

Comparing Human and AI-Generated Editorials in Orthopaedic Oncology: A Blinded Evaluation of Quality and Authorship Perception

Jessa Fogel ¹, Ahlam Ashkar, Philip Khoury, Thomas Schubert, Vincent Ng, Anthony Bozzo, Alexander Christ, Peter Choong, Marcos Galli Serra, Juan Pretell, Raja Bhaskara Rajasekaran, Venkatesan Kumar, Adam Levin, Kenneth Gundle, Christopher Johnson, Ricardo Becker, Alexander Lazarides, Ahmed El Ghoneimy, Panagiotis Tsagkozis, Felasfa Wodajo, Roberto Vélez, Dimosthenis Andreou, Andreas Leithner, Lisa Ercolano, Michelle Ghert

Affiliations + expand

PMID: 42535827 DOI: [10.5435/JAAOS-D-25-01565](https://doi.org/10.5435/JAAOS-D-25-01565) 

Abstract

Background: As the use of artificial intelligence (AI) and large language models (LLMs) is increasingly adopted into scientific writing, it is important to understand AI's ability to produce clear and accurate content that is on par with human-authored content in the field of orthopaedics, including orthopaedic oncology. The aim of this study was to compare a series of editorials written by orthopaedic oncologists with those written by a single LLM (ChatGPT 4.0) using a variety of quality metrics.

Methods: Volunteer orthopaedic oncologists submitted a 3- to 4-paragraph persuasive editorial on a topic of their choice in the field of musculoskeletal oncology. ChatGPT 4.0 was then prompted to write a corresponding editorial for each topic. Each editorial was evaluated by two blinded peer reviewers and graded using a 25-point scale on the following quality metrics: content, clarity, grammar, persuasiveness, and creativity. The evaluators were also asked to indicate whether they believed the editorials were written by humans or by AI.

Results: A total of 20 editorials were submitted by human authors and matched with 20 prompted AI editorials. No notable difference in average total quality score for human versus AI submissions was observed. AI-generated articles scored markedly higher in grammar, but there were no notable differences in any other quality metric. Reviewers correctly identified author type 59% of the time.

Discussion: LLMs such as ChatGPT can generate editorial content in orthopaedic oncology that matches human-written quality, suggesting a potential supportive role for AI in scientific communication, with implications for authorship standards, editorial practices, and peer review.

Copyright © 2026 by the American Academy of Orthopaedic Surgeons.

[PubMed Disclaimer](#)