JSES International 8 (2024) 53–57

Contents lists available at ScienceDirect

JSES International

journal homepage: www.jsesinternational.org

Evaluation of information from artificial intelligence on rotator cuff repair surgery

Eric Warren Jr., BS[a], Eoghan T. Hurley, MB, MCh, PhD[b], Caroline N. Park, MD[b], Bryan S. Crook, MD[b], Samuel Lorentz, MD[b], Jay M. Levin, MD, MBA[b], Oke Anakwenze, MD, MBA[b], Peter B. MacDonald, MD, FRCSC[c], Christopher S. Klifto, MD[b,*]

[a]Duke University School of Medicine, Duke University, Durham, NC, USA
[b]Department of Orthopaedic Surgery, Duke University, Durham, NC, USA
[c]Section of Orthopaedic Surgery & The Pan Am Clinic, University of Manitoba, Winnipeg, MB, Canada

ARTICLE INFO

Keywords:
Rotator cuff repair
Artificial intelligence
ChatGPT
Natural language processing
Shoulder
Patient information

Level of evidence: Education Methodology Study; Validation of Computer Sources

Purpose: The purpose of this study was to analyze the quality and readability of information regarding rotator cuff repair surgery available using an online AI software.
Methods: An open AI model (ChatGPT) was used to answer 24 commonly asked questions from patients on rotator cuff repair. Questions were stratified into one of three categories based on the Rothwell classification system: fact, policy, or value. The answers for each category were evaluated for reliability, quality and readability using The Journal of the American Medical Association Benchmark criteria, DISCERN score, Flesch-Kincaid Reading Ease Score and Grade Level.
Results: The Journal of the American Medical Association Benchmark criteria score for all three categories was 0, which is the lowest score indicating no reliable resources cited. The DISCERN score was 51 for fact, 53 for policy, and 55 for value questions, all of which are considered good scores. Across question categories, the reliability portion of the DISCERN score was low, due to a lack of resources. The Flesch-Kincaid Reading Ease Score (and Flesch-Kincaid Grade Level) was 48.3 (10.3) for the fact class, 42.0 (10.9) for the policy class, and 38.4 (11.6) for the value class.
Conclusion: The quality of information provided by the open AI chat system was generally high across all question types but had significant shortcomings in reliability due to the absence of source material citations. The DISCERN scores of the AI generated responses matched or exceeded previously published results of studies evaluating the quality of online information about rotator cuff repairs. The responses were U.S. 10th grade or higher reading level which is above the AMA and NIH recommendation of 6th grade reading level for patient materials. The AI software commonly referred the user to seek advice from orthopedic surgeons to improve their chances of a successful outcome.

Overall prevalence of rotator cuff abnormalities increases with age, from 9.7% in patients less than 20 year old to 62% in patients of 80 years and older.[37,38] Despite the widespread prevalence, the decision to proceed with conservative or surgical management of a rotator cuff tear is not always obvious. Generally, younger patients with acute or acute on chronic full-thickness tears with loss of function have the greatest risk of disease progression and the highest rate of tendon healing, making them strong candidates for early rotator cuff repair.[5,13,18,29,36] Furthermore, patients with small partial-thickness or full-thickness tears, advanced rotator cuff muscle fatty infiltration, and degenerative tears if older than 65 year old may be better suited for conservative treatment prior to considering surgery.[5,8,18,25,28] For medium-risk disease in the middle of the two extremes, shared-patient decision-making with an informed discussion of conservative and surgical management options may be warranted.[18]

Although patients are often given information resources both physically and virtually by providers, patients often take it upon themselves to search for medical information on the internet. Previously published evaluations of internet information about rotator cuff repairs have found the available information to be low quality, often outdated, and written at a reading level too advanced for the general population. However, websites or videos produced by academic institutions or physicians were generally of higher quality.[6,9,12,22] In recent years, artificial intelligence (AI) natural

Institutional review board approval was not required for this study.

*Corresponding author: Christopher S. Klifto, MD, Duke University Department of Orthopaedic Surgery, 3609 SW Durham Drive, Durham, NC 27707, USA.
E-mail address: christopher.klifto@duke.edu (C.S. Klifto).

https://doi.org/10.1016/j.jseint.2023.09.009

language processing models have become increasingly accessible and popular due to the immediate production of conversation-style responses to a wide range of questions. As the usage of these AI chat bots rises, it is important to evaluate the reliability and quality of this resource to effectively counsel patients on credible online sources.

The purpose of this study was to evaluate the reliability, quality, and readability of information from a free, online AI natural language processing model (ChatGPT) regarding rotator cuff disease and rotator cuff repair using previously validated scoring systems. Our hypothesis was that the AI software would generate answers with poor quality and readability scores regardless of question theme.

Methods

Artificial intelligence and question input

In August 2023, a free AI language processing software, ChatGPT (GPT-3.5), was queried with twenty-four commonly asked questions from patients on rotator cuff repair. The list of questions were curated from a combination of previous studies into patient internet searches about rotator cuff repair and sports medicine.[15-17,19] Questions were categorized according to the Rothwell classification into one of three themes that identify what type of problem is being queried—fact, policy, or value.[16,30] As defined by Rothwell, a question of fact asks whether something is true, and to what extent. Objective evidence can and should be employed to answer these questions. A question of policy asks whether a specific course of action should be undertaken to solve a problem. Finally, a question of value asks for an evaluation of an object, idea, event, or person.[30] The full list of questions is available in Table I and the AI generated responses are available in Supplementary Appendix S1.

Quality analysis

The quality of responses in each group was assessed as a whole using 2 validated tools: DISCERN score and Journal of the American Medical Association (JAMA) benchmark criteria.

The DISCERN instrument is a valid and reliable quality assessment tool of written patient information about management options for a medical problem commonly used in the literature.[7,9,22,33] DISCERN was funded by The British Library and the National Health Service Research and Development Programme and produced by a panel consisting of both clinical experts and nonphysician consumer health literature experts.[7] It comprises three sections: 8 questions on reliability, 7 questions on treatment information and lastly an additional overall quality rating. Each question is scored 1-5, with 1 representing a “definite no” with question not being fulfilled at all, while 5 represents a “definite yes” with complete fulfilment of the quality criteria. Scores of 2-4 represent the material partially meeting the quality criteria, with the specific number judged by the raters based on the extent of the shortcomings. With a maximum score of 80, a score greater than 70 is classified as “excellent” and a score greater than 50 is classified as “good”. Two authors (EW and EH) scored the responses independently, and then openly discussed discrepancies until final scores were agreed upon.

The JAMA[34] benchmark criteria comprises four axiomatic standards to assess the quality of information: authorship, attribution, disclosure and currency; with each standard scoring 1 point. ‘Authorship’ is important in allowing the reader to identify the origin of the information, ‘Attribution’ deals with content referencing, ‘Affiliation’ addresses any potential conflict of interest an author might have and ‘Currency’ addresses whether or not the content is current and therefore relevant to the reader.

Readability analysis

The readability of responses in each group was assessed as a whole using the Flesch-Kincaid Reading Ease Score (FRES)[11] and Flesch-Kincaid Grade Level (FKGL). The FRES is a tool that generates a score from 0 (unreadable) to 100 (very easy to read) for a given input using the formula 206.835 - 1.015*(total words/total sentences) - 84.6*(total syllables/ total words). Flesch then later adapted this into the FKGL, which denotes the minimum level of U.S.-based schooling a patient must have obtained to be able to read the material, ie, higher FKGL correlates with more difficult to understand. The formula for FKGL is as follows: 0.39*(total words/total sentences) + 11.8*(total syllables/total words) - 15.59.

Results

Itemized DISCERN scoring is available for each question class in Table II.

Fact questions

The JAMA Benchmark criteria score was 0, as there was no referencing to available source material used to compile the answers. The DISCERN score was 51, this is considered a “good” score. The reliability portion of the DISCERN score was low, owing to the lack of source material references. Conversely, the quality of information portion scored high. The FRES was 48.3., and the FKGL was 10.3, considered to be slightly above a 10th grade reading level.

Policy questions

The JAMA Benchmark criteria score was 0, as there was no reference of source material used to compile the answers. The DISCERN score was 53, qualified as a “good” score. Again, the reliability portion of the DISCERN score was very low since it is not possible to assess the quality of information sources or assess

Table I
Full question list sorted by Rothwell classification.

Fact—whether something is true, and to what extent
1. Can an x-ray show rotator cuff tear?
2. How do you go to the bathroom after shoulder surgery?
3. How can I tell if I tore my rotator cuff?
4. How much does a rotator cuff surgery cost?
5. Can you wear a bra after shoulder surgery?
6. How long after shoulder surgery can I drive?
7. How long do you have to sleep in a recliner after shoulder surgery?
8. What is the average recovery time for rotator cuff surgery?
Policy—whether a specific course of action should be undertaken to solve a problem
1. What happens if a torn rotator cuff goes untreated?
2. Can you wait too long for rotator cuff surgery?
3. Can a rotator cuff tear heal on its own?
4. How can I speed up recovery after rotator cuff repair?
5. Will a rotator cuff repair prevent me from getting arthritis?
6. What should I do to get back to playing sports after rotator cuff repair?
7. What happens if I don’t do physical therapy after rotator cuff repair?
8. How long can I wait before I get a rotator cuff repair?
Value—an evaluation of an object, idea, event, or person
1. Is arthroscopic shoulder surgery worth it?
2. Why is rotator cuff surgery so painful?
3. How long does a rotator cuff repair last?
4. Why does a rotator cuff tear hurt more at night?
5. Will a rotator cuff repair return my shoulder back to the way it was before?
6. Should I still get a rotator cuff repair if I can't afford physical therapy?
7. Can I still get a rotator cuff repair if I don't have someone to help me at home?
8. Is rotator cuff surgery major surgery?

Table II
Itemized DISCERN scores by question class.

Rothwell classification	Fact	Policy	Value
SECTION 1			
Is the publication reliable?			
1. Are the aims clear?	1	1	1
2. Does it achieve its aims?	1	1	1
3. Is it relevant?	5	5	5
4. Is it clear what sources of information were used to compile the publication (other than the author or producer)?	1	1	1
5. Is it clear when the information used or reported in the publication was produced?	1	1	1
6. Is it balanced and unbiased?	2	2	2
7. Does it provide details of additional sources of support and information?	1	1	1
8. Does it refer to areas of uncertainty?	5	5	5
SECTION 2			
How good is the quality of information on treatment choices?			
9. Does it describe how each treatment works?	4	4	5
10.Does it describe the benefits of each treatment?	5	5	5
11. Does it describe the risks of each treatment?	3	4	4
12. Does it describe what would happen if no treatment is used?	4	5	5
13. Does it describe how the treatment choices affect overall quality of life?	5	5	5
14. Is it clear that there may be more than one possible treatment choice?	5	5	5
15. Does it provide support for shared decision-making?	5	5	5
16. Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices	3	3	4
TOTAL	51	53	55

references for bias. The quality of information portion of the DISCERN score again received high marks. The FRES was 42.0, and the FKGL was 10.9, just below an 11th grade reading level.

Value questions

The JAMA Benchmark criteria score was 0, as there was no referencing to available source material used to compile the answers. The DISCERN score was 55, still considered a "good" score. Similar to the fact and policy questions, the absence of references precluded any high scores on the reliability portion of the DISCERN score. The FRES was 38.4, and the FKGL was 11.6, between an 11th and 12th grade reading level.

Discussion

The most important finding from this study was that the quality of information provided by the open AI chat system was generally good across all three questions classes but had significant shortcomings in reliability of information sources and readability. The DISCERN subsections were somewhat dichotomous for the three question groups with the reliability portion of the score receiving close to the minimum score while the treatment information portion of the score was consistently close to full marks. Our hypothesis was partially correct. The answers required a reading level too high for the general public but proved to be of good quality across all question types, with policy and value subsections scoring higher than the responses to questions of fact. It would be reasonable to think that questions of fact would be the easiest for the open AI software given the ability to provide objective evidence to formulate an answer. However, this study revealed that the AI system can also effectively generate more nuanced reasoning ie, often required to provide advice for frequently subjective questions of policy and value.

AI is exploding in popularity among the general population, and its number of use cases is rapidly expanding due to its ability to immediately synthesize fluent information from unknown sources. The current version of ChatGPT is trained from all publicly available online data prior to 2022 using a machine-learning process called reinforcement learning with human feedback.[23] In medicine, it has shown potential as an academic writing tool, sometimes fooling both humans and other AI platforms.[3,4,10,32] The rising popularity of AI is partially due to the viral news stories of impressive feats such as passing law school final exams and the United States Medical Licensing Examination with high levels of insight in its answer explanations.[21] AI has already shown the potential for impacts in orthopedics beyond writing such as the creation of outcome prediction algorithms in trauma, imaging interpretation, and extraction of data from electronic health records.[3,14,20,26,27] Importantly, the ChatGPT interface specifically discloses that the AI system may occasionally generate incorrect or harmful content, and our study found that the responses consistently recommend speaking to an orthopedic surgeon for diagnosis and treatment. Regardless, there is an increasingly prevalent call for curtailing the use of AI technology in medicine largely due to its lack of transparency.[1,24,35] The AI chat bot in this study never referenced any source material, earning a JAMA benchmark criteria score of 0 across the board. While we did not explicitly ask for accommodating references, this specific AI model has been shown to produce completely fictitious or incorrect references when asked to.[2,3,31] Although AI technology performs impressively, it should be used with extreme caution, especially when the user is untrained in the subject.

The internet provides a wealth of medical information that otherwise would have been very difficult for patients to access. However, there is no regulation on what can be published online, allowing the circumvention of the peer-review process. Studies have found some high-quality medical information on the internet, but commonly these resources require a reading level too high for a layperson, are tedious to locate, or are formatted in a way that does not appeal to the public.[6,9,22] Dalton et al evaluated 59 websites returned after searching "rotator cuff tear" on popular internet search engines. They calculated the DISCERN, JAMA, FRES, and FKGL scores, and found a mean DISCERN score of 39.47, mean JAMA score of 1.72, and mean reading grade level of above 9.[9] Interestingly, they also noted no difference in quality between websites written by physicians and those with a nonphysician author or unlisted author. Although the DISCERN score in the current study was higher than the 59 websites included in Dalton's study, the lack of references provided by the AI chat bot prevents the conclusion that the information was collected from more quality sources than websites alone. At least on the topic of rotator cuff disease, it appears that the open AI chat bot provides higher quality information on the diagnosis and management of rotator cuff tears than the average website returned on a search engine, albeit at the cost of worse readability scores. Readability is a major access concern, as

the average reading level among US adults is no higher than the eighth-grade level, and the AMA and NIH recommend a 6th grade reading level for information intended for laypeople.[39]

Furthermore, Lawson et al performed a comprehensive analysis of 150 websites concerning rotator cuff repairs and found an overall mean DISCERN score of 44, with academic institution-affiliated websites scoring a mean of 51.6 while websites controlled by private physician groups scored a mean of 40.7. They also found a FRES score of 50.17 on average, corresponding to a mean grade level of 10.98; however, no correlation was found between website readability and DISCERN score.[22] They also found that only 26% of websites cited peer-reviewed sources in any capacity; however, this still stands in contrast to no peer-reviewed references at all in the current study. Finally, in a systematic review of online orthopedic sports medicine information quality assessments, Schwarz et al found similar results, with a mean DISCERN score of 40.55 and FKGL of 10.24.[33] The open AI natural language processing model evaluated in this study performed better than the average website found through search engines and at a similar level to websites affiliated with academic institutions.

Limitations

There are several limitations to this study. This evaluation of a single AI natural language processing model took place at a single point in time. The ability of the model to produce consistent responses was not tested. Furthermore, since the model is based on machine learning and feedback, it will continue to evolve and improve by its very nature. As outlined in the discussion, AI may present serious ethical concerns due to its lack of transparency and potential for incorrect information.

Conclusion

The quality of information provided by the open AI chat system was generally high across all question types but had significant shortcomings in reliability due to the absence of source material citations. The DISCERN scores of the AI generated responses matched or exceeded previously published results of studies evaluating the quality of online information about rotator cuff repairs. The responses were U.S. 10th grade or higher reading level which is above the AMA and NIH recommendation of 6th grade reading level for patient materials. The AI software commonly referred the user to seek advice from orthopedic surgeons to improve their chances of a successful outcome.

Disclaimers:

Funding: No outside funding or grants were used for the data collection, data analysis, preparation of or editing of this manuscript. Conflicts of interest: Eoghan T. Hurley: Arthroscopy: Editorial or governing board. European Society for Surgery of the Shoulder and Elbow: Board or committee member. Journal of Shoulder and Elbow Surgery: Editorial or governing board. Oke Anakwenze: Exactech, Inc: Paid consultant. Smith & Nephew: Paid consultant. Peter B. MacDonald: American Shoulder and Elbow Surgeons: Board or committee member. Clinical Journal of sports medicine: Editorial or governing board. Christopher S. Klifto: Acumed, LLC: Paid consultant, Restore3d: Paid consultant and stock options. Smith & Nephew: Paid consultant.
The other authors, their immediate families, and any research foundation with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jseint.2023.09.009.

References

1. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature 2023;613:612. https://doi.org/10.1038/d41586-023-00191-1.
2. Ariyaratne S, Iyengar KP, Nischal N, Chitti Babu N, Botchu R. A comparison of ChatGPT-generated articles with human-written articles. Skeletal Radiol 2023;52:1755-8. https://doi.org/10.1007/s00256-023-04340-5.
3. Bi AS. What's important: the Next academic-ChatGPT AI? J Bone Joint Surg Am 2023. https://doi.org/10.2106/jbjs.23.00269.
4. Biswas S. ChatGPT and the Future of medical writing. Radiology 2023;307: e223312. https://doi.org/10.1148/radiol.223312.
5. Boileau P, Brassart N, Watkinson DJ, Carles M, Hatzidakis AM, Krishnan SG. Arthroscopic repair of full-thickness tears of the supraspinatus: does the tendon really heal? J Bone Joint Surg Am 2005;87:1229-40. https://doi.org/10.2106/jbjs.D.02035.
6. Celik H, Polat O, Ozcan C, Camur S, Kilinc BE, Uzun M. Assessment of the quality and reliability of the information on rotator cuff repair on YouTube. Orthop Traumatol Surg Res 2020;106:31-4. https://doi.org/10.1016/j.otsr.2019.10.004.
7. 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;53:105-11.
8. Chung SW, Oh JH, Gong HS, Kim JY, Kim SH. Factors affecting rotator cuff healing after arthroscopic repair: osteoporosis as one of the independent risk factors. Am J Sports Med 2011;39:2099-107. https://doi.org/10.1177/0363546511415659.
9. Dalton DM, Kelly EG, Molony DC. Availability of accessible and high-quality information on the Internet for patients regarding the diagnosis and management of rotator cuff tears. J Shoulder Elbow Surg 2015;24:e135-40. https://doi.org/10.1016/j.jse.2014.09.036.
10. Else H. Abstracts written by ChatGPT fool scientists. Nature 2023;613:423. https://doi.org/10.1038/d41586-023-00056-7.
11. Flesch R. A new readability yardstick. J Appl Psychol 1948;32:221-33.
12. Goldenberg BT, Schairer WW, Dekker TJ, Lacheta L, Millett PJ. Online resources for rotator cuff repair: what are patients reading? Arthrosc Sports Med Rehabil 2019;1:e85-92. https://doi.org/10.1016/j.asmr.2019.06.002.
13. Harryman DT 2nd, Mack LA, Wang KY, Jackins SE, Richardson ML, Matsen FA 3rd. Repairs of the rotator cuff. Correlation of functional results with integrity of the cuff. J Bone Joint Surg Am 1991;73:982-9.
14. Hill BG, Krogue JD, Jevsevar DS, Schilling PL. Deep learning and imaging for the Orthopaedic surgeon: How machines "read" Radiographs. J Bone Joint Surg Am 2022;104:1675-86. https://doi.org/10.2106/jbjs.21.01387.
15. Hodakowski AJ, McCormick JR, Damodar D, Cohn MR, Carey KD, Verma NN, et al. Rotator cuff repair: what questions are patients asking online and where are they getting their answers? Clin Shoulder Elb 2023;26:25-31. https://doi.org/10.5397/cise.2022.01235.
16. Kanthawala S, Vermeesch A, Given B, Huh J. Answers to health questions: internet search results versus online health Community responses. J Med Internet Res 2016;18:e95. https://doi.org/10.2196/jmir.5369.
17. Karpinski K, Plachel F, Gerhardt C, Saier T, Tauber M, Auffarth A, et al. Different expectations of patients and surgeons with regard to rotator cuff repair. J Shoulder Elbow Surg 2022;31:1096-105. https://doi.org/10.1016/j.jse.2021.12.043.
18. Keener JD, Patterson BM, Orvets N, Chamberlain AM. Degenerative rotator cuff tears: refining surgical indications based on natural history data. J Am Acad Orthop Surg 2019;27:156-65. https://doi.org/10.5435/jaaos-d-17-00480.
19. Khalil LS, Castle JP, Akioyamen NO, Corsi MP, Cominos ND, Dubé M, et al. What are patients asking and reading online? An analysis of online patient searches for rotator cuff repair. J Shoulder Elbow Surg 2023;32:2245-55. https://doi.org/10.1016/j.jse.2023.04.021.
20. Khosravi B, Rouzrokh P, Erickson BJ. Getting more out of large databases and EHRs with natural language processing and artificial intelligence: the future is here. J Bone Joint Surg Am 2022;104:51-5. https://doi.org/10.2106/jbjs.22.00567.
21. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLoS Digit Health 2023;2:e0000198. https://doi.org/10.1371/journal.pdig.0000198.
22. Lawson KA, Codella S, Ciccotti MG, Kane PW, Duncan IC, Cohen SB. Evaluation of internet information about rotator cuff repair. Am J Orthop (Belle Mead NJ) 2016;45:E136-42.
23. Matsuo Y, LeCun Y, Sahani M, Precup D, Silver D, Sugiyama M, et al. Deep learning, reinforcement learning, and world models. Neural Netw 2022;152: 267-75. https://doi.org/10.1016/j.neunet.2022.03.037.
24. Mello MM, Guha N. ChatGPT and physicians' Malpractice risk. JAMA Health Forum 2023;4:e231938. https://doi.org/10.1001/jamahealthforum.2023.1938.
25. Moosmayer S, Gärtner AV, Tariq R. The natural course of nonoperatively treated rotator cuff tears: an 8.8-year follow-up of tear anatomy and clinical outcome in 49 patients. J Shoulder Elbow Surg 2017;26:627-34. https://doi.org/10.1016/j.jse.2016.10.002.

26. Myers TG, Ramkumar PN, Ricciardi BF, Urish KL, Kipper J, Ketonis C. Artificial intelligence and orthopaedics: an introduction for clinicians. J Bone Joint Surg Am 2020;102:830-40. https://doi.org/10.2106/jbjs.19.01128.
27. Oosterhoff JHF, Gravesteijn BY, Karhade AV, Jaarsma RL, Kerkhoffs G, Ring D, et al. Feasibility of machine learning and Logistic Regression algorithms to predict outcome in orthopaedic trauma surgery. J Bone Joint Surg Am 2022;104:544-51. https://doi.org/10.2106/jbjs.21.00341.
28. Park JS, Park HJ, Kim SH, Oh JH. Prognostic factors affecting rotator cuff healing after arthroscopic repair in small to medium-sized tears. Am J Sports Med 2015;43:2386-92. https://doi.org/10.1177/0363546515594449.
29. Petersen SA, Murphy TP. The timing of rotator cuff repair for the restoration of function. J Shoulder Elbow Surg 2011;20:62-8. https://doi.org/10.1016/j.jse.2010.04.045.
30. Rothwell JD. In mixed company : Communicating in small groups and teams. Boston, MA. 8th ed ed. Boston, MA: Wadsworth, Cengage Learning; 2013 (ISBN No. 9781111346850; 1111346852).
31. Sanchez-Ramos L, Lin L, Romero R. Beware of references when using ChatGPT as a source of information to write scientific articles. Am J Obstet Gynecol 2023;229:356-7. https://doi.org/10.1016/j.ajog.2023.04.004.
32. Sánchez-Sotelo J, Jed Kuhn JE, Mallon WJ. Artificial intelligence and the creation of scientific papers. J Shoulder Elbow Surg 2023;32:685-6. https://doi.org/10.1016/j.jse.2023.02.002.
33. Schwarz I, Houck DA, Belk JW, Hop J, Bravman JT, McCarty E. The quality and content of internet-based information on Orthopaedic sports medicine requires Improvement: a systematic review. Arthrosc Sports Med Rehabil 2021;3:e1547-55. https://doi.org/10.1016/j.asmr.2021.05.007.
34. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor–Let the reader and viewer beware. JAMA 1997;277:1244-5.
35. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove. Nature 2023;613:620-1. https://doi.org/10.1038/d41586-023-00107-z.
36. Tan M, Lam PH, Le BT, Murrell GA. Trauma versus no trauma: an analysis of the effect of tear mechanism on tendon healing in 1300 consecutive patients after arthroscopic rotator cuff repair. J Shoulder Elbow Surg 2016;25:12-21. https://doi.org/10.1016/j.jse.2015.06.023.
37. Tashjian RZ. Epidemiology, natural history, and indications for treatment of rotator cuff tears. Clin Sports Med 2012;31:589-604. https://doi.org/10.1016/j.csm.2012.07.001.
38. Teunis T, Lubberts B, Reilly BT, Ring D. A systematic review and pooled analysis of the prevalence of rotator cuff disease with increasing age. J Shoulder Elbow Surg 2014;23:1913-21. https://doi.org/10.1016/j.jse.2014.08.001.
39. Weiss BD, Coyne C. Communicating with patients who cannot read. N Engl J Med 1997;337:272-4.