

Research Article

Temporal Trends in Orthopaedic Scientific Publishing Before and After the Release of ChatGPT: A Bibliometric Analysis

Preprinted: 24 October 2025
Peer-approved: 26 February 2026

© The Author(s) 2026. This is an Open Access article under the CC BY 4.0 license.

Qeios, Vol. 8 (2026)
ISSN: 2632-3834

Gabriel Ferraz Ferreira^{1,2,3}, Thomas Lorchan Lewis⁴, Miguel Viana Pereira Filho^{1,2,3}, Fabricio Machado⁵, Gustavo Gonçalves Arliani¹, Noel Oizerovici Foni⁶, Roberto Zambeli⁷, Caio Nery⁸, Rodrigo Pagnano³

1. Department of Orthopaedics and Traumatology, Prevent Senior, Brazil; 2. Instituto Vita, São Paulo, Brazil; 3. Department of Orthopedics, Rheumatology and Traumatology, Faculty of Medical Sciences, Universidade Estadual de Campinas (Unicamp), Brazil; 4. Orthopaedic and Arthritis Specialist Centre, Australia; 5. Health Lab, Prevent Senior, Brazil; 6. Hospital Israelita Albert Einstein, São Paulo, Brazil; 7. Faculty of Medical Sciences of Minas Gerais, Brazil; 8. Department of Orthopaedics and Traumatology, UNIFESP - Escola Paulista de Medicina, Brazil

Purpose: This study aimed to assess bibliometric trends in orthopaedic research before and after the public release of ChatGPT.

Methods: A bibliometric analysis was conducted using PubMed data from January 2021 to December 2025, encompassing articles from ten high-impact orthopaedic journals. A seven-month washout period (December 2022 – June 2023) was applied to account for the typical lag between manuscript preparation and publication. Trends in daily publication frequency, number of co-authors per article, sentence length, lexical diversity, and sentiment were compared between the pre-ChatGPT period (January 2021 – June 2023) and the post-adoption period (July 2023 – December 2025).

Results: A total of 21,524 articles were analysed (10,403 before; 11,121 after). The mean number of publications per day increased from 11.42 ± 7.32 to 12.15 ± 8.26 ($p = 0.044$). After adjusting for monthly seasonality, the difference remained significant (adjusted increase: $+1.15$ publications/day; $p = 0.002$). The mean number of authors per article rose from 5.94 ± 3.95 to 6.28 ± 4.39 ($p < 0.001$). The average sentence length decreased from 14.81 ± 5.08 to 14.52 ± 5.06 words per sentence ($p < 0.001$), while lexical diversity increased (Type-Token Ratio: 0.51 ± 0.07 to 0.52 ± 0.07 ; $p < 0.001$). Mean sentiment scores rose from 3.28 ± 3.27 to 3.64 ± 3.47 ($p < 0.001$).

Conclusion: Following the public release of ChatGPT, orthopaedic publications have exhibited a measurable rise in daily output, modest increases in co-authorship, and subtle changes in linguistic style. These temporal associations, while not evidence of causality, correspond with the widespread adoption of AI-assisted writing tools and warrant ongoing evaluation.

Corresponding author: Gabriel Ferraz Ferreira, gabriel.ferraz38@yahoo.com.br

Introduction

Artificial intelligence (AI) has rapidly evolved, influencing multiple aspects of scientific research, including data processing, writing, and publication trends.^{[1][2][3]} ChatGPT, a state-of-the-art large language model (LLM) developed by OpenAI (California, USA) and released in November 2022, emerged as a transformative tool, quickly gaining widespread popularity for its ability to generate human-like text.^[4]

Leveraging natural language processing (NLP) and deep learning (DL), ChatGPT processes vast amounts of data to generate coherent and contextually relevant responses.^[5] Researchers continue to explore its potential in various medical fields, leading to an increasing number of publications across different specialities.^{[6][7]}

From early-career researchers to senior academics, there is increasing interest in leveraging ChatGPT's capabilities to support scientific research and scholarly writing. Its potential to streamline literature reviews, assist in data interpretation, and enhance manuscript drafting is reshaping academic publishing, contributing to the evolving landscape of medical and healthcare research.

The introduction of LLMs such as ChatGPT has raised questions about their role in academic publishing, particularly regarding the volume and quality of publications, co-authorship trends, and language complexity.

Recent studies have explored the role of artificial intelligence in scientific publishing, highlighting both its benefits and challenges^{[8][9]}. AI-powered tools, including large language models like ChatGPT, have

demonstrated potential in automating literature reviews, drafting manuscripts, and analysing complex datasets^{[10][11]}. Research by Fayed *et al.*^[11] showed how AI is increasingly integrated into orthopaedic research, assisting in clinical decision-making. However, concerns regarding accuracy, bias, and ethical implications remain central to ongoing discussions in the scientific community.

Using artificial intelligence to write scientific papers raises concerns about plagiarism, compromised methodology, improper contextualisation of findings, and questions about intellectual contribution and scientific integrity^[12].

We hypothesised that the release of ChatGPT and similar AI tools would significantly influence the landscape of scientific publishing in orthopaedic research. We expected to observe an increase in publication rates and changes in authorship patterns. Furthermore, we anticipated a decline in linguistic complexity, as AI-generated texts often exhibit more structured and repetitive language patterns.

The aim of this study was to perform a bibliometric analysis of orthopaedic research articles published before and after the release of ChatGPT. By evaluating publication trends and linguistic characteristics, we sought to quantify the impact of AI on scientific communication. We believe these findings will contribute meaningfully to the ongoing discussion regarding the role of AI in academic publishing, highlighting both its potential benefits and the challenges it may pose to scientific integrity.

Design and methods

Study Design

A bibliometric review was conducted following the BIBLIO guideline for reporting bibliometric analyses of biomedical literature.^[13] Data were retrieved from PubMed, focusing on ten high-impact orthopaedic journals. These journals were purposefully selected to ensure broad representation across the major orthopaedic subspecialties. Each journal was chosen based on its established influence within its respective domain, consistent indexing in PubMed, and high citation impact. The included journals were: The Journal of Bone & Joint Surgery, The American Journal of Sports Medicine, The Journal of Arthroplasty, Knee Surgery Sports Traumatology Arthroscopy, Foot and Ankle International, European Spine Journal, Archives of Orthopaedic and Trauma Surgery, Clinical Orthopaedics and Related Research, International Orthopaedics, and Journal of Orthopaedic Research.

Search Strategy

A structured bibliometric search was conducted in PubMed to identify articles published between January 2021 and December 2025 in ten high-impact orthopaedic journals, selected based on their ISSNs. Records were retrieved using automated queries stratified by journal and publication year, ensuring completeness and mitigating download limitations. All retrieved data, including abstracts and bibliographic metadata, were converted into a structured text format using the easyPubMed and bibliometrix packages in R.^[14] To ensure data integrity, the retrieval process incorporated batch controls, duplicate checks, and verification of publication dates against inclusion criteria.

Group Definition

The dataset was stratified into two temporal groups using a washout design. Given that manuscripts published immediately after the release of ChatGPT (30 November 2022) were almost certainly written, submitted, and reviewed before its availability, a seven-month washout period was applied. Articles with an entry date (EDAT) before 1 July 2023 were classified as the pre-ChatGPT group ("Before"), while those indexed from 1 July 2023 onwards were included in the post-adoption group ("After"). This resulted in a pre-ChatGPT period of 911 days (January 2021 – June 2023) and a post-adoption period of 915 days (July 2023 – December 2025), providing approximately balanced observation windows.

Data Processing and Analysis

All data processing and statistical analyses were conducted using R.^[14] Following data retrieval, records were cleaned and structured for analysis. The number of authors per article was extracted from standardised metadata fields.

Abstracts were processed using natural language processing techniques to evaluate linguistic features. Lexical diversity was assessed using the Type-Token Ratio (TTR), and sentence complexity was estimated by calculating the average number of words per sentence.

Sentiment analysis was conducted using the *syuzhet* package in R^[15], which applies a lexicon-based sentiment extraction method. This package has been widely used in bibliometric studies to assess emotional tone in academic writing and has demonstrated reasonable performance in scientific contexts,

as it captures broad affective dimensions (positive/negative valence) rather than requiring domain-specific terminology.

We acknowledge that dictionary-based sentiment analysis has inherent limitations when applied to technical medical literature, as clinical terminology may be misclassified or lack sentiment annotations. The lexicon-based method was chosen for its transparency, reproducibility, and established use in similar bibliometric analyses. Each abstract received a continuous sentiment score reflecting its overall emotional tone, with higher values indicating more positive sentiment.

For temporal assignment, we used PubMed's EDAT as a standardised operational proxy of indexing time; this choice ensures reproducible period classification within the same retrieval framework, while acknowledging potential index-lag variability.

Statistical Analysis

Descriptive statistics were reported as means and standard deviations. Comparisons between the pre- and post-ChatGPT periods were performed using Welch's t-test for continuous variables, given the presence of unequal variances. In addition, a linear regression model was used to assess differences in daily publication rates, adjusting for monthly effects to account for seasonality. A two-sided p-value of less than 0.05 was considered indicative of statistical significance. All analyses were conducted using R.^[14]

Results

Publication Trends

A total of 21,524 articles were included in the analysis, with 10,403 published before and 11,121 in the post-adoption period (from 1 July 2023). To ensure that daily publication rates accurately reflected calendar time, including days with zero publications, a complete calendar grid was constructed and joined with observed counts. The mean number of publications per day was 11.42 (\pm 7.32) in the pre-ChatGPT period and 12.15 (\pm 8.26) in the post-adoption period (Welch's t-test: $t = 2.01$, $p = 0.044$; Cohen's $d = 0.094$). This difference remained statistically significant after adjusting for monthly seasonality using a linear regression model with monthly fixed effects, which estimated an adjusted increase of 1.15 publications per day attributable to the period ($p = 0.002$).

These findings indicate a measurable rise in publication activity in orthopaedic journals during the period following the public release of ChatGPT (Figure 1).

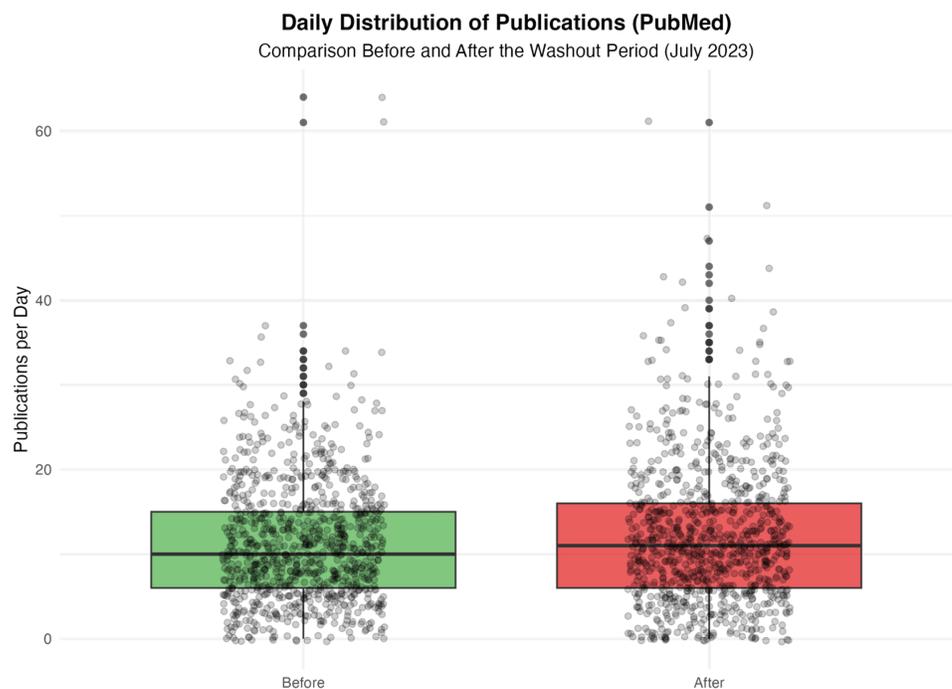


Figure 1. Boxplot showing the daily number of orthopaedic publications before (January 2021 – June 2023) and after (July 2023 – December 2025) the washout period.

Authorship Patterns

The mean number of authors per article increased from 5.94 (\pm 3.95) in the pre-ChatGPT period to 6.28 (\pm 4.39) in the post-adoption period (Welch's t-test: $t = 5.87$, $p < 0.001$; Cohen's $d = 0.08$). Figure 2 displays the proportional distribution of articles according to the number of authors, restricted to those with between one and ten contributors.

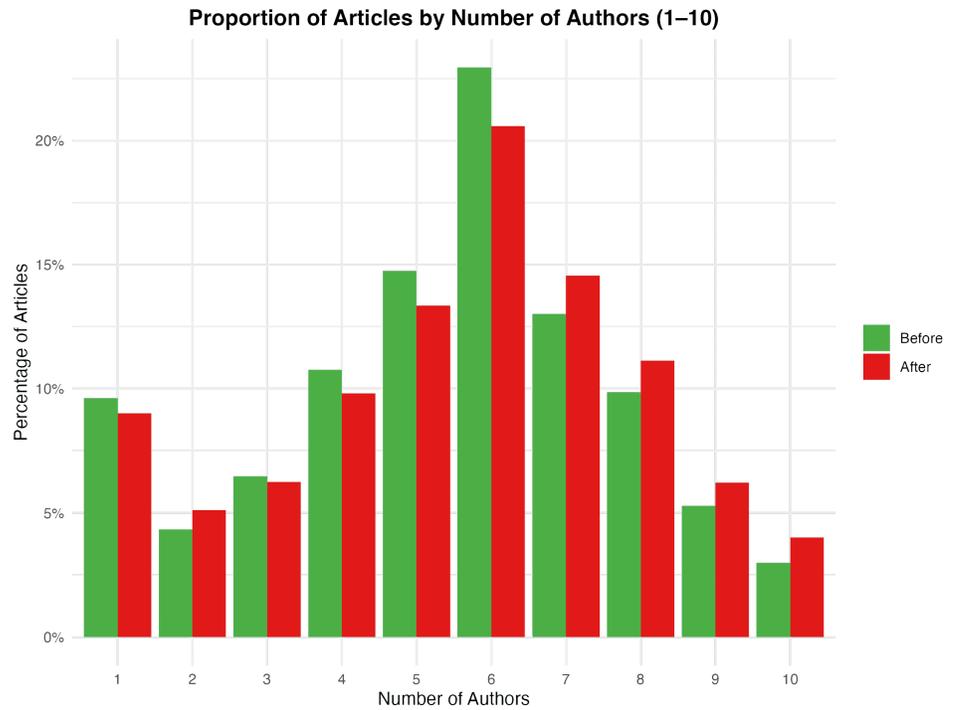


Figure 2. Bar chart showing the percentage of articles by number of authors, stratified by period.

Textual Complexity

The mean number of words per sentence in abstracts decreased from 14.81 (\pm 5.08) in the pre-ChatGPT period to 14.52 (\pm 5.06) in the post-adoption period (Welch's t-test: $t = -3.69$, $p < 0.001$; Cohen's $d = -0.056$). Figure 3 presents the density distribution of sentence length, normalised by the total number of abstracts in each period.

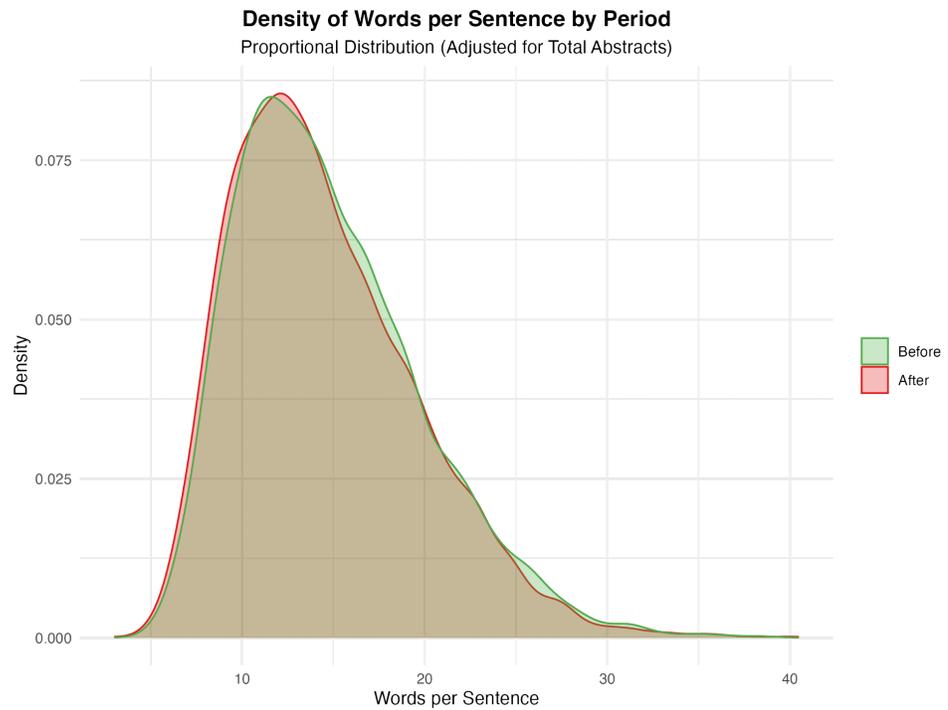


Figure 3. Density plot of the number of words per sentence in abstracts, stratified by period.

Lexical Diversity

Vocabulary diversity was evaluated using the Type-Token Ratio (TTR), calculated individually for each abstract. The mean TTR increased from 0.5195 (± 0.0666) in the pre-ChatGPT period to 0.5299 (± 0.0695) in the post-adoption period (Welch's t-test: $t = 10.04$, $p < 0.001$; Cohen's $d = 0.152$), suggesting an increase in lexical diversity. Figure 4 presents the density distribution of TTR values across both periods.

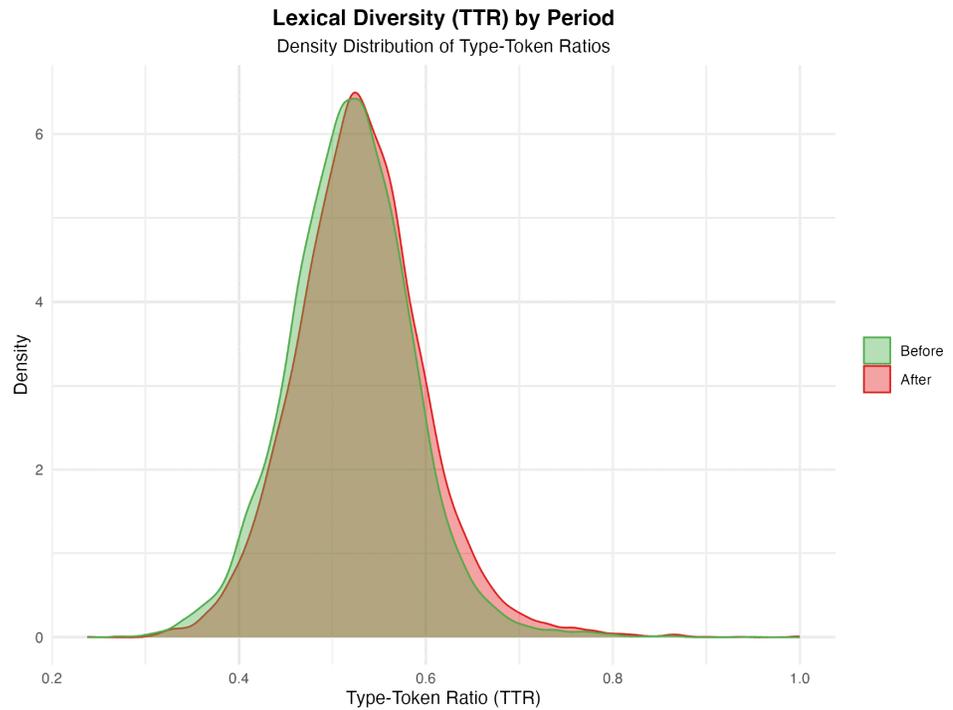


Figure 4. Density plot showing the distribution of Type-Token Ratios (TTR) in abstracts, stratified by period.

Sentiment Analysis

The mean sentiment score increased from 3.28 (\pm 3.27) in the pre-ChatGPT period to 3.64 (\pm 3.47) in the post-adoption period (Welch's t-test: $t = 7.01$, $p < 0.001$; Cohen's $d = 0.106$), indicating a shift towards a more positive tone in abstracts over time. To facilitate interpretation, sentiment scores were categorised as positive, negative, or neutral. Figure 5 presents the proportional distribution of abstracts across these sentiment categories for each period.

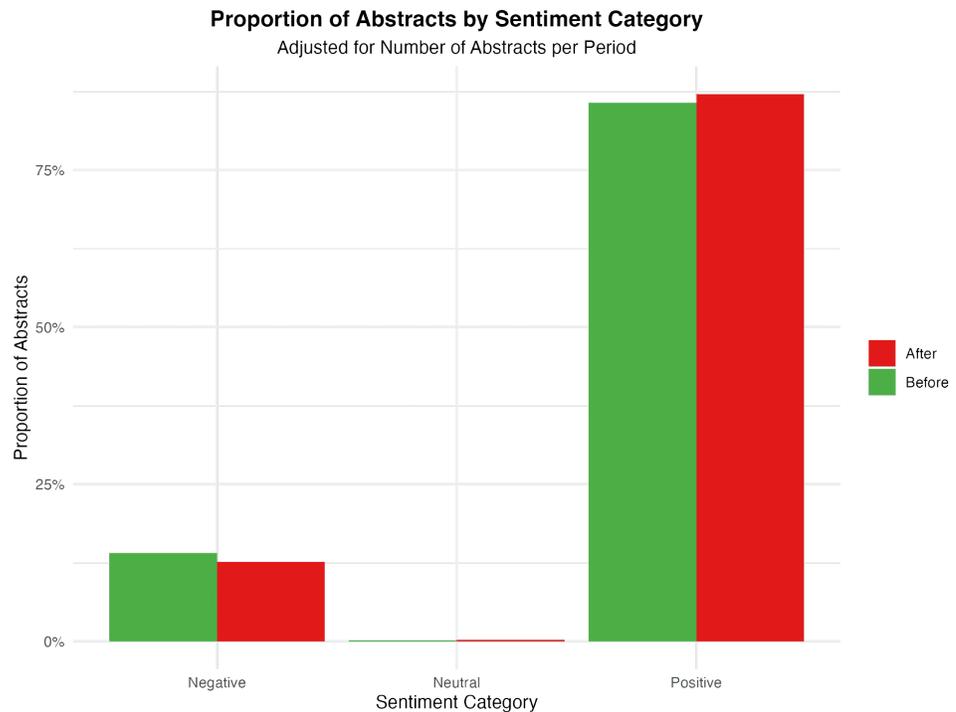


Figure 5. Bar chart showing the proportion of abstracts by sentiment category (positive, neutral, negative), stratified by period.

Discussion

This study examined whether the release of ChatGPT was associated with changes in orthopaedic scientific publishing. Following the application of a seven-month washout period to account for the typical lag between manuscript preparation and publication, our findings revealed a statistically significant increase in the mean number of articles published per day (11.42 vs. 12.15; $p = 0.044$), which remained significant after adjusting for monthly seasonality ($p = 0.002$). However, all observed effect sizes were small (Cohen's $d < 0.2$), warranting cautious interpretation.

While our analysis does not establish causality, the observed increase in publication output should be interpreted with caution. This trend may reflect not only the contemporaneous availability of large language models but also a confluence of other factors, including post-pandemic recovery in research activities and pre-existing secular growth trends in academic publishing. It is plausible, however, that some researchers have begun incorporating AI-based tools into their writing workflows, which may contribute to enhanced efficiency in manuscript preparation.

The observed increase in the average number of authors per article from 5.94 to 6.28 ($p < 0.001$; Cohen's $d = 0.08$) is consistent with broader shifts in research collaboration dynamics. This trend is more plausibly explained by the ongoing rise of team science and multi-centre studies, alongside the post-pandemic acceleration of digital connectivity, rather than by the direct influence of AI-assisted writing tools.

The widespread adoption of virtual communication platforms has lowered barriers to collaboration, enabling researchers from diverse locations to co-author work more readily ^[16]. This trend aligns with findings that highlight a surge in digital collaboration and co-authorship patterns following the COVID-19 pandemic.

Furthermore, Fayed et al. (2023) discuss how the integration of AI in orthopaedics often necessitates interdisciplinary collaboration, involving experts in machine learning, data science, and clinical practice. Such collaborations inherently increase the number of co-authors on publications, reflecting the multifaceted nature of AI-driven research in the field.^[11]

Our analysis identified statistically detectable but modest changes in the aggregate linguistic structure of orthopaedic abstracts following the washout period. The mean number of words per sentence decreased from 14.81 to 14.52 ($p < 0.001$; Cohen's $d = -0.056$), indicating a tendency towards more concise expression.

In parallel, lexical diversity, as measured by the type-token ratio (TTR), increased from 0.5195 to 0.5299 ($p < 0.001$; Cohen's $d = 0.152$).

Although statistically significant, the small magnitude of these effects suggests they do not represent a meaningful shift in writing style at the individual manuscript level. These subtle patterns may, however, reflect the early influence of large language models, which are designed to generate fluent, grammatically accurate, and easily readable text, often prioritising clarity and directness.

While the increased clarity can enhance accessibility, there is also a risk of homogenisation in scientific writing. If many researchers begin relying on the same AI tools, there may be an unintentional convergence in tone and phrasing, leading to a loss of individual authorial voice. Alkaissi and McFarlane have warned that the standardised output of ChatGPT, although fluent, may lack the depth or nuance of expert-crafted narratives.^[17]

Another notable finding was the significant increase in positive sentiment observed in abstracts. Mean sentiment scores rose from 3.28 to 3.64 ($p < 0.001$; Cohen's $d = 0.106$) in the post-adoption period. Although scientific writing is expected to maintain a neutral tone, our analysis suggests a subtle shift towards more positively framed language. This may be partly attributable to the generative tendencies of LLMs, such as ChatGPT, which often default to polite, affirmative, and constructive phrasing unless explicitly instructed otherwise.

Challenges and Ethical Considerations

Despite its advantages, AI is not infallible. One study highlights inconsistencies in AI-generated academic articles, underscoring the need for human oversight in verifying outputs.^[18] Furthermore, the reliance on proprietary AI systems could lead to restricted access to scientific knowledge, as concerns about paywalled AI-generated research platforms grow. If AI-driven research tools become monetised, there is a risk that only well-funded institutions will benefit from these advancements, limiting open science initiatives.^[6]
^{[19][20]}

While the integration of AI tools into scientific workflows shows promise, it is important to recognise their current limitations. ChatGPT and similar models are known to generate factual inaccuracies, including fabricated references and misrepresented data trends. Farhat et al. demonstrated that such tools may produce plausible-sounding but incorrect scientific content, particularly when applied to data interpretation or citation generation.^[21] Although LLMs can assist in refining the clarity and structure of text, their output still requires careful human oversight.

Nevertheless, these shortcomings may diminish over time. As statistical architectures become more sophisticated and training datasets expand, the accuracy and reliability of AI-generated content are likely to improve. The trajectory of model refinement suggests that future versions may offer enhanced factual precision, potentially increasing their utility in academic writing without compromising scientific integrity.

Additionally, Kim *et al.* examined the role of ChatGPT in medical education and its performance in standardised assessments, highlighting both its strengths and limitations.^[19] While AI demonstrates promising applications in synthesising information and assisting in academic tasks, concerns remain regarding plagiarism, reliability, and disparities in educational access.

Dave *et al.* further highlight the ethical concerns surrounding AI-generated scientific content, particularly in relation to plagiarism and transparency in authorship, reinforcing the need for rigorous validation and oversight.^[6] These findings align with our observations that AI, despite improving research efficiency, raises ethical concerns regarding transparency and the risk of diminishing critical evaluation in scientific writing. As AI tools become increasingly integrated into academia, it is crucial to establish clear guidelines to mitigate these risks and ensure their responsible use in research and publishing.

Maintaining Scientific Integrity in AI-Assisted Research

While qualitative research exhibits greater resistance to AI, fearing the loss of human interpretation, other fields appear to adopt these tools more flexibly, particularly for enhancing writing and data analysis. In both cases, there is a clear need for well-defined guidelines to ensure that AI serves as a valuable aid to research without compromising originality and academic rigour. The risks of AI-generated errors, including data misinterpretation and fabricated references, must be mitigated through structured human-AI collaboration.

Similarly, Alkaissi *et al.* demonstrated that ChatGPT can generate artificial hallucinations, where factually incorrect statements and fabricated references are confidently presented as accurate. Their findings highlight the potential risks of unchecked AI-generated content, particularly in scientific writing, where misinformation could propagate if not rigorously verified.^[17]

In 2023, Cheng et al. found that AI-generated scientific abstracts were of significantly lower quality than those written by human authors, with some containing incorrect conclusions—highlighting the importance of human oversight in AI-assisted research.^[22] While these findings underscore the limitations of earlier language models, the rapid pace of AI development suggests that such tools may already have improved. As AI systems continue to evolve, distinguishing between human- and AI-generated manuscripts could become increasingly difficult, raising important questions about authorship attribution, academic integrity, and the evolving role of human researchers in scientific writing.

The phenomenon of AI-generated inaccuracies aligns with our observation that while AI-assisted tools may improve efficiency, they also pose challenges to the integrity of scientific communication. This underscores the urgent need for robust validation frameworks to ensure that AI-generated texts uphold the highest standards of academic reliability and transparency.

Democratisation of Research Through AI

One of the major advantages of AI in publishing is its ability to bridge the gap between well-funded research institutions and smaller hospitals or under-resourced research teams. By providing automated support in data analysis and manuscript preparation, AI can empower researchers who lack access to extensive academic infrastructure. This democratisation of research tools allows clinicians and scientists from diverse backgrounds to engage in academic publishing without the same level of institutional resources traditionally required.

Despite ongoing concerns, AI holds promise for democratising scientific research by assisting individuals in resource-limited settings. As highlighted by Granjeiro et al., tools such as ChatGPT and Grammarly can help non-native English speakers and early-career researchers overcome linguistic barriers, improve the clarity and coherence of their manuscripts, and increase their visibility within the global academic community.^[8]

For these benefits to be equitably distributed, however, access to AI tools and training must be widened. Supporting infrastructure, open-source alternatives, and targeted capacity-building initiatives can all contribute to a more inclusive research ecosystem.

Implications for Journals and Peer Review

The proliferation of AI-generated content compels journals to refine their editorial workflows in a constructive and integrative manner. Rather than solely focusing on AI detection or restriction, editorial teams may increasingly harness AI tools to assist in the standardisation and quality control of submissions. For instance, AI can be applied to automatically verify whether manuscripts comply with established reporting guidelines and checklists, ensuring that authors have completed all required sections and methodological disclosures before peer review begins.

Moreover, artificial intelligence has the potential to enhance the peer review process itself. In the future, a dual-layered system may emerge wherein AI performs a parallel review, focusing on structural, statistical, and compliance aspects, alongside traditional human peer review, which continues to evaluate conceptual, scientific, and interpretative quality. This complementary model could improve the efficiency, rigour, and consistency of manuscript evaluation while preserving the irreplaceable insight provided by expert reviewers. To achieve this, journals will need to develop clear frameworks that outline the role of AI in editorial decision-making and ensure transparency and accountability throughout the process.

AI and the Future of Scientific Publishing

The increasing integration of AI in academia suggests a future where scientific research is no longer centred on human-driven analysis and interpretation. Instead of reading and synthesising literature, researchers may simply upload raw data, allowing AI to perform statistical analyses, identify patterns, compare findings with existing guidelines, and generate comprehensive reports. In this scenario, the role of the orthopaedic surgeon as a researcher could shift towards data curation and clinical validation, ensuring that AI-generated insights align with real-world practice.

While this transformation could enhance efficiency and streamline scientific production, it raises ethical and methodological concerns regarding transparency, potential algorithmic bias, and the diminishing role of human critical thinking in evaluating medical evidence. Additionally, as AI-driven research platforms grow, there is a risk that access to scientific data may become restricted, with proprietary databases limiting the availability of findings to exclusive research networks or commercial entities. What is now open-access knowledge could, in the future, be locked behind paywalls, creating disparities in who can access and contribute to medical advancements.

A recent development in this context is the launch of the FAIR² data management framework (Findable, Accessible, Interoperable, Reusable, and AI-Ready) by Frontiers.^[23] This system encourages researchers to upload structured datasets enriched with metadata, facilitating AI-driven analyses and enabling direct

citation of datasets. By promoting transparency and reproducibility, FAIR² aims to enhance the integrity of AI-assisted research and ensure that data remains a citable and integral component of scientific communication.

Editorial Implications and AI Disclosure Policies

Our findings hold several implications for editorial practice. Rather than focusing on the detection or restriction of AI-assisted writing, journals may consider a proactive approach centred on transparency. Leading editorial bodies, including the International Committee of Medical Journal Editors (ICMJE) and the Committee on Publication Ethics (COPE), have issued guidance that, while not prohibiting the use of AI tools, reaffirms that human authors retain full responsibility for the integrity and accuracy of the work. Journals can support this principle by developing clear disclosure policies requiring authors to specify which, if any, AI tools were used in manuscript preparation. The descriptive, non-punitive monitoring of bibliometric trends, such as that presented in this study, may serve as a useful tool for editors to understand publishing dynamics in the AI era without resorting to enforcement.

Strengths

This study is one of the first large-scale, orthopaedics-focused evaluations of AI's impact on scientific writing. By applying natural language processing techniques to a robust, multi-year dataset spanning multiple journals, we were able to quantify subtle but important shifts in collaboration, writing style, lexical diversity, and sentiment. Focusing on high-impact journals ensured that the observed trends reflect key changes within influential research outputs.

Limitations

Several important limitations must be acknowledged. Our analysis focused exclusively on titles and abstracts, which limits insight into full-text evolution and potentially overlooks changes in methodology sections, discussion depth, or reference patterns. Additionally, we examined only ten high-impact journals, which may not adequately represent trends across all orthopaedic publications, particularly lower-impact or regional journals, thereby limiting the generalisability of our findings.

Additionally, our analysis was restricted to published articles indexed in PubMed, which may introduce selection bias. We were unable to assess the total number of manuscript submissions to these journals, as such data are not publicly available. Consequently, whilst we observed an increase in published articles following ChatGPT's release, we cannot determine whether submission rates increased proportionally or whether acceptance rates remained stable. It is conceivable that submission volumes rose without corresponding changes in publication output, or that editorial policies evolved during this period. Future studies with access to editorial databases could provide valuable insights into submission-to-publication ratios.

Crucially, our study is correlational and cannot definitively attribute observed changes to ChatGPT or AI adoption alone. Multiple confounding factors warrant consideration. The COVID-19 pandemic fundamentally altered research collaboration through the widespread adoption of virtual platforms, potentially facilitating increased co-authorship independently of AI tools. Natural temporal trends in scientific publishing—including pressure to increase output, evolving editorial policies, and changes in submission systems—may have contributed to rising publication rates.

Moreover, ChatGPT's release coincided with broader technological advances in academic writing tools (grammar checkers, reference managers, other AI platforms), making it difficult to isolate large language model effects specifically. The linguistic changes we observed—increased lexical diversity and more positive sentiment—could alternatively reflect evolving writing standards, changing editorial preferences, or the greater participation of non-native English speakers using various language assistance tools.

Academic publishing has experienced sustained growth over recent decades, driven by factors including expanding researcher populations, institutional publication pressures, and the proliferation of open-access journals. Whilst our study period (2021–2025) was selected to minimise the influence of longer-term secular trends, we acknowledge that disentangling the specific contribution of large language models from ongoing publication growth remains methodologically challenging.

Our study period (2021–2025) was intentionally selected to provide a focused snapshot around the release of ChatGPT while minimising the direct confounding effects of the COVID-19 pandemic's most disruptive phase in 2020. We acknowledge, however, that this limited timeframe prevents a broader analysis of long-term secular trends. Future work incorporating longer historical baselines, extending ten or more years prior, would be invaluable for contextualising our findings.

Furthermore, our reliance on a binary pre-post design is a simplification. More sophisticated analytical approaches, such as interrupted time-series analyses, would be required to model underlying trends more robustly, although such causal inference was beyond the scope of our descriptive study. The use of EDAT

may also introduce some temporal smoothing due to indexing lags, potentially blunting the detection of any sharp discontinuities immediately following November 2022.

To address the methodological concern regarding the contamination of the post-ChatGPT group by legacy submissions, the present revision incorporates a seven-month washout period (December 2022 – June 2023). This ensures that articles included in the "After" group were indexed sufficiently late to plausibly reflect awareness of, or access to, large language models during the writing process. The revised study period (January 2021 – December 2025) also provides approximately balanced observation windows of 911 and 915 days for the pre- and post-adoption periods, respectively.

Finally, our sentiment analysis employed a general-purpose lexicon that, whilst validated for scientific texts, may not fully capture the nuanced emotional tone specific to medical literature. The dictionary-based approach may misclassify technical terminology or fail to account for context-dependent sentiment in clinical abstracts. Despite these limitations, our findings provide valuable insights into temporal trends in orthopaedic publishing during a period of rapid technological change.

Conclusion

In the period following the public release of ChatGPT—and after applying a seven-month washout period to mitigate the inclusion of legacy submissions—orthopaedic journals exhibited a modest but statistically significant rise in daily publication output, increases in co-authorship, and subtle changes in the linguistic style and sentiment of abstracts. All observed effect sizes were small (Cohen's $d < 0.2$). Although bibliometric data cannot establish causality, these temporal associations are compatible with the growing uptake of AI-assisted tools. However, they must be interpreted in the context of concurrent factors, including long-term publication growth and evolving editorial practices.

It remains essential for the research community to address the ethical and practical implications of AI use, promoting transparent author disclosures and clear editorial guidance. Future work should examine specific pathways by which AI affects manuscript preparation and review, and evaluate policies that maintain scientific quality while enabling the responsible, productive use of these tools.

Statements and Declarations

Funding

No specific funding was received for this work.

Potential Competing Interests

No potential competing interests to declare.

Ethics

This study was conducted using publicly available bibliometric data and did not involve human participants or require ethical approval.

Data Availability

The dataset used in this study was derived from PubMed-indexed journals using automated queries conducted between January 2021 and December 2025. The complete dataset, including bibliographic metadata and processed abstracts, is available upon reasonable request to the corresponding author. The data extraction and analysis procedures are fully described in the Methods section.

Author Contributions

All authors contributed to the conception, analysis, interpretation, and drafting of the manuscript.

References

1. ^{a, b, c}Fayed AM, Mansur NSB, de Carvalho KA, Behrens A, DHooghe P, de Cesar Netto C (2023). "Artificial Intelligence and ChatGPT in Orthopaedics and Sports Medicine." *J Exp Orthop*. **10**(1):74. doi:[10.1186/s40634-023-00642-8](https://doi.org/10.1186/s40634-023-00642-8).
2. ^ΔMarshall DT, Naff DB (2024). "The Ethics of Using Artificial Intelligence in Qualitative Research." *J Empir Res Hum Res Ethics*. doi:[10.1177/15562646241262659](https://doi.org/10.1177/15562646241262659).
3. ^ΔPhillips F (2024). "From My Perspective: Artificial Intelligence and Research Productivity." *Technol Forecast Soc Change*. doi:[10.1016/j.techfore.2024.123937](https://doi.org/10.1016/j.techfore.2024.123937).
4. ^ΔChakraborty C, Bhattacharya M, Lee SS (2024). "Need an AI-Enabled, Next-Generation, Advanced ChatGPT or Large Language Models (LLMs) for Error-Free and Accurate Medical Information." *Ann Biomed Eng*. **52**

- (2):134135. doi:[10.1007/s10439-023-03297-9](https://doi.org/10.1007/s10439-023-03297-9).
5. [△]Adamopoulou E, Moussiades L (2020). "An Overview of Chatbot Technology." In: *Artificial Intelligence Applications and Innovations*. Cham: Springer International Publishing. doi:[10.1007/978-3-030-49186-4_31](https://doi.org/10.1007/978-3-030-49186-4_31).
 6. [△][♠][♣]Dave T, Athaluri SA, Singh S (2023). "ChatGPT in Medicine: An Overview of Its Applications, Advantages, Limitations, Future Prospects, and Ethical Considerations." *Front Artif Intell*. 6:1169595. doi:[10.3389/frai.2023.1169595](https://doi.org/10.3389/frai.2023.1169595).
 7. [△]Ray PP (2023). "ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope." *Internet Things Cyber Phys Syst*. 3:121154. doi:[10.1016/j.iotcps.2023.04.003](https://doi.org/10.1016/j.iotcps.2023.04.003).
 8. [△][♠]Granjeiro JM, Del Bel Cury AA, Cury JA, Bueno M, Sousa-Neto MD, Estrela C (2025). "The Future of Scientific Writing: AI Tools, Benefits, and Ethical Implications." *Braz Dent J*. 36:e256471. doi:[10.1590/0103-644020256471](https://doi.org/10.1590/0103-644020256471).
 9. [△]Khalifa M, Albadawy M (2024). "Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool." *Comput Methods Programs Biomed Update*. 5:100145. doi:[10.1016/j.cmpbup.2024.100145](https://doi.org/10.1016/j.cmpbup.2024.100145).
 10. [△]Marshall IJ, Kuiper J, Wallace BC (2016). "RobotReviewer: Evaluation of a System for Automatically Assessing Bias in Clinical Trials." *J Am Med Inform Assoc*. 23(1):193201. doi:[10.1093/jamia/ocv044](https://doi.org/10.1093/jamia/ocv044).
 11. [△]van de Schoot R, de Bruin J, Schram R, Zahedi P, de Boer J, Weijdema F, Kramer B, Huijts M, Hoogerwerf M, Ferdinands G, Harkema A, Willemsen J, Ma Y, Fang Q, Hindriks S, Tummers L, Oberski DL (2021). "An Open Source Machine Learning Framework for Efficient and Transparent Systematic Reviews." *Nat Mach Intell*. 3(2):125133. doi:[10.1038/s42256-020-00287-7](https://doi.org/10.1038/s42256-020-00287-7).
 12. [△]Gravel J, DAmours-Gravel M, Osmanlliu E (2023). "Learning to Fake It: Limited Responses and Fabricated References Provided by ChatGPT for Medical Questions." *Mayo Clin Proc Digit Health*. 1(3):226234. doi:[10.1016/j.mcpdig.2023.05.004](https://doi.org/10.1016/j.mcpdig.2023.05.004).
 13. [△]Montazeri A, Mohammadi S, Hesari PM, Ghaemi M, Riazi H, et al (2023). "Preliminary Guideline for Reporting Bibliometric Reviews of the Biomedical Literature (BIBLIO): A Minimum Requirements." *Syst Rev*. 12:239. doi:[10.1186/s13643-023-02410-2](https://doi.org/10.1186/s13643-023-02410-2).
 14. [♠][♣]R Core Team (2020). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.r-project.org/>.
 15. [△]Jockers M (2023). "syuzhet: Extracts Sentiment and Sentiment-Derived Plot Arcs from Text (R Package)." *Comprehensive R Archive Network (CRAN)*. doi:[10.32614/CRAN.package.syuzhet](https://doi.org/10.32614/CRAN.package.syuzhet).
 16. [△]Tao Y, Steckel D, Kleme JJ, You F (2021). "Trend Towards Virtual and Hybrid Conferences May Be an Effective Climate Change Mitigation Strategy." *Nat Commun*. 12(1):7324. doi:[10.1038/s41467-021-27251-2](https://doi.org/10.1038/s41467-021-27251-2).
 17. [♠][♣]Alkaiissi H, McFarlane SI (2023). "Artificial Hallucinations in ChatGPT: Implications in Scientific Writing." *Cureus*. 15(2):e35179. doi:[10.7759/cureus.35179](https://doi.org/10.7759/cureus.35179).
 18. [△]Ariyaratne S, Iyengar KP, Nischal N, Chitti Babu N, Botchu R (2023). "A Comparison of ChatGPT-Generated Articles with Human-Written Articles." *Skeletal Radiol*. 52:17551758. doi:[10.1007/s00256-023-04340-5](https://doi.org/10.1007/s00256-023-04340-5).
 19. [♠][♣]Kim TW (2023). "Application of Artificial Intelligence Chatbots, Including ChatGPT, in Education, Scholarly Work, Programming, and Content Generation and Its Prospects: A Narrative Review." *J Educ Eval Health Prof*. 20:38. PMID [38148495](https://pubmed.ncbi.nlm.nih.gov/38148495/).
 20. [△]Salvagno M, Taccone FS, Gerli AG (2023). "Can Artificial Intelligence Help for Scientific Writing?" *Crit Care*. 27:75. doi:[10.1186/s13054-023-04380-2](https://doi.org/10.1186/s13054-023-04380-2).
 21. [△]Farhat F, Sohail SS, Madsen D (2023). "How Trustworthy Is ChatGPT? The Case of Bibliometric Analyses." *Cogent Eng*. 10(1):2222988. doi:[10.1080/23311916.2023.2222988](https://doi.org/10.1080/23311916.2023.2222988).
 22. [△]Cheng S-L, Tsai S-J, Bai Y-M, Tsai P-H, Liang C-S, Chen Y-Y, Hu W-S, Hsu C-N, Su T-P (2023). "Comparisons of Quality, Correctness, and Similarity Between ChatGPT-Generated and Human-Written Abstracts for Basic Research: Cross-Sectional Study." *J Med Internet Res*. 25:e51229. doi:[10.2196/51229](https://doi.org/10.2196/51229).
 23. [△]Frontiers (2025). "FAIR Data Management." *Frontiers*. <https://www.frontiersin.org/about/fair-data-management>.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.