

Research Article

Citation Inaccuracies in Orthopaedic Surgery: A Novel Classification and Precautions for AI-Generated Bibliographies

Erik R Nakken, M.D.¹^a, Victoria Aquilino, B.S.²^b, Jacqueline K Kobayashi, M.D.¹^c, Emily K Porter¹^d, William R Aibinder, M.D.¹^e

¹ Department of Orthopaedic Surgery, University of Michigan, ² Michigan State University

Keywords: Citation accuracy, Artificial intelligence, Orthopaedic surgery, Bibliography, Scholarly integrity

<https://doi.org/10.60118/001c.147398>

Journal of Orthopaedic Experience & Innovation

Vol. 7, Issue 1, 2026

Background

The integration of artificial intelligence (AI) in medical research is becoming increasingly common. While AI has the potential to assist in various research tasks, its effectiveness in accurately identifying citations and compiling bibliographies is hampered by existing human errors linked to citation inaccuracies. This study aims to: 1) determine the

-
- a Dr. Erik Nakken is an orthopaedic surgery resident at the University of Michigan pursuing a sports medicine fellowship. His research interests include citation accuracy and the use of artificial intelligence in research, contact pressure reduction during lower extremity immobilization, outcomes after total shoulder arthroplasty, and prosthetic joint infection. He enjoys spending time with his family, as well as camping, hiking, grilling, and watching Michigan football.

[Connect with Dr. Erik Nakken on LinkedIn](#)

[Conflicts of Interest Statement for Dr. Erik Nakken](#)

[Visit the Open Payments Data Page for Dr. Erik Nakken](#)

- b Victoria Aquilino is a medical student at Michigan State University College of Human Medicine with a strong interest in orthopaedic surgery, sparked by her work in a multidisciplinary orthopaedic surgery clinic prior to starting medical school. Since then, she has continually sought out opportunities to deepen her exposure to both the clinical and research aspects of orthopaedics, strengthening her commitment to advancing patient care within the specialty.

[Conflicts of Interest Statement for Victoria Aquilino](#)

- c Jacqueline Kobayashi, MD, is an orthopaedic surgery resident at the University of Michigan. She initially earned an Honors Bachelor of Science degree in Kinesiology and Human Sciences at Oregon State University. She then completed her Doctorate of Medicine degree at University of Utah. She is currently a fourth-year resident pursuing a fellowship in shoulder and elbow surgery.

[Conflicts of Interest Statement for Dr. Jacqueline Kobayashi](#)

[Visit the Open Payments Data Page for Dr. Jacqueline Kobayashi](#)

- d Emily Porter is a current student at the University of Michigan Ann Arbor completing a major in Molecular, Cellular, and Developmental Biology. She has previous research experience in Orthopaedics and Osteoarthritis and continues to be interested in this field.

[Conflicts of Interest Statement for Emily Porter](#)

- e Dr. Aibinder is an associate professor, co-director of the Sports Medicine and Shoulder Surgery Fellowship, medical student clerkship director, and a shoulder and elbow surgeon at Michigan Medicine. He graduated from the Macaulay Honors College at the City University of New York (CUNY) Brooklyn College, and then went on to earn his medical degree from SUNY Downstate Health Sciences University as part of a combined B.A.-M.D. program. After medical school, he completed his orthopaedic surgery residency at the Mayo Clinic in Rochester, MN followed by fellowship training in upper extremity surgery with a special focus on shoulder and elbow at the internationally renowned Roth McFarlane Hand and Upper Limb Centre at the University of Western Ontario in London, ON.

Dr. Aibinder specializes in all disorders of the shoulder and elbow, including replacement surgery, rotator cuff surgery, instability surgery, and fracture surgery. Prior to joining the faculty at the University of Michigan, Dr. Aibinder served as the Director of Shoulder and Elbow Surgery and the Associate Residency Program Director at SUNY Downstate Health Sciences University. He was awarded the Teacher of the Year Award in 2021 at SUNY Downstate and at the University of Michigan in 2024. He has been selected as an Emerging Leader by the American Orthopaedic Association (AOA) and was awarded the ASES Young Scholars Memorial Award in 2022.

[Visit Dr. William Aibinder](#)

[Conflicts of Interest Statement for Dr. William Aibinder](#)

[Visit the Open Payments Data Page for Dr. William Aibinder](#)

prevalence of citation inaccuracies in orthopaedic surgery literature, and 2) propose a pragmatic classification system to clearly define citation accuracy to determine the validity of AI application in citation collection.

Methods

We reviewed existing literature on citation accuracy to develop a novel citation accuracy classification system to determine if information is correctly cited. This scoring system distinguishing between major and minor errors, with subcategories that include secondary article citations, oversimplifications, statistical inaccuracies, missing facts, and contradictory conclusions, rated on a scale of 1 to 5. We analyzed 229 references from seven scientific articles and applied our novel classification system to assess accuracy. We also evaluated two artificial intelligence (AI) platforms for their potential to automate citation accuracy checks.

Results

Improper citations were identified in 27.5% of the 229 references analyzed. Twenty references (8.3%) cited a secondary article. Minor errors were present in 25 references (10.9%), while major errors were found in 21 references (9.2%). Both tested AI platforms failed to complete a thorough review due to an inability to apply the classification system or access full-text sources.

Conclusions

We present a novel citation accuracy classification system that can be applied across scientific disciplines. Citation inaccuracies are prevalent in orthopaedic surgery literature, with nearly 10% of citations containing major errors. These findings highlight the high prevalence of citation inaccuracies in orthopaedic surgery literature and raise concern about the unchecked use of AI tools that may inadvertently propagate such errors. Combining rigorous human oversight with carefully integrated AI support may help improve the accuracy and reliability of scholarly references in the future.

Statement of Clinical Significance

The proposed citation accuracy classification system provides a foundation for characterizing inaccuracies, and we emphasize concerns about the unchecked deployment of AI tools that may inadvertently perpetuate the citation errors already present in the literature.

INTRODUCTION

References play a central role in scientific literature, serving multiple critical purposes. Primarily, they prevent plagiarism by crediting original authors. Citations provide evidence for claims, enhancing the validity of research findings. Additionally, they create a structured network of information, allowing readers to trace methodologies and findings with precision. When studies are frequently cited, they have the power to shape their field by establishing foundational knowledge and driving further research. The importance of citations is indisputable.

The pace of scientific discovery has dramatically accelerated in recent decades, especially considering rapid advancements in generative artificial intelligence (AI), such as ChatGPT (OpenAI 2025). Large language models, often trained on scientific literature, depend on accurate citations to generate dependable responses. The importance of references within scientific literature is once again underscored by the role they play in AI. Yet, existing publications demonstrate some inaccuracies within the published literature, ranging from as low as 13.6% (Gazendam et al. 2021) to as high as 84.1% (Montenegro et al. 2021). Multiple

studies examining orthopaedic literature have attempted to quantify inaccuracies within subspecialties including sports medicine (Gazendam et al. 2021), spine (Montenegro et al. 2021), pediatrics (Davids et al. 2010), and foot and ankle (Luo et al. 2013). Error classification methodology varied, with some literature documenting merely the prevalence of errors, while others used more granular methods to additionally differentiate between the severity of each inaccuracy. For example, in two separate analyses of reference accuracy, both Mogull (Mogull 2017) and Montenegro et al (Montenegro et al. 2021). initially determined the prevalence of errors and subsequently classified them as either major or minor. However, to further convolute matters, the criteria for major and minor errors differed between the two papers. Citation fidelity remains crucial, but there is a demonstrated need for a framework to classify the severity of citation inaccuracies.

The primary aim of our study is to assess the prevalence of citation inaccuracies in orthopedic surgery literature. The secondary aim of our study is to propose a pragmatic and reproducible classification system for defining and categorizing citation inaccuracies. We hypothesize the rate of citation inaccuracies in orthopedic surgery literature to be

Table 1. A Novel Citation Accuracy Classification System

Classification Number	Error Type	Definition	Formal Clarification
1	Accurate	Primary reference cited correctly	duplicate of another cited fact in the referenced article (i.e. the original source was not cited)
2a		Secondary reference cited accurately	
2b	Minor Error	Secondary reference cited inaccurately	duplicate of another cited fact in the referenced article (i.e. the original source was not cited), that includes oversimplification, overgeneralization, or trivial inaccuracies
3a		Inaccurate statistics or population	statistical misrepresentations or the extrapolation of findings to vastly different populations, that only includes oversimplification, overgeneralization, or trivial inaccuracies
3b	Major Error	Inaccurate statistics or population	statistical misrepresentations or the extrapolation of findings to vastly different populations that also fail to substantiate, are unrelated to, or contradict the assertion of the primary reference
4		Fact not present	statements cited, either direct or implied, are not supported by the referenced article
5		Contradictory conclusion	claims contradict the referenced article's conclusion

similar to prior studies, and thus we should be cautious about the appropriateness of AI for citation collection and bibliography generation.

METHODS

We performed a literature search using the PubMed database to assess for prior reports on citation accuracy as well as citation accuracy definitions. Search terms included a combination of “citation accuracy,” “citation accuracies,” “citation inaccuracy,” “citation inaccuracies,” “definition,” and “orthopaedic surgery,” with and without quotations. No timeline restriction was applied. Citation error definitions were combined to create a comprehensive and novel citation accuracy classification, adopted predominantly from the works of Choi, et al. and Mogull (Table 1) (Mogull 2017; Choi et al. 2021). Our goal was to assess the level of accuracy of the information cited, rather than errors in article location (i.e. journal name, volume, edition, page numbers, etc.)

The first issue of 2024 from the American edition of *The Journal of Bone and Joint Surgery* (JBJS, Volume 106, Issue 1) was selected for review. All 229 citations from all seven scientific articles from this issue were included in our study with no exclusions (Larson, Marks, Gonzalez Sepulveda, et al. 2024; Anastasio, Kim, Wixted, et al. 2024; Skolasky et al. 2024; Richardson et al. 2024; Harris et al. 2024; MacK-echnie, Shearer, Verhofstad, et al. 2024; Singh, Jolissaint, Kohler, et al. 2024). The full-text articles for each citation were obtained electronically via institutional access or interlibrary loan methods and assessed for citation accuracy via manual review. If a secondary source was cited, the orig-

inal source was searched for citation accuracy. For example, if paper A cited a statement from paper B when paper B was quoting paper C, then paper C was also searched for the information stated in paper A.

Each citation was given an alpha-numeric rating based on our novel classification system (Table 1). The classification scale ranges from 1 to 5 and citations are further delineated as accurate, minor errors, or major errors. Class 1 includes citations that are accurate and support the primary author's claims and/or their statistics. Class 2 errors include secondary citations and are subcategorized into 2A and 2B. A class 2A error signifies accuracy from the primary source but does not earn a class 1 designation due to citation of a secondary article. Class 2B errors include an oversimplification, overgeneralization, or trivial inaccuracy from either the primary or secondary source. Class 3 errors also include inaccurate statistics or population applications, subclassified into minor or major errors. A class 3A error is equivalent to a class 2B error, but from a primary source. A class 3B error involves statistical misrepresentations, inaccurate population applications, and contradictions. If direct or implied statements were absent or not supported, then a class 4 error was designated. This also included citations that were numbered or placed incorrectly in the manuscript, as such errors would prevent AI from accurately identifying the correct source. Lastly, a class 5 error includes statements that are contradictory to or refute the cited paper's conclusion. This is a stronger contradiction than may be observed in a class 3B error, one only based on a simple statement. Minor errors include error classes 2B and 3A that contain an oversimplification, overgeneralization, or trivial inaccuracies. Major errors include error classes 3B,

4, and 5, and fail to substantiate claims, are contradictory, or are absent entirely. Citations were not classified by typographical errors of the reference section as all cited publications were able to be obtained.

Error rates and 95% confidence intervals were calculated using Microsoft Excel (Redmond, VA). Significance was determined as $p < 0.05$. A 2-proportion Z-test was used to assess for association of degree, professorship, and training status of the first, senior, and corresponding authors with citation error rates. Finally, the citation accuracy classification system was entered into ChatGPT (OpenAI 2025) and OpenEvidence ("OpenEvidence" 2025) to test AI's application of the rubric for these same seven articles.

RESULTS

All full-text articles from the 229 citations referenced by the seven scientific articles in the first issue of JBJS in 2024 were reviewed for accuracy (Larson, Marks, Gonzalez Sepulveda, et al. 2024; Anastasio, Kim, Wixted, et al. 2024; Skolasky et al. 2024; Richardson et al. 2024; Harris et al. 2024; MacKechnie, Shearer, Verhofstad, et al. 2024; Singh, Jolissaint, Kohler, et al. 2024). First authors included an attending physician, research professor, director of global programs, three residents, and a medical student. Five of the first authors had a doctorate degree. Senior authors included five attending professors, an associate professor, and an assistant professor, all of orthopaedic surgery. Corresponding authors included four professors of orthopaedic surgery, a research professor, an assistant professor of orthopaedic surgery, and a resident. The median number of named authors was 7, not including additional study groups (range, four to fifteen). The mean number of citations was 32.7 (range, 24 to 48).

Improper citations were noted in 27.5% of all references (63 of 229, 95% confidence interval (CI): 21.7%-33.3%; see Table 2, Figure 1). This included minor errors, major errors, and secondary citations that were accurate. Twenty (8.7%) of the references cited a secondary article, of which 3 were inaccurate (95% CI: 5.1%-12.4%). The total citation error rate was 20.1% (46 of 229, 95% CI: 14.9%-25.3%). A minor error was found in 25 references (10.9%, 95% CI: 6.9%-15.0%) and a major error in 21 references (9.2%, 95% CI: 5.4%-12.9%).

Citations were given a class 1 rating for accuracy 72.5% of the time (166 of 229 references, 95% CI: 66.7%-78.3%) with an additional seventeen accurate citations that were from secondary sources for a total accuracy rate of 79.9% (183 of 229, 95% CI 68.7%-80.0%). The rate of a class 2B error was 1.3% (3 of 229, 95% CI: -0.2%-2.8%). A class 3A or 3B error was observed in 9.6% (22 of 229, 95% CI: 5.8%-13.4%) and 0.4% (1 of 229, 95% CI: -0.4%-1.3%) respectively. We observed seventeen (7.4%, 95% CI: 4.0%-10.8%) class 4 errors and three (1.3%, 95% CI: -0.2%-2.8%) class 5 errors.

There was no statistical significance reached to show an association between degree or training level of the first author and any error type or prevalence. Additionally, there was no association between senior author professorship

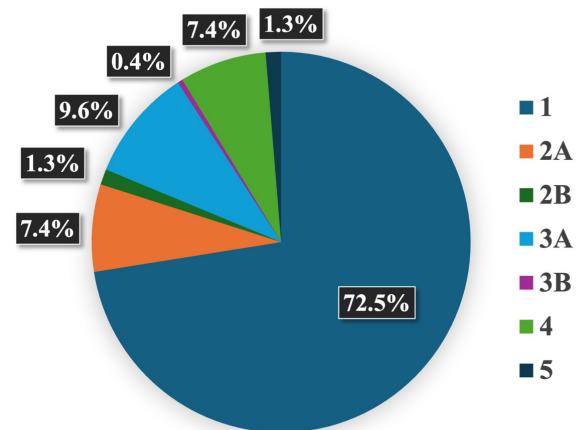


Figure 1. Graphical representation of the accuracy of 229 reviewed citations based on our novel citation accuracy classification system (Table 1).

and citation inaccuracy prevalence. All senior authors were attending orthopaedic surgeons, therefore, no calculation for association of training type was indicated. Lastly, there was also no association between professorship or training level of the corresponding author and error prevalence (Table 3).

ChatGPT self-reported the ability to evaluate the accuracy of citations by "direct comparison of the quoted text to the original text," and via "context and nuance checks (misrepresented conclusions, ignored limitations, exaggerative language)" (OpenAI 2025). It could also apply our provided rubric for citation review. However, it could not directly verify this information against the original 229 source citations and instead provided accuracy ratings based on a "logical fact-check using well-established knowledge and standard authoritative sources." (OpenAI 2025) Additionally, although ChatGPT requested the full-text PDFs of each reference for a more in-depth citation review, the platform did not allow file uploads, which ultimately made a thorough source verification impossible, therefore, no comparison of human versus AI review using Chat GPT was possible. The OpenEvidence platform was unable to apply the classification system entirely.

DISCUSSION

We found a total citation error rate of 20.1%, with 9.2% categorized as major errors. This aligns with previous studies on the prevalence of citation inaccuracies. As such, we advise caution when relying solely on AI for citation collection and bibliography generation, given the inherent human errors present in scientific literature. Additionally, we identified multiple definitions of citation inaccuracy, contributing to inconsistent representations of citation error prevalence. To address this, we propose a novel classification system for citation accuracy that can be easily applied across all scientific disciplines (Table 1).

Table 2. Prevalence of citation errors based on our novel citation accuracy classification system (Table 1). Superscript correlates to citation number in the reference section.

Error Class	<u>1 (%)</u>		<u>2A (%)</u>		<u>2B (%)</u>		<u>3A (%)</u>		<u>3B (%)</u>		<u>4 (%)</u>		<u>5 (%)</u>		Total
Article 1 (Larson, Marks, Gonzalez Sepulveda, et al. 2024)	22	(75.9%)	1	(3.4%)	0	(0%)	6	(20.7%)	0	(0%)	0	(0%)	0	(0%)	29
Article 2 (Anastasio, Kim, Wixted, et al. 2024)	22	(75.9%)	0	(0%)	2	(6.9%)	3	(10.3%)	0	(0%)	2	(6.9%)	0	(0%)	29
Article 3 (Skolasky et al. 2024)	26	(81.3%)	0	(0%)	0	(0%)	2	(6.3%)	1	(3.1%)	3	(9.4%)	0	(0%)	32
Article 4 (Richardson et al. 2024)	30	(62.5%)	7	(14.6%)	1	(2.1%)	2	(4.2%)	0	(0%)	6	(12.5%)	2	(4.2%)	48
Article 5 (Harris et al. 2024)	25	(80.6%)	1	(3.2%)	0	(0%)	4	(12.9%)	0	(0%)	1	(3.2%)	0	(0%)	31
Article 6 (MacKechnie, Shearer, Verhofstad, et al. 2024)	16	(66.7%)	7	(29.2%)	0	(0%)	0	(0%)	0	(0%)	1	(4.2%)	0	(0%)	24
Article 7 (Singh, Jolissaint, Kohler, et al. 2024)	25	(69.4%)	1	(2.8%)	0	(0%)	5	(13.9%)	0	(0%)	4	(11.1%)	1	(2.8%)	36
	166	(72.5%)	17	(7.4%)	3	(1.3%)	22	(9.6%)	1	(0.4%)	17	(7.4%)	3	(1.3%)	229

Table 3. Association of level of degree, training, and professorship with citation error prevalence. All senior authors were non-trainees and no training level association calculation was indicated. No statistically significant relationship was observed.

	Degree	#	Any Improper Citation			Secondary Citation			Minor Error			Major Error			Any Error		
			n	%	p	n	%	p	n	%	p	n	%	p	n	%	p
First Author	Doctorate	5	37	(16.2%)	0.674	5	(2.2%)	0.765	22	(9.6%)	0.665	12	(5.2%)	0.918	34	(14.8%)	0.475
	Non-Doctorate	2	26	(11.4%)		15	(6.6%)		3	(1.3%)		9	(3.9%)		12	(5.2%)	
	<u>Training Level</u>																
	Trainee	4	42	(18.3%)	0.442	12	(5.2%)	0.892	17	(7.4%)	0.769	16	(7.0%)	0.745	33	(14.4%)	0.506
	Non-Trainee	3	21	(9.2%)		8	(3.5%)		8	(3.5%)		5	(2.2%)		13	(5.7%)	
Senior Author	<u>Professorship</u>																
	Attending Professor	5	38	(16.6%)	0.621	10	(4.4%)	1.000	17	(7.4%)	0.769	11	(4.8%)	0.972	28	(12.2%)	0.718
	Attending Non-Professor	2	25	(10.9%)		10	(4.4%)		8	(3.5%)		10	(4.4%)		18	(7.9%)	
Corresponding Author	<u>Professorship</u>																
	Attending/Research Professor	5	38	(16.6%)	0.621	10	(4.4%)	1.000	17	(7.4%)	0.769	11	(4.8%)	0.972	28	(12.2%)	0.718
	Non-Professor	2	25	(10.9%)		10	(4.4%)		8	(3.5%)		10	(4.4%)		18	(7.9%)	
	<u>Training Level</u>																
	Trainee	1	7	(3.1%)	0.232	2	(0.9%)	0.740	5	(2.2%)	0.674	2	(0.9%)	0.729	7	(3.1%)	0.396
	Non-Trainee	6	56	(24.5%)		18	(7.9%)		20	(8.7%)		19	(8.3%)		39	(17.0%)	
	All Errors	7	63	(27.5%)		20	(8.7%)		25	(10.9%)		21	(9.2%)		46	(20.1%)	

= number of authors per category, n = number of citation errors, % = percentage, p = p-value.

Four authors participated in the citation accuracy review. Before the formal evaluation, each author independently assessed the same ten citations to establish inter-observer reliability. Given the subjective nature of the classification system, we encountered varying opinions regarding appropriate error designations. To address this, we agreed to classify errors by selecting the more severe designation when two ratings were deemed applicable. Ultimately, we concluded that the most significant contributions of our study lie in the introduction of a novel classification system and the identification of errors, rather than solely quantifying the exact incidence of a specific error type. While we recognize that the subjectivity inherent in our classification may limit reproducibility with exact precision, it highlights the need for ongoing dialogue and refinement in the assessment of citation accuracy.

In scientific literature, authors hold primary responsibility for ensuring the accuracy of citations, a task that demands both diligence and integrity (Cowell 1998). Once an inaccuracy is published, it may be inadvertently propagated, particularly when subsequent authors lack full-text access due to licensing restrictions or cost barriers. As artificial intelligence continues to permeate academic workflows, it is important to recognize its potential limitations in this context. AI, such as ChatGPT, may generate inaccurate reference lists due to pre-existing human errors. The validity of the articles it compiles can also be questionable, since AI may not be able to verify the accuracy on its own. Recent studies have shown that between 20.6% and 98.1% of articles compiled by various ChatGPT models were fabricated (Chen and Chen 2023), and up to 70% of citations generated by AI-only articles may be inaccurate (Kacena, Plotkin, and Fehrenbacher 2024). These findings underscore the importance of authors to maintain “citation integrity” (Howard and Englesbe 2022) and to implement appropriate oversight to prevent AI tools from compounding existing challenges in reference accuracy.

This study further illustrates both the promise and the current limitations of applying AI to citation accuracy assessment. While ChatGPT was able to interpret and apply

our structured classification system, it lacked the ability to cross-verify citations against the original 229 source documents. As a result, accuracy ratings were derived from general knowledge and logical inference rather than direct evidence from the primary literature. Although the platform indicated that full-text access would enable a more robust audit, the inability to upload source files remains a significant constraint. OpenEvidence can access full-text scientific literature, however, it was unable to apply our classification system. This highlights an important gap between emerging AI tools and the practical requirements of rigorous citation verification. Future integration of AI tools with institutional full-text databases or linked research repositories may help bridge this gap, enhancing the feasibility and rigor of AI-assisted citation verification within the literature.

CONCLUSION

In this study, we present a novel citation accuracy classification system that has the potential to be applied across scientific disciplines. Our findings confirm that citation inaccuracies remain prevalent in orthopaedic surgery literature, with nearly 10% of references containing major errors. While artificial intelligence, such as ChatGPT, shows promise for assisting with citation screening, limitations remain, including the inability to directly verify original sources. Recent evidence shows that AI-generated citations may themselves be inaccurate or fabricated, underscoring the critical need for oversight. Authors hold primary responsibility for maintaining citation integrity to prevent the propagation of inaccuracies in future research. Moving forward, combining human diligence with thoughtfully integrated AI tools may help strengthen the accuracy and reliability of scholarly references.

Submitted: September 09, 2025 EDT. Accepted: November 15, 2025 EDT. Published: May 27, 2026 EDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-ND-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-nc-nd/4.0> and legal code at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode> for more information.

REFERENCES

- Anastasio, A. T., B. I. Kim, C. M. Wixted, et al. 2024. "Younger Patients Undergoing Total Ankle Arthroplasty Experience Higher Complication Rates and Worse Functional Outcomes." *Journal of Bone and Joint Surgery* 106 (1): 10–20. <https://doi.org/10.2106/JBJS.23.00122>.
- Chen, A., and D.O. Chen. 2023. "Accuracy of Chatbots in Citing Journal Articles." *JAMA Netw Open* 6 (8): e2327647. <https://doi.org/10.1001/jamanetworkopen.2023.27647>.
- Choi, J., A. Gupta, A. Kaghazchi, T. S. Htwe, M. Baiocchi, and D. A. Spain. 2021. "Citation Inaccuracies in Influential Surgical Journals." *JAMA Surgery* 156 (8): 791–92. <https://doi.org/10.1001/jamasurg.2021.1445>.
- Cowell, H. R. 1998. "Editorial - Ethics of Medical Authorship." *Journal of Bone and Joint Surgery* 80-A (2): 151–53. <https://doi.org/10.2106/00004623-199802000-00001>.
- Davids, J. R., D. M. Weigl, J. P. Edmonds, and D. W. Blackhurst. 2010. "Reference Accuracy in Peer-Reviewed Pediatric Orthopaedic Literature." *The Journal of Bone and Joint Surgery-American Volume* 92 (5): 1155–61. <https://doi.org/10.2106/JBJS.I.00063>.
- Gazendam, A., D. Cohen, S. Morgan, S. Ekhtiari, and M. Ghert. 2021. "Quotation Errors in High-Impact-Factor Orthopaedic and Sports Medicine Journals." *JBJS Open Access* 6 (3). <https://doi.org/10.2106/JBJS.OA.21.00019>.
- Harris, A. B., G. Benes, D. Ghanem, M. Cartagena-Reyes, A. Jain, and D. M. Laporte. 2024. "Using a Modern Linked Research Database to Examine Gender Disparities in Orthopaedic Grant Funding from 2010 to 2022." *Journal of Bone and Joint Surgery* 106 (1): 39–46. <https://doi.org/10.2106/JBJS.23.00358>.
- Howard, R., and M. Englesbe. 2022. "Citation Accuracy vs Citation Integrity." *JAMA Surg* 157 (4): 361. <https://doi.org/10.1001/jamasurg.2021.6552>.
- Kacena, M. A., L. I. Plotkin, and J. C. Fehrenbacher. 2024. "The Use of Artificial Intelligence in Writing Scientific Review Articles." *Curr Osteoporos Rep* 22 (1): 115–21. <https://doi.org/10.1007/s11914-023-00852-0>.
- Larson, A. N., M. C. Marks, J. M. Gonzalez Sepulveda, et al. 2024. "Non-Fusion Versus Fusion Surgery in Pediatric Idiopathic Scoliosis: What Trade-Offs in Outcomes Are Acceptable for the Patient and Family?" *Journal of Bone and Joint Surgery* 106 (1): 2–9. <https://doi.org/10.2106/JBJS.23.00503>.
- Luo, M., C. C. Li, D. Molina, C. R. Andersen, and V. K. Panchbhavi. 2013. "Accuracy of Citation and Quotation in Foot and Ankle Surgery Journals." *Foot Ankle Int* 34 (7): 949–55. <https://doi.org/10.1177/1071100713475354>.
- MacKechnie, M. C., D. W. Shearer, M. H. J. Verhofstad, et al. 2024. "Establishing Consensus on Essential Resources for Musculoskeletal Trauma Care Worldwide: A Modified Delphi Study." *Journal of Bone and Joint Surgery* 106 (1): 47–55. <https://doi.org/10.2106/JBJS.23.00387>.
- Mogull, S. A. 2017. "Accuracy of Cited 'Facts' in Medical Research Articles: A Review of Study Methodology and Recalculation of Quotation Error Rate." Edited by N. R. Smalheiser. *PLoS ONE* 12 (9): e0184727. <https://doi.org/10.1016/j.nicl.2023.103496>.
- Montenegro, T. S., K. Hines, P. P. Partyka, and J. Harrop. 2021. "Reference Accuracy in Spine Surgery." *Journal of Neurosurgery: Spine* 34 (1): 22–26. <https://doi.org/10.3171/2020.6.SPINE20640>.
- OpenAI. 2025. "ChatGPT." 2025. <https://openai.com>.
- "OpenEvidence." 2025. 2025. <https://www.openevidence.com>.
- Richardson, M. K., K. C. Liu, C. K. Mayfield, N. M. Kistler, J. R. Lieberman, and N. D. Heckmann. 2024. "Tranexamic Acid Is Safe in Patients with a History of Venous Thromboembolism Undergoing Total Joint Arthroplasty." *Journal of Bone and Joint Surgery* 106 (1): 30–38. <https://doi.org/10.2106/JBJS.23.00254>.
- Singh, V., J. E. Jolissaint, J. G. Kohler, et al. 2024. "Precision or Pitfall? Evaluating the Accuracy of ICD-10 Coding for Cemented Total Hip Arthroplasty: A Multicenter Study." *Journal of Bone and Joint Surgery* 106 (1): 56–61. <https://doi.org/10.2106/JBJS.23.00325>.
- Skolasky, R. L., A. M. Cizik, A. Jain, and B. J. Neuman. 2024. "Measuring Value in Spine Care Using the PROMIS-Preference Scoring System." *Journal of Bone and Joint Surgery* 106 (1): 21–29. <https://doi.org/10.2106/JBJS.23.00113>.