Should I Ask Doctor Google? Reliability of Israeli Web-based Information Regarding General Orthopedic Injuries and Symptoms

Yaniv Steinfeld MD[1,2], Omri Besor MD[3], Yaniv Yonai MD[1], Yaron Berkovich MD[1,2], Elad Apt MD[4], Nadav Rinott MD[4], and Yaniv Keren MD[4]

[1]Department of Orthopedics B, Hillel Yaffe Medical Center, Hadera, Israel
[2]Rappaport Faculty of Medicine, Technion–Institute of Technology, Haifa, Israel
[3]Medical Corps, Surgeon General's Headquarters, Israel Defense Forces, Ramat Gan, Israel
[4]Department of Orthopedics, Rambam Health Care Campus, Haifa, Israel

ABSTRACT **Background:** Websites serve as a source of medical information for a large part of the public, some claim to be a substitute for a physician's consultation. Many patients meet a physician after conducting internet research. Medical staff are concerned that internet sources of information are unreliable and may lead to erroneous decisions by patients.

Objectives: To examine the reliability of web-based sources of information (through the Google™ search engine) regarding five common orthopedic complaints and injuries.

Method: We performed a search of five common orthopedic complaints and injuries using the Google search engine. The reliability of web-based information was measured by the DISCERN tool, which is a valid and verified tool for examining the reliability of medical information sources to the public. The reliability of 47 websites was examined by two orthopedic surgeons and two senior residents.

Results: The overall average score given to the sites was 2.8, on a scale of 1 to 5. We found that the higher the site appeared in the search results, the higher the quality of its information. Commercial sites scored higher than general internet information sources.

Conclusions: The internet network is a very broad source of information. For those who lack scientific education and training it is not easy to distinguish between reliable and unreliable or biased sources. The trend of searching for medical information and self-healing is increasing. We must strengthen the network with reliable sources by creating official scientific position papers by medical teams and promoting them online.

IMAJ 2023; 25: 688–691

KEY WORDS: internet, orthopedic injuries, reliability, quality, website information

Web-based searches today serve as a source of medical information for a large part of the public. Some claim it is a substitute for a physician's consultation. Many patients come to the physician's office after conducting extensive web-based research by themselves [1]. Such web-based research exists in all areas of medicine, including orthopedic surgery [2,3]. Past scientific literature has indicated that there is a significant difference in the use of web-based information among different populations, which may create inequity in treatment [4]. The inequity is narrowing because most searching is conducted via the popular search engine Google™. In fact, as early as 2014, over 75% of the population used the internet to search for medical information, and 92% used Google's search engine for this purpose [5]. In addition, according to recent research, the sites that appear on the first page of the search (10 selected results) receive about 88% of the hits in a search for a specific result. However, medical staff are concerned and claim that some of the web-based information is not reliable. As a result, patients may make wrong decisions. In addition, a physician must address the excess of uncontrolled information that a patient reads. Some medical staff see this as a threat to their authority [6-8].

In 2022, Byrne and colleagues [9] examined the quality of medical information on webpages related to parathyroid diseases. They reviewed the top 75 Google search results for *parathyroidectomy*, *parathyroid surgery*, and *parathyroid gland removal*. They found that most websites were of poor quality. Sullivan and co-authors [10] examined Google search results regarding stem cell therapy and knee osteoarthritis. They found that search results often directed the user to websites lacking transparency and verifiable information. Özbek et al. [11] examined the reliability of medical information regarding carpal tunnel syndrome from YouTube, another relatively popular internet platform. They found videos that were of low quality and low reliability. Our literature search revealed no similar studies on medical information sources in Hebrew.

OBJECTIVES

The aim of our study was to examine the reliability of internet sources of information available on the Google search engine regarding five common orthopedic complaints and injuries.

PATIENTS AND METHOD

We examined the quality of internet information sources in quantitative terms using the DISCERN tool [12].

The study was conducted on the first 10 results of five common orthopedic complaints and injuries using the Google search engine. The terms included *hip degenerative arthritis*, *herniated lumbar disc*, *Achilles tendon rupture*, *Meniscus tear*, and *recurrent ankle sprains*. The search was performed by a new user after clearing the browsing history on the computer to prevent search biases.

A total of 50 webpages were analyzed regarding their reliability level. The inclusion criteria were:

a. Appearance on the first page of the search results (non-paid results)
b. Hebrew-based webpage
c. Free access to the site

The exclusion criteria were:

a. Pay-only website
b. Non-verbal page (a link to a video, for example)
c. Site that appeared in the search results as an advertisement

We then subdivided the sites into groups based on their origin: institutional website (such as a public hospital website), commercial website, and general website (such as Wikipedia, a free online encyclopedia that is created and edited by volunteers and hosted by the Wikimedia Foundation). Those sites were then subdivided according to their position in the search results: high (places 1–3), middle (4–6), and low (7–10). The text from the websites was copied. All identifications were removed from the file to prevent bias.

The files were sent to two senior orthopedic surgery residents and two orthopedic surgery specialists along with an accompanying questionnaire used to score each website. The score was given using the DISCERN tool. DISCERN is designed to help users of consumer health information determine the quality. The DISCERN questionnaire consists of 16 questions divided into three scores: reliability, treatment options, and overall quality score (question 16). The questions are designed to characterize the quality of medical information and its reference to treatment options on a Likert scale of 1–5. The questionnaire consisted of questions that evaluated the site on a Likert-type scale of 1 (absolutely not reliable) to 5 (completely reliable).

The research tool was translated into Hebrew. Questions 1–8 were considered part of reliability, while 9–15 tested quality. Question 16 rated overall quality. The scores ranged from 16–80. A score of 80 indicated maximum quality. Before ranking the webpages, the evaluators read the instruction leaflet for DISCERN and made sure they understood how it was used by the lead researcher. The DISCERN scores were independently applied by two reviewers whose results were compared to determine the quality of a publication written about treatment questions. When significant discrepancies were present, the two reviewers discussed the results and came to agreement.

STATISTICAL ANALYSIS

Statistical analyses were performed using IBM Statistical Package for the Social Sciences statistics software, version 26 (SPSS, IBM Corp, Armonk, NY, USA). The degree of agreement between the reviewers through kappa regarding the results of the DISCERN was calculated. We qualitatively assessed the level of reliability of the articles that were listed first in the search results and compared the different groups of sites using tools of descriptive statistics, which were then compared among the three study groups by the ANOVA test and ANOVA-post HOC. A correlation was calculated to test whether there was a relationship between the location of a source of information within the results in the search results and its reliability score.

RESULTS

We included 47 webpages in the study; three did not meet the inclusion criteria. In the search for the complaint Achilles tendon rupture, two resulting websites were completely irrelevant to the subject and therefore were excluded. In the search for the complaint *herniated disc*, one result was a link to an explanatory video and therefore not comparable. It was excluded from the study.

A significant number of sites in the search results were marketing sites from orthopedic surgeons or other commercial sites (23/47, 48.9%). For certain orthopedic issues, the relative share of commercial sites was greater than for other topics [Figure 1]. In the search for terms *hip degenerative arthritis* and *recurrent ankle sprains*, commercial results increased by 60% compared to 33.3% for *herniated lumbar disc*.

Figure 1. Sites subdivision by their type

Blue = commercial, Red = general, Green = institutional

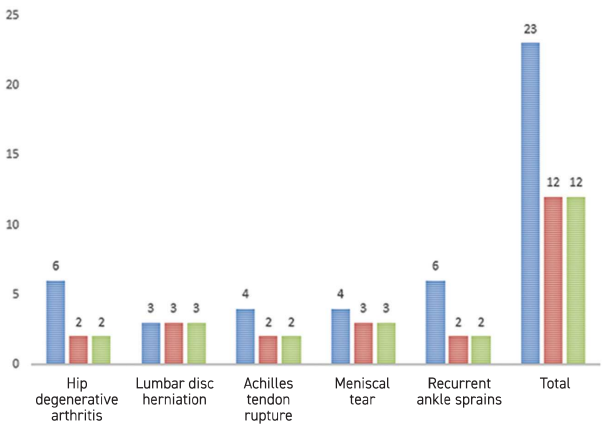

Table 1. The average score for each question, based on a Likert-type scale of 1–5

Question	Senior surgeon	Resident
	Mean ± SD	Mean ± SD
1. Are the aims clear?	3.81 ± 0.88	4.23 ± 0.45
2. Does it achieve its aims?	3.52 ± 1.00	3.89 ± 0.71
3. Is it relevant?	3.95 ± 0.83	4.26 ± 0.62
4. Is it clear what sources of information were used to compile the publication (other than the author or producer)?	1.66 ± 1.30	1.64 ± 1.24
5. Is it clear when the information used or reported in the publication was produced?	1.69 ± 1.25	1.72 ± 1.35
6. Is it balanced and unbiased?	3.00 ± 1.41	3.38 ± 1.29
7. Does it provide details of additional sources of support and information?	1.59 ± 0.65	1.46 ± 0.61
8. Does it refer to areas of uncertainty?	2.48 ± 0.92	2.26 ± 0.85
9. Does it describe how each treatment works?	2.93 ± 1.10	3.11 ± 1.14
10. Does it describe the benefits of each treatment?	3.11 ± 1.14	3.25 ± 1.03
11. Does it describe the risks of each treatment?	2.47 ± 1.20	2.32 ± 1.05
12. Does it describe what would happen if no treatment was offered?	2.62 ± 1.19	2.29 ± 1.00
13. Does it describe how the treatment choices affect overall quality of life?	2.89 ± 0.92	2.92 ± 0.95
14. Is it clear that there may be more than one possible treatment choice?	3.53 ± 1.74	3.54 ± 1.35
15. Does it provide support for shared decision making?	2.44 ± 1.00	2.47 ± 0.94
16. Based on the answers to all the above questions, rate the overall quality of the publication as a source of information about treatment choices.	2.76 ± 1.24	2.88 ± 1.01

SD = standard deviation

The average scores for each question are listed by senior surgeons vs. residents in Table 1. When we compared the scores, we found several statistically significant factors. There was a variance between the evaluators, mainly between the residents ($P < 0.05$). When we compared groups of questions (reliability questions 1–8, quality questions 9–15) we found that there was statistically significant variance between the residents and the senior surgeons ($P = 0.008$) as well as between the senior surgeons themselves ($P = 0.03$) and between the residents themselves ($P = 0.004$).

When we tested the variance using the kappa test for agreement between examiners, the test was clear, with a level of agreement between the examiners of 0.5. That is, in half of the cases residents and senior surgeons agreed with each other regarding results.

When we compared groups of websites (general vs. commercial vs. institutional), we found that the general sites had a higher overall average score (48.47) compared to the commercial sites (42.6) and the institutional sites (46.2). These differences were statistically significant. There was a significant difference between the information from commercial websites and general websites ($P = 0.004$) and between the general and the institutional ones in the reliability score of the residents ($P = 0.02$). There was a significant difference between the information from the advertisement sites and general sites with a reliability score of the senior surgeons ($P = 0.046$). There was a significant difference between the information from the commercial sites and general sites ($P = 0.01$) and between the general and institutional sites ($P = 0.029$) in indicating the reliability.

When examining the results by position in the search results, we found some differences. The average overall score was 47.6 in the high position group, 45.5 in the middle, and 42.5 in the low. When we examined the correlation between the location of each site and the results of the questionnaire, no correlation was found between location and score. When we eliminated question number 1 from the results (Are the aims clear?), a significant correlation was found of medium intensity: the lower the rating of the page, the lower the score.

DISCUSSION

We assessed the quality and reliability of web-based orthopedic related information. Five common orthopedic injuries and complaints were selected to represent orthopedic related information accessible to the public. We tested the first 10 sites in each search result (excluding sponsored results) because in at least 88% of the cases users chose to further examine these results. Most of the sites (96%) detected by this method were completely relevant to the search question.

Our results showed that the accessible information is inaccurate and even partially incorrect. The average score given to the final question (Based on the answers to all the above questions, rate the overall quality of the publication as a source of information about treatment choices.) was 2.8 of 5, an unsatisfactory score in our opinion. Even when looking at each question individually, very few questions received an average score over 4, and they were all in the first part (the reliability section). In the second section of the questionnaire representing the quality aspect, most of the assessments had a score of less than 3.

A conspicuous result of the study was that the commercial sites were less reliable than the other sites. The commercial sites consisted mainly of personal pages of surgeons and of companies that sell a certain product (e.g., accessories for conservative treatment). When examining the pages, it seemed that large parts were similar and therefore received similar scores. Our conclusion was that those websites should have relied more on an up-to-date literature review given the added value that every publisher would like to have. A prominent site that was repeated in each of the search results was Wikipedia. Wikipedia results were highly valued both in terms of quality and credibility from the evaluators.

Another finding in our study was the increase in reliability and quality of a site based on its position in the search results. Since we were aware that many websites are promoted and marketed by professionals, our hypothesis about this finding was that professional sites were promoted more. These professional sites may have had more funding with which to promote the site and write the content. We concluded that a site is more likely to be trustworthy if it is listed high on the search results.

Originally our intention was to examine readability on the webpages using scientific tools. It was not possible to analyze the data because those tools are validated appropriately only in the English language and are unique to it. Those tools are not suitable for use in the Hebrew language.

This study has several limitations. First, the DISCERN tool is not validated to the Hebrew language and we used translations. In addition, we saw a relatively large variance between the evaluators, both between the residents and the senior surgeons, but also among the evaluators, which showed the possibility of less credibility in the use of the DISCERN tool in Hebrew. We believe that it is essential to expand research to other topics and with a greater number of evaluators and more articles. Validation for this tool is also needed.

CONCLUSIONS

The internet is a very broad source for information. Professional medical staff often use it as a substitute for a formal scientific literature search, as does the public. For those who lack training and scientific education, it is not easy to distinguish between reliable sources of information and those that are less reliable or biased. The trend of searching for medical information and self-healing is not going to change, so experts must strengthen the internet network with reliable sources of information. This task can be completed by professional medical societies endorsing position papers and promoting them to the public on the internet.

Correspondence

Dr. Y. Steinfeld
Dept. of Orthopedic B Hillel, Yaffe Medical Center, Hadera 38425, Israel
Phone: (972-4) 774-4738
Fax: (972-4) 774-4736
Email: yanivsteinfeld@gmail.com

References

1. Cline RJ, Haynes KM. Consumer health information seeking on the internet: the state of the art. *Health Educ Res* 2001; 16 (6): 671-92.
2. Cassidy JT, Baker JF. Orthopaedic patient information on the world wide web: an essential review. *J Bone Joint Surg Am* 2016; 98 (4): 325-38.
3. Proude EM, Shourie S, Conigrave KM, Wutzke SE, Ward JE, Haber PS. Do elective surgery patients use the internet to look for information about their condition? *ANZ J Surg* 2004; 74 (5): 304-7.
4. Walsh KP, Rehman S, Goldhirsh J. Disparities in internet use among orthopedic outpatients. *Orthopedics* 2014; 37 (2): e133-40.
5. Aitken M, Altmann T, Rosen D. Engaging patients through social media. Parsippany (NJ): IMS Institute for Healthcare Informatics. 2014.
6. Sommerhalder K, Abraham A, Zufferey MC, Barth J, Abel T. Internet information and medical consultations: experiences from patients' and physicians' perspectives. *Patient Educ Couns* 2009; 77 (2): 266-71.
7. Dieterich A. The modern patient - threat or promise? Physicians' perspectives on patients' changing attributes. *Patient Educ Couns* 2007; 67 (3): 279-85.
8. McMullan M. Patients using the internet to obtain health information: how this affects the patient-health professional relationship. *Patient Educ Couns* 2006; 63 (1-2): 24-8.
9. Byrne J, Keogh S, Cullinane C, Razzaq Z, Redmond HP. Readability and quality of online health information regarding parathyroidectomy. *OTO Open* 2022; 6 (4): 2473974X221133308.
10. Sullivan B, Platt B, Joiner J, Jacobs C, Conley C, Landy DC, Stone AV. An investigation of Google searches for knee osteoarthritis and stem cell therapy: what are patients searching online? *HSS J* 2022; 18 (4): 485-9.
11. Özbek EA, Armangil M, Karaca MO, Merter A, Dursun M, Kocaoğlu H. Evaluation of the reliability and quality of information in carpal tunnel syndrome shared on YouTube. *J Wrist Surg* 2021; 11 (4): 295-301.
12. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999; 53 (2): 105-11.

Learning is acquired by reading books; but the much more necessary learning, the knowledge of the world, is only to be acquired by reading man, and studying all the various editions of them.

Philip Dormer Stanhope, 4th Earl of Chesterfield (1694–1773), British statesman, diplomat, man of letters, and an acclaimed wit of his time