OPEN

PRS GLOBAL OPEN

ORIGINAL ARTICLE
Peripheral Nerve

Quality of ChatGPT Responses to Frequently Asked Questions in Carpal Tunnel Release Surgery

Troy B. Amen, MD, MBA*
Kaveh A. Torabian, MS†
Tejas Subramanian, BE*
Brian W. Yang, MD*
Adriana Liimakka, MBI‡
Duretti Fufa, MD*

Background: Although demonstrating remarkable promise in other fields, the impact of artificial intelligence (including ChatGPT in hand surgery and medical practice) remains largely undetermined. In this study, we asked ChatGPT frequently asked patient-focused questions surgeons may receive in clinic from patients who have carpel tunnel syndrome (CTS) and evaluated the quality of its output.

Methods: Using ChatGPT, we asked 10 frequently asked questions that hand surgeons may receive in the clinic before carpel tunnel release (CTR) surgery. Included questions were generated from the authors' own experiences regarding conservative and operative treatment of CTS.

Results: Responses from the following 10 questions were included: (1) What is CTS and what are its signs and symptoms? (2) What are the nonsurgical options for CTS? (3) Should I get surgery for CTS? (4) What is a CTR and how is it preformed? (5) What are the differences between open and endoscopic CTR? (6) What are the risks associated with CTR and how frequently do they occur? (7) Does CTR cure CTS? (8) How much improvement in my symptoms can I expect after CTR? (9) How long is the recovery after CTR? (10) Can CTS recur after surgery?

Conclusions: Overall, the chatbot provided accurate and comprehensive information in response to most common and nuanced questions regarding CTS and CTR surgery, all in a way that would be easily understood by many patients. Importantly, the chatbot did not provide patient-specific advice and consistently advocated for consultation with a healthcare provider. *(Plast Reconstr Surg Glob Open 2024; 12:e5822; doi: 10.1097/GOX.0000000000005822; Published online 16 May 2024.)*

INTRODUCTION

ChatGPT has brought the power of artificial intelligence (AI) to every household. Although demonstrating remarkable promise in other fields, the impact of these technologies in medical practice and education is undetermined. Patients with orthopedic concerns frequently have questions that require nuanced answers. To that end, patient education is one avenue that may see remarkable change, as patients equipped with AI technology can produce answers to their very specific questions. Rather than relying on difficult to navigate, static websites such as Wikipedia or WebMD,[1,2] patients can now input seemingly any question and receive very focused responses. This may serve as an excellent resource for patients learning more about their health and options. Given this broad access to AI technology, it therefore becomes important to evaluate the quality of these responses. Previous literature within orthopedic surgery has been hesitantly optimistic regarding the accuracy and quality of responses given by ChatGPT.[3–5] However, to date, a few studies have evaluated ChatGPT's responses within the context of hand surgery.[5] In this study, we asked ChatGPT frequently asked patient-focused questions surgeons may receive in clinic from patients who are offered carpal tunnel release (CTR) surgery and evaluated the quality of its output.

METHODS

Using ChatGPT (version 3.5), we asked 10 frequently asked questions that hand surgeons may receive in the clinic before CTR surgery. Included questions were generated from the authors' own experiences. The

*From *Department of Orthopaedic Surgery, Hospital for Special Surgery, New York, N.Y.; †Rush Medical College of Rush University, Chicago, Ill.; and ‡Harvard Medical School, Boston, Mass.*

Received for publication December 18, 2023; accepted March 22, 2024.

DOI: 10.1097/GOX.0000000000005822

Disclosure statements are at the end of this article, following the correspondence information.

Related Digital Media are available in the full-text version of the article on www.PRSGlobalOpen.com.

conversation with the chatbot began with the following statement, "Please answer the following question as a surgeon speaking to their patient." Questions were inputted sequentially into the ChatGPT user interface. Responses were then reviewed by the authors for accuracy and quality. Questions included were:

1. What is carpal tunnel syndrome and what are its signs and symptoms?
2. What are the nonsurgical options for carpal tunnel syndrome?
3. Should I get surgery for carpal tunnel syndrome?
4. What is a carpal tunnel release and how is it preformed?
5. What are the differences between open and endoscopic CTR?
6. What are the risks associated with carpal tunnel release and how frequently do they occur?
7. Does carpal tunnel release cure carpal tunnel syndrome?
8. How much improvement in my symptoms can I expect after carpal tunnel release?
9. How long is the recovery after carpal tunnel release?
10. Can carpal tunnel syndrome recur after surgery?

RESULTS

All questions along with their corresponding responses generated by ChatGPT are provided in detail in Supplemental Digital Content 1. (**See appendix, Supplemental Digital Content 1,** which displays the questions and answers generated. **http://links.lww.com/PRSGO/D213**.)

DISCUSSION

Among the wide range of fields with possible AI application, medicine stands out as one with tremendous potential along with equally substantial challenges. The advent of the digital age has dramatically increased patient access to medical information, ushering in a new era of patient education.[6,7] However, the widespread use of the internet as a source of medical information has also brought unique challenges, many of which are still problematic today. For example, Hutchinson et al found that online education materials most frequently used by patients were written at reading levels far beyond that recommended by the National Institutes of Health.[8] In fact, several prior studies have reported poor readability of current online resources, leading to poor comprehension, suboptimal treatment adherence, and increased healthcare disparities.[9,10] As AI chatbots, such as ChatGPT, continue to experience remarkable growth in popularity, they have the potential to become the basis of health education. Unlike static websites, responses from ChatGPT are personalized and dynamic. The ability of the system to keep track of the context of an ongoing conversation conveys a more useful and natural feeling. Furthermore, a program that is constantly being updated with new information has the potential to provide up-to-date information while tailoring its response to a given query. Although there is ample reason for optimism, it is important to learn from the past and identify the unique limitations of AI in patient education before its widespread adoption. Thus, we sought to evaluate ChatGPT's effectiveness as a patient education tool in the context of hand surgery, specifically CTS and CTR.

Takeaways

Question: In this study, we asked ChatGPT frequently asked patient-focused questions surgeons may receive in clinic from patients who have carpel tunnel syndrome (CTS) and evaluated the quality of its output.

Findings: Overall, the chatbot provided accurate and comprehensive information in response to most common and nuanced questions regarding CTS and carpal tunnel release (CTR).

Meaning: ChatGPT has the potential to serve as an effective patient education tool in the context of hand surgery and specifically CTS and CTR.

Overall, the chatbot provided accurate and comprehensive information in response to most common and nuanced questions regarding CTS and CTR surgery, all in a way that would be easily understood by many patients. Importantly, the chatbot did not provide patient-specific advice, but rather highlighted the fact that CTS can vary substantially depending on a number of patient-specific factors and consistently advocated for consultation with a healthcare provider in every response. This recommendation to seek the advice of a physician is crucial and encouragingly, has also been reported in other studies investigating ChatGPT responses in the context of patient education.[3,5]

When asked to differentiate between open carpal tunnel release (OCTR) and endoscopic carpal tunnel release (ECTR), a debated topic in the literature, the chatbot provided an acceptable explanation. Specifically, it answered that ECTR "usually has a shorter recovery time" and "may involve less postoperative pain and swelling compared with OCTR." Indeed, recent meta-analyses agree that ECTR is associated with accelerated recovery and quicker return to work.[11–14] Similarly, randomized controlled trials have noted ECTR to result in less postoperative pain and swelling at the incision site.[15–17] Ultimately, however, it concluded the response by claiming that both procedures are effective and the choice between OCTR and ECTR depends on various factors, including the surgeon's expertise. Given the long history of conflicting studies comparing ECTR and OCTR, the chatbot's ability to selectively present only the strongest supported claims is encouraging. Similar to the increasing patient interest and advertisement of direct anterior total hip arthroplasty in the joint reconstruction realm, the importance of this nuanced output is further underscored in the context of increased emphasis on patient-reported outcomes.[18,19] Notably, its conclusion was that both procedures were safe with a similarly low complication profile despite several studies reporting on higher rates of transient nerve injury after ECTR.

Although the majority of responses were well-informed and accurate, a notable limitation we observed

in ChatGPT's responses was its tendency to include broad statements or recommendations that lack conclusive evidence. For example, when discussing postoperative care, ChatGPT advocated for wearing a wrist splint for a few weeks postoperatively to immobilize the wrist, even implying that nonadherence could impact long-term success. However, strong evidence suggests that there is no benefit to routine postoperative immobilization after CTR,[20] with some studies suggesting that it may adversely affect rehabilitation.[21] Although some patient-specific factors may necessitate the use of a wrist split for postoperative immobilization, this was not clarified by the chatbot. Additionally, when asked about nonsurgical options for CTS, ChatGPT recommended nonsteroidal antiinflammatory drugs (NSAIDs). However, moderate evidence suggests that there is no benefit of NSAIDs compared with placebo,[20,22] and excessive use of NSAIDs may potentiate the risk of gastrointestinal adverse effects, especially if used with oral glucocorticoids. Again, NSAID use may be warranted in certain circumstances, but this was not clarified. A major criticism of ChatGPT is its tendency to include blatant falsehoods in its responses, often referred to as a "hallucination."[23] Hallucinations can be particularly dangerous in medical scenarios, especially when subtle and stated in a convincing manner.[23] However, the discrepancies noted in the present response should not be considered hallucinations, as there may be circumstances where NSAID use and postoperative immobilization are appropriate, and several prominent patient education websites also recommend them.[24,25] Nevertheless, there is no way to know what information was utilized in generating these outputs, or whether that information was high quality. This is known as the "black box" concern, which refers to the opacity and lack of transparency in AI models, making it challenging to understand how they arrive at their decisions or predictions.[26] This can also have dangerous implications in medical scenarios and is an important limitation in the use of chatbots for patient education. Furthermore, a similar study by Crook et al[5] found that ChatGPT responses were written at a college reading level, far surpassing the sixth grade reading level recommended by the National Institutes of Health. Converting responses to patient-selected reading levels and including supplemental illustrations would significantly improve AI's use in patient education, although these are not current features of ChatGPT. The purpose of this study was to evaluate ChatGPT as tool to inform patients on the diagnosis and treatment of carpal tunnel syndrome. Taken as a whole, it generated responses that were highly accurate and comprehensive without any false or potentially misleading information provided.

Although it is essential to acknowledge the current limitations of AI and ChatGPT in healthcare, it is equally crucial to underscore the remarkable progress that has been made in recent years. The journey toward integrating AI as a valuable source of patient education is undoubtedly underway, and its potential is both vast and encouraging. The applications of AI in medicine have expanded significantly, demonstrating its capacity to revolutionize the healthcare landscape. Notable examples include its pivotal role in the analysis of medical images,[27] its proficiency in detecting drug interactions,[28] and its ability to identify high-risk patients to enable early intervention.[29] However, the current limitations of ChatGPT, including the possibility for hallucinations, its lack of citations, and outstanding legal considerations,[30] may preclude its widespread adoption into clinical practice as a clinician-recommended education tool. Although some of these limitations are unanimous among generative AI chatbots today, given the speed at which the technology is evolving, it is highly plausible that chatbots may achieve the consistency and reliability necessary to be deemed clinician-recommended education tools in the near future.[31] As we move forward, it is vital to recognize that the pivotal role that today's physicians and healthcare professionals hold in shaping the future of healthcare, as they will be instrumental in guiding AI's integration into healthcare, ensuring ethical and responsible implementation, and harnessing the potential of AI as a source of patient education. The collaboration between human expertise and AI's capabilities is poised to enhance patient outcomes, reduce healthcare disparities, and create a healthcare system that is more efficient, accessible, and patient-centered.

Troy B. Amen, MD, MBA
Hospital for Special Surgery
535 E 70th St
NY, NY 10021
E-mail: ament@hss.edu

DISCLOSURE

The authors have no financial interest to declare in relation to the content of this article.

REFERENCES

1. Smith DA. Situating wikipedia as a health information resource in various contexts: a scoping review. *PLoS One.* 2020;15:e0228786.
2. Tulbert BH, Snyder CW, Brodell RT. Readability of patient-oriented online dermatology resources. *J Clin Aesthet Dermatol.* 2011;4:27–33.
3. Subramanian T, Shahi P, Araghi K, et al. Using artificial intelligence to answer common patient-focused questions in minimally invasive spine surgery. *J Bone Joint Surg Am.* 2023;105:1649–1653.
4. Mika AP, Martin JR, Engstrom SM, et al. Assessing ChatGPT responses to common patient questions regarding total hip arthroplasty. *J Bone Joint Surg.* 2023;105:1519–1526.
5. Crook BS, Park CN, Hurley ET, et al. Evaluation of online artificial intelligence-generated information on common hand procedures. *J Hand Surg Am.* 2023;48:1122–1127.
6. Diaz JA, Griffith RA, Ng JJ, et al. Patients' use of the internet for medical information. *J Gen Intern Med.* 2002;17:180–185.
7. Trotter MI, Morgan DW. Patients' use of the internet for health related matters: a study of internet usage in 2000 and 2006. *Health Informatics J.* 2008;14:175–181.
8. Hutchinson N, Baird GL, Garg M. Examining the reading level of internet medical information for common internal medicine diagnoses. *Am J Med.* 2016;129:637–639.
9. Rodriguez F, Ngo S, Baird G, et al. Readability of online patient educational materials for coronary artery calcium scans and implications for health disparities. *J Am Heart Assoc.* 2020;9:e017372.
10. Johnson PT, Chen JK, Eng J, et al. A comparison of world wide web resources for identifying medical information. *Acad Radiol.* 2008;15:1165–1172.

11. Sayegh ET, Strauch RJ. Open versus endoscopic carpal tunnel release: a meta-analysis of randomized controlled trials. *Clin Orthop Relat Res.* 2015;473:1120–1132.
12. Chen L, Duan X, Huang X, et al. Effectiveness and safety of endoscopic versus open carpal tunnel decompression. *Arch Orthop Trauma Surg.* 2014;134:585–593.
13. Vasiliadis HS, Nikolakopoulou A, Shrier I, et al. Endoscopic and open release similarly safe for the treatment of carpal tunnel syndrome. a systematic review and meta-analysis. *PLoS One.* 2015;10:e0143683.
14. Li Y, Luo W, Wu G, et al. Open versus endoscopic carpal tunnel release: a systematic review and meta-analysis of randomized controlled trials. *BMC Musculoskelet Disord.* 2020;21:272.
15. Atroshi I, Larsson GU, Ornstein E, et al. Outcomes of endoscopic surgery compared with open surgery for carpal tunnel syndrome among employed patients: randomised controlled trial. *BMJ.* 2006;332:1473.
16. Kang HJ, Koh IH, Lee TJ, et al. Endoscopic carpal tunnel release is preferred over mini-open despite similar outcome: a randomized trial. *Clin Orthop Relat Res.* 2013;471:1548–1554.
17. Brown RA, Gelberman RH, Seiler JG, et al. Carpal tunnel release. A prospective, randomized assessment of open and endoscopic methods. *J Bone Joint Surg Am.* 1993;75:1265–1275.
18. Moore ML, Brinkman JC, Pollock JR, et al. Patients are most interested in which hip arthroplasty approach? A 15-year Google Trends analysis. *Arthroplast Today.* 2022;17:192–197.
19. Shofoluwe AI, Naveen NB, Inabathula A, et al. Internet promotion of direct anterior approach total hip arthroplasty by members of the American Association of Hip and Knee Surgeons. *J Arthroplasty.* 2018;33:167–170.e1.
20. Graham B, Peljovich AE, Afra R, et al. The American Academy of Orthopaedic Surgeons evidence-based clinical practice guideline on. *J Bone Joint Surg Am.* 2016;98:1750–1754.
21. Huemer GM, Koller M, Pachinger T, et al. Postoperative splinting after open carpal tunnel release does not improve functional and neurological outcome. *Muscle Nerve.* 2007;36:528–531.
22. Chang MH, Chiang HT, Lee SSJ, et al. Oral drug of choice in carpal tunnel syndrome. *Neurology.* 1998;51:390–393.
23. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI Chatbot for medicine. *N Engl J Med.* 2023;388:1233–1239.
24. Mayo Clinic. Carpal tunnel syndrome—diagnosis and treatment. Available at https://www.mayoclinic.org/diseases-conditions/carpal-tunnel-syndrome/diagnosis-treatment/drc-20355608#. Updated February 6, 2024. Accessed October 28, 2023.
25. Johns Hopkins Medicine. Carpal tunnel release. Available at https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/carpal-tunnel-release#. Accessed October 28, 2023.
26. Price WN, Gerke S, Cohen IG. How much can potential jurors tell us about liability for medical artificial intelligence? *J Nucl Med.* 2021;62:15–16.
27. Ker J, Wang L, Rao J, et al. Deep learning applications in medical image analysis. *IEEE Access.* 2018;6:9375–9389.
28. Han K, Cao P, Wang Y, et al. A review of approaches for predicting drug–drug interactions based on machine learning. *Front Pharmacol.* 2022;12:814858.
29. Beaulieu-Jones BK, Yuan W, Brat GA, et al. Machine learning for patient risk stratification: standing on, or looking over, the shoulders of clinicians? *NPJ Digit Med.* 2021;4:62.
30. Altamimi I, Altamimi A, Alhumimidi AS, et al. Artificial intelligence (AI) chatbots in medicine: a supplement, not a substitute. *Cureus.* 2023;15:e40922.
31. Singhal K, Tu T, Gottweis J, et al. Towards expert-level medical question answering with large language models. Arxiv.org. Available at https://arxiv.org/abs/2305.09617. Published May 16, 2023. Accessed April 19, 2024.