Since the 1990s, the Internet has made information much more accessible. Both availability and access to health care information on the Internet have also increased. Recent studies showed that 72% of Americans have searched the Internet for health or medical information within the past year, but the quality of online material is extremely variable, with a minority of sites providing peer-reviewed material. Patients are often unaware of bias, with 52% of users believing that “almost all” or “most” health care information on the Internet is credible. With minimal regulation of these websites, patients routinely consume misinformation that may lead to anxiety and unrealistic expectations for treatment and potentially unsafe and inappropriate care.

Online information on musculoskeletal conditions has been shown to be of poor quality. Garcia et al showed that information on shoulder instability is often inaccurate and is written in overly complicated language. Similar deficiencies have been found in information on metal-on-metal total hip arthroplasty (THA) have generated much attention in the media because of early failure of certain implant systems. This study assessed the quality, accuracy, and readability of online information on metal-on-metal THA. The search terms “metal-on-metal hip replacement” and “metal hip replacement” were entered into the 3 most popular search engines. Information quality was assessed with the DISCERN score and a specific metal-on-metal THA content score. Accuracy of information was assessed with a customized score. Readability of the websites was assessed with the Flesch-Kincaid grade level score. A total of 61 unique websites were assessed. For 56% of websites, the target audience was patients. Media or medicolegal sources accounted for 44% of websites. As assessed by DISCERN (range, 16-80) and metal-on-metal THA (range, 0-25) scores, quality of the websites was moderate at best (47.1 and 9.6, respectively). Accuracy (range, 0-8) of the information presented also was moderate, with a mean score of 6.6. Media and medicolegal websites had the lowest scores for both quality and accuracy, despite making up the greatest proportion of sites assessed. Only 1 website (2%) had a Flesch-Kincaid grade level at or less than the recommended level of 8th grade. This study found that online information on metal-on-metal THA was of poor quality, often was inaccurate, and was presented at an inappropriately high reading level, particularly for media and medicolegal websites. Health care providers should counsel patients on the quality of information available and recommend appropriate online resources.

The authors are from Tallaght Hospital, Dublin, Ireland.
The authors have no relevant financial relationships to disclose.
Correspondence should be addressed to: Geoff Crozier-Shaw, MB, BCh, BAO, 2 Griver St, Cottesloe, WA 6011, Australia (gcrozshaw@hotmail.com).
Received: April 8, 2016; Accepted: October 3, 2016.
doi: 10.3928/01477447-20161116-02
cies in the quality of online information have been reported for other musculoskeletal conditions, such as carpal tunnel syndrome, back pain, lumbar disk herniation, cauda equina syndrome, fracture of the distal radius, and developmental dysplasia of the hip. A recent study of total hip arthroplasty (THA) showed that the quality of online information was poor, with 54% of websites having extensive shortcomings. Websites with a Health on the Net (HON) code endorsement appear to provide higher-quality information, with authors advising physicians to direct patients to access peer-reviewed websites that have markers of quality assurance, such as the HON code.

Metal-on-metal bearings in THA or hip resurfacing have generated significant controversy in recent years because of early high failure rates with certain implant systems, such as the ASR resurfacing system (DePuy, Leeds, United Kingdom) and the ASR XL acetabular cup (DePuy). High revision rates (up to 25% at 6 years) reported by the Australian National Joint Registry and the National Joint Registry in the United Kingdom resulted in worldwide medical device alerts, with ASR implants being recalled and health regulatory authorities initiating surveillance programs for metal-on-metal bearing implant systems. The recall and the subsequent drive for compensation for those affected resulted in extensive media coverage and increased the negative public perception of metal-on-metal THA. However, only a minority of metal-on-metal THA systems are associated with early revision, and failure to communicate this information may lead to unnecessary anxiety and suffering among patients who underwent THA with a metal-on-metal bearing or any bearing type. This study assessed the quality, accuracy, and readability of online information on metal-on-metal THA.

**Materials and Methods**

An Internet search was performed with the terms “metal-on-metal hip replacement” and “metal hip replacement” on February 15, 2015. Searches were performed in Ireland, an English-speaking nation. The search engine history and cookies were erased and reset before the study was undertaken to avoid clinically orientated bias based on previous searches. These simple terms were chosen to represent the most likely search terms that patients would use when searching the Internet for information on metal-on-metal THA. The most popular search engines, Google, Yahoo, and Bing, were used. There are used in more than 90% of Internet searches. To reflect the market share and popularity of these search engines, the first 50, 25, and 25 websites from the Google, Bing, and Yahoo searches, respectively, were analyzed. Search results were cross-referenced to create a list of unique websites for further analysis. Exclusion criteria included repeated websites (36), material unrelated to metal-on-metal THA (1), videos (2), and unreadable websites (1). Ultimately, 61 unique websites were included in the study. Websites were classified into the following categories: commercial, academic, physician, medicolegal, governmental or nonprofit, and media. Each website was assessed for the HON code, which recognizes health care information on websites and provides certification to websites that meet the required criteria. The content of each website was then assessed for quality, accuracy, and readability by 2 assessors (G.C.S., J.M.Q.).

The primary outcome was website quality, as assessed by 2 instruments, the DISCERN score and a metal-on-metal THA content score developed by the authors. Two authors independently assessed each website. The DISCERN score (range, 16-80, with a higher score indicating higher quality) is a widely used health care website quality assessment tool developed by the National Health Service Executive Research and Development program. This scale consists of 16 questions, with each question rated on a scale of 1 to 5, for a minimum score of 16 and a maximum score of 80. Websites are classified as excellent (range, 68-80), good (range, 55-67), fair (range, 42-54), poor (range, 29-41), or very poor (range, 16-28). The second score used was a metal-on-metal THA-specific content score (range, 0-25, with a higher score indicating higher quality) developed by the authors (Table 1). With this scale, 1 point was allocated for the mention of 25 predefined terms that were chosen as being relevant to metal-on-metal THA and failure of metal-on-metal THA. These terms are listed in Table 1, based on guidelines from the Medicines and Healthcare Products Regulatory Agency, the British Orthopaedic Association, and the American Academy of Orthopaedic Surgeons.

Accuracy was assessed with a scoring system used in previous studies. Two authors (G.C.S., J.M.Q.) ranked each website on a scale of 1 to 4 for accuracy, with 1 indicating that less than 25% of the content was accurate; 2 indicating that 25% to 50% of the information was accurate; 3 indicating that 50% to 75% of the information was accurate, and 4 indicating that more than 75% of the information was accurate. The scores from each author were added to provide a composite score (out of 8) for each website.

Readability was assessed with the Flesch-Kincaid grade level score, which indicates the maximal level of education (United States) that a patient must have attained to be able to read and understand the material. The recommended readability level was set at 8th grade or less. The Flesch-Kincaid grade level score was determined with an online readability calculator.

**Statistics**

Comparative statistics were used to analyze quality and accuracy scores among the groups as follows. For normally distributed data, analysis of variance with Bonferroni post hoc correction was used to assess for differences between multiple means. The Student’s t test was used to as-

---

Table 1

<table>
<thead>
<tr>
<th>Term</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal-on-metal THA</td>
<td>5</td>
</tr>
<tr>
<td>Metal-on-metal</td>
<td>5</td>
</tr>
<tr>
<td>Bearing</td>
<td>5</td>
</tr>
<tr>
<td>Hip</td>
<td>5</td>
</tr>
</tbody>
</table>

---
sess for the difference when 2 means were compared. The threshold for significance was set at $P < .05$ for all tests. Statistical analysis was performed with GraphPad Prism, version 6 (GraphPad Software, Inc, La Jolla, California).

**RESULTS**

A total of 61 unique websites were identified (Figure 1). The categories are shown in Figure 2. The primary outcome, website quality as determined by mean DISCERN score (range, 16-80, with a higher score indicating better quality), was 47.1 (SD, 11.6; range, 27-68) (Table 2), showing that the quality of websites is fair. A physician website scored highest (68), and a medicolegal website scored lowest (27). The category with the highest score was nonprofit organizations, which scored significantly higher than the overall mean (57.8 vs 47.1, $P < .05$). Media and medicolegal websites scored significantly lower than the overall mean (14 vs 9.6, $P < .05$). Nonprofit organization, physician, and academic sites scored better than media and medicolegal websites, respectively ($P < .05$, Table 2).

The second tool used to assess website quality was a score specific to metal-on-metal THA. Mean metal-on-metal content score (range, 0-25, with a higher score indicating better quality) was 9.6 (SD, 4.6; range, 2-23) (Table 2). A nonprofit website had the highest score (23), whereas a media website had the lowest score (2). Overall, the nonprofit organization category had the highest score, significantly higher than the overall mean (14 vs 9.6, $P < .05$). Media websites scored lowest, with a mean score significantly lower than the overall mean (5.6 vs 9.6, $P < .05$). Nonprofit organization, physician, and academic sites scored better than media, commercial, and medicolegal categories, respectively ($P < .05$, Table 2).

For website accuracy (range, 0-8, with a higher score indicating greater accuracy), mean score was 6.6 (SD, 1.2) (Table 3). The physician category had information accuracy of greater than 75% (score,
≥6) across all websites reviewed. The media category showed the poorest performance, with a mean score of 5.8 (SD, 0.8), and only 50% of the information was accurate. Medicolegal, media, and commercial categories scored significantly lower than academic, physician, and nonprofit organization categories (\(P<.05\)).

For the final outcome, website readability, mean Flesch-Kincaid grade level score was 11.7 (range, 5.7-12; SD, 0.88) (Table 4), indicating that most of the websites had readability at the 11th grade level or above. Only 1 website (physician) scored less than the recommended 8th grade readability level. No significant differences were found in Flesch-Kincaid grade level scores between categories of websites (\(P>.05\) for all comparisons, Table 4).

**DISCUSSION**

This study showed that information on metal-on-metal THA that is available online is generally of poor quality, with moderate accuracy. In addition, most sites are written above the recommended 8th grade reading level. These findings are concerning because of the increasing Internet usage by patients worldwide as a resource for health care information. The use of the Internet by patients has led to a significant change in the patient-physician dynamic, with many patients attending consultations with information of variable quality about health conditions and treatment options. Inaccurate information may lead to frustration for both patient and physician and may lead to unrealistic expectations for treatment or unnecessary anxiety.3,32 Numerous studies have shown deficiencies in the quality of online information on a wide range of conditions and treatments, including chronic pain, chronic obstructive pulmonary disease, depression, and ear, nose, and throat surgery.8-9,33,34

This study showed that overall website quality was moderate at best, with a mean DISCERN score of 47.1 and a metal-on-metal THA content score of 9.6. Website authorship had a significant effect on quality, with media and medicolegal websites having significantly lower scores (DISCERN score, 39.2). This finding indicated that these websites were unlikely to refer to their source of evidence, were likely to be biased, and did not provide complete information, such as all treatment options. In contrast, nonprofit organization, physician, and academic websites scored significantly higher. Accuracy (contrasted with completeness) was also variable, with website authorship again having a significant effect on scores. Media, medicolegal, and commercial sites scored significantly lower than nonprofit organization, physician, and academic websites. This finding is in agreement with a previous study of information on the Internet on THA in general, in which nonprofit organization, academic, and physician websites were most likely to provide information of good to excellent quality.35

In terms of readability, only 2% of websites in this study were at the recommended 8th grade level or below, and this finding is similar to findings on websites on other health topics, such as shoulder...
instability, in which only 12% of websites were written at less than the 8th grade level. It may be unreasonable to expect that all website categories will be written at an 8th grade level because some websites are not necessarily directed at patients (ie, academic websites). However, this study showed no difference between categories, with most websites written at the 11th grade level or above. This finding indicated that almost all websites, including nonprofit organization and physician categories, present information at an excessively high readability level.

Similar studies showed variable quality, accuracy, and readability of online information on minimally invasive THA and hip resurfacing and other musculoskeletal topics, such as shoulder instability. However, the current findings differ from the other studies in the context in which patients were likely to search online for information on metal-on-metal THA. Although most studies to date have assessed online information on a particular condition and its treatment, none have assessed online information on a major controversy with a treatment or procedure. The background to this study is the DePuy hip recall in 2010 in which ASR hip resurfacing and ASR XL acetabular cup systems were recalled because of excessively high failure rates. This recall led to significant media attention, with patients who had THA in situ uncertain whether they had metal-on-metal implants and, if so, what the implications were. Inaccurate information in this setting can be particularly harmful and can lead to significant anxiety among patients, with some media websites at the time referring to THA “poisoning patients,” “toxic” THA components, and an increased risk of “tumor” formation. Of note, in the current study, media websites were both the most popular category (28% of websites) that offered information on metal-on-metal THA and had the lowest scores in terms of both quality and accuracy of information.

An ideal website providing information for patients on metal-on-metal THA should follow the DISCERN guidelines for health and medical sites on the Internet. In short, websites should provide information that is complete and accurate, is written using layman’s terminology, and has clear references to relevant sources, such as local guidelines and authorship of the website. Any relevant conflicts of interest should be disclosed. Specifically, the website should outline the general concept of metal-on-metal THA in broad terms, including the specific benefits and risks of metal-on-metal THA and alternative treatment options (eg, other types of THA implants). Specific information also should be provided that addresses the metal-on-metal THA recall, possible outcomes (positive and negative) in patients who have had a metal-on-metal THA, and how to seek assessment if they have a metal-on-metal THA. Physicians must be able to direct patients to appropriate websites for information. A study of 1200 patients showed that 4% received recommendations from their physician, although 70% stated that they would like to receive this information.

Limitations

Limitations of the current study include the fact that the search results represented a snapshot in time and location. In addition, the search results were in English only and the search was performed in 1 country only. Searches performed in different locations (eg, different countries) would likely provide different results because search algorithms are locally oriented. Similarly, the order and content of search results change over time as search algorithms give different results to reflect new online material and increasing popularity of certain websites. Further, limiting the search to the top 50 websites for Google and the top 25 for the Yahoo and Bing search engines, respectively, may exclude higher-quality information. However, Internet users are most likely to access the first few websites that are presented, rarely going beyond the first 25 results. Allowing for these limitations, the authors still believe that the websites generated from the search used in this study represent an accurate snapshot of those that a typical patient would access.

Another potential limitation is the inclusion of terms such as “metallosis, hypersensitivity” in the metal-on-metal THA content score. It is possible that inclusion of these terms may result in a high readability score. However, these terms should be included because patients may encounter them when searching the Internet. Where these terms are included, they...
should be appropriately deconstructed and simplified to facilitate patient comprehension. Including these terms on websites is unlikely to affect the Flesch-Kincaid grade level score because it is calculated based on words per sentence and syllables per word. In an article that includes several hundred words, occasional (or single) references to more complex words, such as “metalllosis” or “hypersensitivity,” are not likely to influence the Flesch-Kincaid grade level score significantly. Finally, subjective scores were used to assess quality and accuracy, with the possibility of detection bias. To reduce this risk, 2 authors were used to generate mean scores for DISCERN, metal-on-metal THA, and accuracy scores.

**Conclusion**

To the authors’ knowledge, this is the first study to assess the quality and accuracy of online information on metal-on-metal THA. Information was of moderate quality and accuracy. Although using the Internet helps patients to empower themselves in making health care decisions, the potential effect of harmful misinformation also must be considered. Physicians must recognize this potential and counsel patients on these risks. Further, in the case of metal-on-metal THA, physicians should direct patients toward appropriate and accurate online resources to reduce the risk that patients will access potentially harmful misinformation.

**References**

30. Cotugna N, Vickery CE, Carpenter-Haefele KM. Evaluation of literacy level of patient


