

Large Language Models in Orthopaedic Publications

The Good, the Bad and the Ugly

The past few years has seen an explosion in the availability and usage of artificial intelligence (AI) in all fields. Medicine, and more specifically orthopaedics, has been no exception. AI using machine learning has the potential to help researchers analyze huge datasets to develop predictive models for injury and improve treatment plans. The use of large language models (LLMs) like Chat GPT-4 in scientific production has also opened new avenues for efficiency in content generation and summarization.³ However, their integration into the writing and peer review processes in medical journals raises concerns surrounding ethics, legality, copyright, and the integrity of peer review.^{3,5,11} This editorial will examine the benefits, risks and dangers of the “wild-west” of LLMs in orthopaedic publications (Figure 1).³



Figure 1. Image created by Dall-e based on content of this editorial.

This editorial has been copublished in *The American Journal of Sports Medicine*.

Keywords: orthopaedic publishing; artificial intelligence; large language models

The Orthopaedic Journal of Sports Medicine, 12(8), 23259671241265705
DOI: 10.1177/23259671241265705
© The Author(s) 2024

THE BENEFITS OF LLMS

Chat GPT-4 and other LLMs can assist with the production of a manuscript.³ Generative AI can be utilized to quickly identify background articles and help create a bibliography. In the writing process, LLMs can assist in formatting an article,¹³ correcting grammatical mistakes, improving style, and making manuscripts more readable.¹² Most orthopaedic journals are published in the English language. For non-native speakers, LLMs can quickly translate articles from their native language into English.

Editors and publishers can also use AI for their benefit.¹⁵ LLMs can ensure compliance with submission guidelines, scan references, and detect duplicate submissions or potential plagiarism.¹⁴ AI can also assist in identifying potential reviewers and provide timely copyediting. LLMs can help expedite the manuscript processing system and allow more timely publication of orthopaedic research.⁴

THE RISKS OF LLMS

LLMs can also be used to generate ideas and texts. A recent study showed an increase in the use of LLMs in the writing or articles submitted to an orthopaedic journal.^{2,9} The use of chatbots to produce content comes with risks. It is well documented that generative AI can produce hallucinations, creating facts that have not been proven.^{1,7} Additionally, since LLMs synthesize responses based on existing data patterns, there is a risk that their outputs could closely resemble existing literature, leading to accusations of plagiarism. Additionally, the risk of data privacy violations is significant if patient information or sensitive clinical data is inadvertently disclosed in published articles. The World Association of Medical Editors (WAME) has stated that chatbots cannot be authors as they don't have the ability to give “final approval” of a manuscript or understand conflict of interest statements.¹⁸ WAME has additionally recommended that authors must disclose the use of AI and that they are responsible for all material generated by chatbots.^{13,17}

For the *American Journal of Sports Medicine (AJSM)* and the *Orthopaedic Journal of Sports Medicine (OJSM)*,

authors are required to declare any use of AI and specify the use. AI should only be used to correct language, not to generate text. Fabrication of research by any means is a major violation of ethics.

Some investigators have explored whether LLMs can be used in the peer review system.⁶ Chat GPT-4—generated reviews have been shown to be somewhat useful in analyzing research papers.⁴ However, AI will not replace the expertise of human peer review given that AI is based on existing knowledge while research involves the exploration of innovative ideas.¹⁵ We have identified the use of LLMs in some reviews submitted to *AJSM*. LLM-generated reviews could unintentionally perpetuate biases or overlook methodological flaws due to their reliance on patterns rather than a deep understanding of the research context.⁵ Reviewers relying on these models to evaluate submissions might produce assessments that lack the nuanced, critical judgment required to validate research findings. A further concern is that using an Open AI system to generate peer-review breaches the confidentiality of the submitted papers.¹⁶ Authors retain the copyright of their work and uploading manuscripts to a chatbot violates that copyright. It is the policy of *AJSM/OJSM* that reviewers should not use AI to write a review.

THE DANGERS OF LLMs

Medical publishing requires stringent ethical standards to ensure the quality and reliability of research. LLMs could inadvertently compromise these standards by propagating biases embedded in their training data, potentially affecting how research is interpreted and applied in clinical practice.¹¹ Additionally, the risk of data privacy violations is significant if patient information or sensitive clinical data is inadvertently disclosed in published articles.⁸ Medical publishers must navigate strict data protection. If LLMs introduce sensitive information into manuscripts, journals and authors could be exposed to liability.¹⁶

Even more concerning is the ability of artificial intelligence to be used to create intentionally fraudulent manuscripts. Chat GPT-4 has been shown to be capable of creating a highly convincing fraudulent manuscript in approximately 1 hour.⁸ Paper mills are businesses that sell authorship of fake or poor-quality manuscripts. It has been estimated that 3% of all medical publications in 2022 resembled paper mill productions.¹⁷ The ability of generative AI to quickly produce fraudulent research papers will only make this problem worse. Authors, reviewers, editors, and readers must be vigilant in identifying fraudulent manuscripts.¹⁰

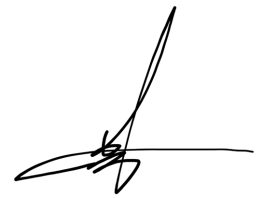
CONCLUSION

LLMs offer transformative potential for medical publishing, enhancing productivity and access to information. However, their use in manuscript drafting and peer review processes brings ethical, legal, copyright, and review integrity issues to the forefront. Editors and publishers of orthopaedic journals must develop clear guidelines on the use of

LLMs to ensure transparency in authorship, originality of content, and fairness in peer review. Ultimately, fostering collaboration between researchers, publishers, and AI developers will help address these challenges while maximizing the benefits that LLMs can bring to scientific advancement. Until then, the use of artificial intelligence in orthopaedic publishing is in a “wild-west” phase. For the time being, it is critical that authors, reviewers, and editors are aware of the good, the bad, and the ugly of these LLMs.



Daniel C. Wascher, MD
Albuquerque, New Mexico, USA



Matthieu Ollivier, MD
Marseille, France

REFERENCES

- Alkaiissi H, McFarlane SI. Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*. 2023;15(2):e35179. doi:10.7759/cureus.35179.
- Bisi T, Risser A, Clavert P, Migaud H, Dartus J. What is the rate of text generated by artificial intelligence over a year of publication in Orthopedics & Traumatology: Surgery & Research? Analysis of 425 articles before versus after the launch of ChatGPT in November 2022. *Orthop Traumatol Surg Res*. 2023;109(8):103694. doi:10.1016/j.otsr.2023.103694
- Dahmen J, Kayaalp ME, Ollivier M, et al. Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(4):1187-1189. doi:10.1007/s00167-023-07355-6
- Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res*. 2023;13(4):1148-1154.
- Kayaalp ME, Ollivier M, Winkler PW, et al. Embrace responsible ChatGPT usage to overcome language barriers in academic writing. *Knee Surg Sports Traumatol Arthrosc*. 2024;32(1):5-9. doi:10.1002/ksa.12014
- Krishnan V. Artificial intelligence in scientific peer review. *J World Fed Orthod*. 2024;13(2):55-56. doi:10.1016/j.ejwf.2024.03.004
- Leffer L. AI chatbots will never stop hallucinating. *Scientific American*. April 5, 2024. Accessed June 17, 2024. <https://www.scientificamerican.com/article/chatbot-hallucinations-inevitable/>
- Májovský M, Černý M, Kasal M, Komarc M, Netuka D. Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *J Med Internet Res*. 2023;25:e46924. doi:10.2196/46924
- Maroteau G, An JS, Murgier J, Hulet C, Ollivier M, Ferreira A. Evaluation of the impact of large language learning models on articles submitted to Orthopaedics & Traumatology: Surgery & Research (OTSR): A significant increase in the use of artificial intelligence in 2023. *Orthop Traumatol Surg Res*. 2023;109(8):103720. doi:10.1016/j.otsr.2023.103720.
- Nahai F. General Data Protection Regulation (GDPR) and data breaches: What you should know. *Aesthet Surg J*. 2019;39(2):238-240. doi:10.1093/asj/sjy296

11. Ollivier M, Pareek A, Dahmen J, et al. A deeper dive into ChatGPT: history, use and future perspectives for orthopaedic research. *Knee Surg Sports Traumatol Arthrosc.* 2023;31(4):1190-1192. doi:10.1007/s00167-023-07372-5
12. Preparing your manuscript. SAGE Publications Inc. <https://us.sagepub.com/en-us/nam/preparing-your-manuscript>. Accessed June 17, 2024
13. Reider B. Author, author!. *Am J Sports Med.* 2002;30(5):635. doi:10.1177/03635465020300050101.
14. Reider B. Duplicates, derivatives, and salamis. *Am J Sports Med.* 2004;32(3):579. doi:10.1177/0363546504264932
15. Reider B. Under surveillance. *Am J Sports Med.* 2010;38(12):2391-2393. doi:10.1177/0363546510390477
16. Reider B. Write; Copyright. *Am J Sports Med.* 2016;44(2):295-296. doi:10.1177/0363546516628526
17. Thorp HH. ChatGPT is fun, but not an author. *Science.* 2023; 379(6630):313. doi:10.1126/science.adg7879.
18. Zielinski C, Winker MA, Aggarwal R, et al. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. *Colomb Med (Calij).* 2023;54(3):e1015868. doi:10.25100/cm.v54i3.5868