Qeios

Peer Review

Review of: "The Influence of an Artificial Intelligence Large Language Model (ChatGPT) on Orthopaedic Scientific Publishing: A Bibliometric Analysis"

Alexandru Tăbușcă[1]

1. School of Computer Science for Business Management, Romanian-American University, Romania

My review is related to the version **Preprint 2** of the paper “The Influence of an Artificial Intelligence Large Language Model (ChatGPT) on Orthopaedic Scientific Publishing: A Bibliometric Analysis”.

---

1. Summary of the Revised Manuscript

Version 2 retains the core objective of the original submission: to examine whether the release of ChatGPT is temporally associated with measurable changes in orthopaedic scientific publishing output, authorship patterns, and abstract-level linguistic features. The study continues to analyse a large PubMed-derived corpus (≈19,000 articles) across ten high-impact orthopaedic journals, using bibliometric and natural language processing techniques.

The revised version improves clarity in several key areas, particularly in methodological justification, interpretative caution, and ethical contextualisation. The authors maintain a non-causal framing and reinforce the role of confounding structural factors (post-pandemic publishing recovery, collaborative trends, journal expansion). The paper remains well aligned with Qeios' mission as an open, descriptive, policy-relevant contribution rather than a prescriptive or causal claim.

doi.org/10.32388/DIVAIB

2. Ethics, Transparency, and Research Integrity (Updated Assessment)

Addressed in Version 2:

- The ethical discussion has been expanded and clarified, particularly regarding the distinction between AI-assisted writing and scientific misconduct.
- The authors now more explicitly acknowledge the limits of AI-detection and bibliometric inference, reducing the risk of misinterpretation or stigmatization.
- The revised text better situates AI use within evolving editorial norms rather than framing it as an anomalous disruption.

Remaining recommendations:

- While ethical framing is improved, a short, explicit subsection on AI disclosure policies in biomedical journals (e.g., alignment with ICMJE or COPE guidance) would further strengthen the contribution.
- Discussion of privacy and consent remains indirect, though this is understandable given the metadata-based design.

Overall, ethical positioning is now more mature and balanced than in Version 1.

3. Methodological Quality and Rigor (Updated Assessment)

Addressed in Version 2:

- The Methods section shows improved structural clarity, particularly in how linguistic variables are defined and operationalised.
- Limitations related to **abstract-only analysis** are now more explicitly acknowledged and better contextualised.
- The authors strengthen the explanation of statistical testing choices and reinforce the descriptive nature of findings.

Partially addressed:

- Concerns about Type-Token Ratio sensitivity to text length are acknowledged, but no alternative lexical diversity metrics are introduced. While acceptable for a revision, this remains an opportunity

for future work.

Still recommended:

- Stratification by article type (original research vs reviews vs editorials) is still absent.
- Full-text analysis or validation against deeper linguistic metrics remains a future direction rather than an implemented enhancement.

Methodologically, Version 2 demonstrates greater reflexivity and transparency, even if the analytical scope remains unchanged.

4. Novelty and Contribution (Updated Assessment)

The revised manuscript does not attempt to artificially inflate novelty, which is a strength. Instead, it clarifies its position as:

- An early, discipline-specific empirical snapshot of AI-era publishing trends.
- A baseline against which future post-2025 analyses can be compared.

By sharpening this framing, Version 2 improves conceptual honesty and strengthens the paper's value as a reference point rather than a definitive account.

5. Impact and Relevance (Updated Assessment)

Improved in Version 2:

- The discussion now more clearly identifies editors, reviewers, and policy-makers as primary stakeholders.
- The revised conclusions avoid speculative overreach and emphasize that quantity and stylistic change do not equate to quality improvement or degradation.

Remaining gap:

- Practical recommendations for journals (e.g., monitoring trends without punitive enforcement) could be more explicitly summarized, possibly as a short “Editorial Implications” subsection.

The paper's impact remains primarily agenda-setting, which is appropriate for Qeios.

6. Clarity, Organization, and Presentation (Updated Assessment)

Addressed in Version 2:

- Redundancies in the Discussion have been reduced.
- The revised manuscript exhibits improved narrative coherence, particularly when transitioning from results to interpretation.
- Terminology related to AI use is more precise, reducing ambiguity.

Minor remaining issues include opportunities for:

- A concise summary table contrasting pre- and post-ChatGPT findings.
- Visual emphasis on effect sizes to avoid misinterpretation of statistical significance.

7. Interpretation and Balance (Updated Assessment)

In my opinion, this is one of the most improved aspects of Version 2.

Clearly addressed:

- The authors now consistently reinforce that observed changes are modest in magnitude.
- Causal language is avoided more rigorously than in Version 1.
- The role of systemic publishing dynamics (team science, journal expansion, submission backlogs) is more explicitly acknowledged.

This strengthens the manuscript's credibility and makes it particularly suitable for open peer review contexts.

Final Assessment of Version 2

Version 2 represents a clear and substantive improvement over the initial submission, particularly in ethical framing, interpretative balance, and clarity of scope. While the core analytical design remains unchanged, the authors have demonstrably engaged with prior critical observations and refined the manuscript accordingly.

The paper now functions more effectively as:

- A descriptive bibliometric benchmark,

- A conversation starter for editors and research policy-makers,
- A measured contribution to discussions on AI-assisted scientific writing.

My Recommendation:

- Publish Version 2

- Encourage future extensions incorporating article-type stratification, alternative linguistic metrics, and longitudinal follow-up beyond 2025.

Declarations

Potential competing interests: No potential competing interests to declare.