinformed about the nature of their medical condition and the treatment options that are presented online. Recent evaluation of online health sites demonstrated poor-quality information.

Reports over the past ten years have suggested that many online health sites are more likely to lack important diagnostic criteria and information compared with peer-reviewed medical references. Previous investigations have evaluated the quality of web sites by entering search terms for topics such as back pain, carpal tunnel syndrome, scoliosis, and lumbar disc herniation. They concluded that over 50% of online materials were of poor quality and that patients are best served by being directed to accurate online resources by their physicians. More recently, studies by Dy et al. and Fabricant et al. evaluated online content relating to fractures of the distal part of the radius, lateral humeral epicondylitis, and developmental hip dysplasia. Because there is no standard of quality or standard method of review, patients may be misinformed about the nature of their medical condition and the treatment options that are presented online. Recent evaluation of online health sites demonstrated poor-quality information.

Materials and Methods

During this web site investigation, we selected three search terms to represent those chosen by patients with shoulder instability: "shoulder instability," "loose shoulder," and "shoulder dislocation." Each of the three search terms was entered as written, without the addition of any other words for specification, into three search engines (Google, Yahoo, and Bing) on March 16, 2013. These search engines were chosen because they represent over 90% of searches performed on the Internet. We evaluated the first twenty-five results from each search. After duplicates were eliminated, 100 web sites remained (Fig. 1). Eighteen web sites were excluded because they lacked informational content (including those with only video content [three sites]), were unrelated to shoulder instability (eight sites), or contained materials intended for peer review (seven sites). Ultimately, eighty-two web sites met the inclusion criteria and underwent blinded assessment (i.e., reviewers were blinded to the type of search term used to locate the web site).

The accuracy and quality of the web sites were assessed by methods similar to those used by previous investigators for scoliosis, back pain, and fractures of the distal part of the radius. As such, twenty-five items that included elements of diagnosis, anatomy, treatment plan, and complications specific to shoulder instability, based upon guidelines written by the American Academy of Orthopaedic Surgeons (AAOS) with collaboration from the American Shoulder and Elbow Surgeons, were included in the quality grading sheet (Table I).

Web site quality was scored by awarding one point per criteria contained in the web site, with a maximum score of 25 points. Four supplemental surgical options were scored separately (Table II), and were used to assess the frequency with which surgical options were presented to a reader. Three orthopaedic surgery residents were trained by the senior author (J.S.D.) in the evaluation of quality and accuracy of web site information, and they subsequently evaluated each captured web site. Quality scores for the three reviewers were combined and averaged to generate a mean quality score (maximum of 25 points) for each captured web site (Table III).

To determine web site accuracy, the three reviewers independently rated the information on a scale of 1 to 4, following a practice used in prior...
investigations. The reviewers assigned an accuracy score of 1 if they agreed with <25% of the web site content, a score of 2 for agreement with 26% to 50% of the web site content, a score of 3 for agreement with 51% to 75% of the web site content, and a score of 4 for agreement with >75% of the web site content. The scores from the three reviewers were added to create a composite accuracy score for each web site, and a maximum total score of 25 points. There was a high intraclass correlation coefficient of 0.86. The Flesch-Kincaid (FK) method was used to assess readability. FK grade level indicates the maximum level of schooling a patient must have obtained to be able to read and comprehend the material. Therefore, a higher FK level indicates information that would be more difficult to understand. FK level (readability threshold) and those at or above the eighth grade level, based upon recommendations for readable patient education material and the average reading level (eighth grade) of the U.S. population. Online resources were also assessed for commercial gain; it was assumed that the web site received profits from the advertisement and that the company in the advertisement received profits from the product. Finally, we categorized the web sites by authorship: health-care provider (nurse, physician, or physical therapist with state credentials), non-health-care provider (blogs, boards, or personal web sites without explicitly stated credentials), and, lastly, commercial. Before each electronic capture was performed, our data researcher (R.M.P.), who was not involved in the scoring, identified the authorship of each site.

### Statistical Methodology

Comparative statistics were used to analyze the accuracy assessment, the FK level, and the quality score. Analysis of variance (for normally distributed data) and Kruskal-Wallis tests (for non-normally distributed data) with post hoc pairwise comparisons were used to determine any difference in quality, accuracy, and readability based on the search term used, the order of search results, and the web site author. Independent sample t tests (for normally distributed data) and Mann-Whitney U tests (for non-normally distributed data) were used to determine any difference in quality, accuracy, or readability based on whether a web site was seeking commercial gain. Correlation analysis (Pearson correlation coefficient for normally distributed data and Spearman rank correlation coefficient for non-normally distributed data) was also used to determine associations between web site quality and FK scores, as well as web site accuracy and FK scores. Interrater reliability for determining quality and accuracy ratings was evaluated with use of the intraclass correlation coefficient. The threshold for significance was \( p < 0.05 \) in all statistical tests. The statistical analysis was conducted with SPSS 16.0 (IBM; Armonk, New York) by a coauthor (C.J.D.) with graduate-level training in biostatistics.

### Results

#### Web Site Content

Of the eighty-two unique web sites that were evaluated, there were thirty specific sites for “shoulder dislocation,” twenty-nine for “shoulder instability,” and twenty-three for “loose shoulder.” Thirteen sites were authored by physicians affiliated with an academic institution, thirty-two sites were authored by physicians without an affiliation to an academic institution, eleven sites had commercial affiliations, twelve sites were authored by non-physician health-care providers, and fourteen were personal blogs. With 91% of the web sites (seventy-five web sites), the target audience was the patient.

As shown in Table IV, the average quality for all web sites was 9.48 ± 5.11 of a maximum of 25 points. There was a significant difference among the overall quality scores of the three search terms (\( p < 0.001 \)). For “shoulder instability,” the mean quality score was 11.94 ± 4.5 in comparison to “loose shoulder,” which was 6.10 ± 5.41; post hoc pairwise analysis demonstrated a

### TABLE II Supplemental Surgical Criteria

<table>
<thead>
<tr>
<th>Types of Surgery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentions thermal capsulorrhaphy</td>
</tr>
<tr>
<td>Mentions arthroscopic Bankart</td>
</tr>
<tr>
<td>Mentions Latarjet or bone block procedure</td>
</tr>
<tr>
<td>Mentions open Bankart</td>
</tr>
</tbody>
</table>

The reviewers assigned an accuracy score of 1 if they agreed with <25% of the web site content, a score of 2 for agreement with 26% to 50% of the web site content, a score of 3 for agreement with 51% to 75% of the web site content, and a score of 4 for agreement with >75% of the web site content. The scores from the three reviewers were added to create a composite accuracy score for each web site, and a maximum total score of 25 points. There was a high intraclass correlation coefficient of 0.86. The Flesch-Kincaid (FK) method was used to assess readability. FK grade level indicates the maximum level of schooling a patient must have obtained to be able to read and comprehend the material. Therefore, a higher FK level indicates information that would be more difficult to understand. FK level (readability threshold) and those at or above the eighth grade level, based upon recommendations for readable patient education material and the average reading level (eighth grade) of the U.S. population. Online resources were also assessed for commercial gain; it was assumed that the web site received profits from the advertisement and that the company in the advertisement received profits from the product. Finally, we categorized the web sites by authorship: health-care provider (nurse, physician, or physical therapist with state credentials), non-health-care provider (blogs, boards, or personal web sites without explicitly stated credentials), and, lastly, commercial. Before each electronic capture was performed, our data researcher (R.M.P.), who was not involved in the scoring, identified the authorship of each site.
scores of 9.70 ± 0.065) (Table III). As shown in Table IV, the average quality score for sites with a reading level above eighth grade standards was 10.16 ± 4.93, compared with a score of 4.63 ± 3.63 if the sites were written below the eighth grade reading level (p = 0.001). Web sites without commercial bias (seventy-two sites) had an average quality score of 9.94 ± 5.13, which was significantly higher than those with commercial bias, which had a score of 6.16 ± 3.56 (p = 0.027). Web sites authored by physicians (forty-five sites) had significantly higher quality ratings (p < 0.001) compared with all other authors. Additionally, the average reading level of web sites authored by physicians was an FK score of 11.1. The interobserver reliability of the quality rating was high, with an intraclass correlation coefficient of 0.91.

As shown in Table V, the average accuracy for all web sites was 8.61 ± 2.6 out of a maximum score of 12. The accuracy scores of “loose shoulder” site scores (6.96 ± 2.9) were significantly lower than the accuracy scores of the “shoulder dislocation” (9.20 ± 2.4) and “shoulder instability” (9.31 ± 1.9) sites (p = 0.001). There was a significant difference in the accuracy based on reading level of the web site; above an eighth grade level, the average score was 8.99 ± 2.30, compared with 5.90 ± 3.11 if the reading level was below an eighth grade level (p = 0.001). Additionally, if a physician had authored the site, there was a significantly higher accuracy rating compared with all of the other non-physician authors (p = 0.001). The accuracy rating was highly reliable among the raters, with an intraclass correlation coefficient of 0.86.

There were seventy-two sites with FK scores above 8. The overall FK scores were 10.94 ± 4.5 (range, 5 to 16). The average FK score for “shoulder instability” was 12.32 ± 2.5, and these sites were written at a significantly higher reading level than “shoulder dislocation” sites with FK scores of 10.6 ± 2.27 (p = 0.005) or “loose shoulder” sites with scores of 9.717 ± 1.95 (p < 0.001). Finally, web sites authored by a physician (forty-five web sites) had a significantly higher reading level compared with those authored by non-physicians (thirty-seven web sites) (p < 0.001).

**Surgical Management**

Out of eighty-two web sites, fifty-nine mentioned conservative treatment only (immobilization, physical therapy), with no reference to specific surgical options. Only twenty-three sites (28%) mentioned specific types of surgery for shoulder instability. Of those sites, fourteen described one surgical intervention, while nine described two or more surgical options. The most commonly described surgery, open or arthroscopic Bankart repair, was noted in eighteen sites. Thermal capsulorrhaphy was described in ten (12%) web sites (as an alternative surgical option in eight web sites, and for primary management in two web sites); no web sites mentioned the harmful effects of this option. Finally, the Latarjet procedure or bone block was described in seven web sites, always as a second surgical option after Bankart repair.
Discussion

Patients are increasingly turning to the Internet as a source of health-care information\textsuperscript{2,26}. There is concern that online resources used by patients may be inaccurate\textsuperscript{3,4,5,27}. The fact that shoulder instability predominantly affects younger patients who have heightened technological prowess amplifies these concerns\textsuperscript{15,16}. We demonstrated that online shoulder instability resources available to patients are not reliably accurate, and that the quality of the information is highly dependent on the specific search term. Inaccurate information found during a patient’s Internet investigation may be a source of frustration and may interfere with the physician-patient relationship.

The evaluation of the quality and the accuracy of web sites reveals biases that favor patients with higher levels of education. Although the search term “shoulder instability” yielded the sites with the highest overall quality and accuracy of sources, the readability of these sites was above a twelfth grade reading level; these sites were also more likely to have been written by a physician. Furthermore, only 12% (ten of eighty-two) of all web sites were written below the recommended eighth grade reading level\textsuperscript{24,25}. This is particularly worrisome because more accurate information may not be understandable by readers with a lower reading ability.

The findings related to the presentation of surgical options were interesting. Only 28% of sites (twenty-three of eighty-two) mentioned a surgical procedure despite the fact that 50% of patients younger than forty years old with shoulder instability have a surgical repair\textsuperscript{28}. The paucity of discussion about surgical intervention for shoulder instability may result in a delayed presentation of young, healthy, active first-time patients with dislocations to an orthopaedic surgeon. Such a delay has the potential to adversely impact a patient’s level of activity or participation in sports\textsuperscript{28}. Additionally, our findings regarding thermal capsulorrhaphy were somewhat alarming. Despite literature emphasizing the high failure rates and postoperative complications resulting from thermal capsulorrhaphy\textsuperscript{29-31}, this procedure was noted by ten web sites to be a viable, primary surgical option.

To the best of our knowledge, reports about the quality and accuracy of information on the Internet regarding shoulder instability previously have not been published in the literature, so we have no reference for comparison. Yet, data on other orthopaedic and medical terminology do exist. A number of publications demonstrate that online medical information is often inaccurate\textsuperscript{3,4,10,11,18,32}. Two studies reported similar results: “less medically advanced” search terms result in lower quality results. Dy et al. investigated online sources for fractures of the distal part of the radius and found that the search term “distal radius fracture” yielded higher quality and more accurate results compared with the term “broken wrist”\textsuperscript{12}. In contrast to fractures of the distal part of the radius, shoulder instability commonly occurs in young active patients\textsuperscript{33}. Typically, these younger patients are more facile with the use of computers and mobile devices for Internet searches than older patients\textsuperscript{34}. We believe that this demographic profile is an important consideration. Similar to Dy et al.\textsuperscript{12}, we found that individuals with lower reading levels may be at a disadvantage due to the fact that more accurate web sites tend to be written at a more advanced reading level. Fabricant et al. found that that the most accurate medical terminology (e.g., “developmental dysplasia of the hip”) yielded more accurate and higher quality information compared with results found with the general search term “congenital hip dislocation”\textsuperscript{13}. These studies concluded that
physicians should strive to direct patients to accurate and understandable Internet sites, and they emphasized the need for additional efforts by orthopaedic societies to optimize the search hierarchy of current sites and develop better online patient resources.

A major hurdle for patients doing medical research on the Internet is the sheer volume of information available. Without strict criteria or a governing body, much of the information goes uncensored and could potentially mislead patients. Recently, in a poll of 1200 patients, only 4% had received web site recommendations from their physicians, yet 70% of them stated that they would like such recommendations. The Federal Trade Commission has warned that much of the information available on the Internet is misleading and can be harmful to patients. There is an opportunity for physicians to direct patients to medically appropriate "preapproved" web sites. Sethuram and Weerakkody also emphasized the Internet's capabilities of educating patients with appropriate physician guidance. As a result, we developed a list of the top five web sites used in our study with the highest overall accuracy and quality, and we provided a quick link in a handout for physicians to distribute (Fig. 2). We realize that the Internet is dynamic, but we hope this will serve as guidance for physicians to distribute a list of "prescreened sites" to patients.

One of the major limitations of our study is the fluidity of the Internet, and because of open-source posting, the order of search results and rank may change on a daily basis. As a result, our data may be only a snapshot of a specific time point for shoulder instability web sites. An additional limitation is that our study only involved English-language searches, despite the Internet being multilingual and an international resource. Search engines often return results in multiple languages, but non-English results were not included in our analysis. As such, it is possible that other surgical procedures (e.g., Latarjet) may be more commonly encountered by patients who reside in regions where these procedures are more frequently performed. Other orthopaedic search terms may have yielded different search engine results, so no conclusions about the accuracy and quality of sites outside of the shoulder instability realm should be drawn. Furthermore, it is possible that web sites ranked outside of the top twenty-five results for each search engine may have contained important and accurate information. However, Internet users are more likely to access the links listed on the first two pages of results, making it less likely that lower ranked web sites will be accessed. Our final limitation was that when a web site was found with multiple search terms, we only analyzed it under the search term in which it was ranked the highest; thus, we may have eliminated higher quality web sites from the results of the use of the other search terms. Although this may have been the case, the rank of the search is important. For example, if higher quality sites are ranked lower in one search when compared with another search, this would require the searcher to "scroll" further. This helps to emphasize the fact that the specific search term influences the information that is obtained.

The overall quality and accuracy of shoulder instability information available on the Internet varies substantially with the search term that is used. We suggest the use of the term "shoulder instability," which we have demonstrated yields the most accurate and highest quality results. Unfortunately, these results may not be fully understandable to the average U.S. citizen (with an eighth grade reading level). This trade-off between quality and readability must be improved before physicians can conclude that the Internet is a good source of health-care-related information. One suggestion is for physicians to be the authors of these sites since they generally write with higher quality and provide more accurate information. Concerns include the need to focus on using more layman's terms and to continually update online content to provide patients with up-to-date diagnostics and therapeutic information. Although physician authorship is not the current trend for online information, physicians should be responsible for suggesting quality sites with better defining diagnoses and for suggesting proper search terms to their patients. Overall, our study demonstrates that improvement is needed in the regulation of the Internet to prevent flow of misinformation to a generation of patients with unlimited access to online health information.

References