


# The Transparency Paradox: Why Researchers Avoid Disclosing AI Assistance in Scientific Writing

Ahmed S BaHamam <sup>1,2</sup>

<sup>1</sup>The University Sleep Disorders Center, Department of Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia; <sup>2</sup>King Saud University Medical City, Riyadh, Saudi Arabia

Correspondence: Ahmed S BaHamam, The University Sleep Disorders Center, Department of Medicine, College of Medicine, King Saud University, Box 225503, Riyadh, 11324, Saudi Arabia, Email [ashammam2@gmail.com](mailto:ashammam2@gmail.com)

The integration of artificial intelligence (AI) tools in scientific writing has accelerated dramatically; yet a concerning transparency gap persists.<sup>1</sup> While current data indicate that researchers are increasingly adopting AI assistance for manuscript preparation, disclosure rates remain surprisingly low, despite evolving publisher requirements.<sup>2</sup> This phenomenon reflects deeper tensions between technological advancement and traditional academic evaluation systems, creating what may be termed a “transparency paradox” in scholarly publishing, where the tools most capable of enhancing research productivity also pose the greatest risks to scientific accountability when used without disclosure.

The imperative for transparency in AI-assisted scientific writing has become a defining challenge for contemporary academic integrity. Scientific progress fundamentally depends on reproducibility and methodological transparency. However, the rapid integration of AI tools into research workflows poses a threat to these foundational principles.<sup>3</sup> As generative AI capabilities advance beyond simple grammar correction to substantive content generation, the stakes of disclosure failures escalate dramatically. When sophisticated language models contribute to literature synthesis, hypothesis development, or manuscript drafting, undisclosed usage compromises the ability of readers to evaluate the validity and reproducibility of research.<sup>4</sup> The problem intensifies because modern generative AI systems exhibit increasing sophistication alongside decreasing predictability, making human oversight more challenging precisely when it becomes most essential.<sup>5</sup>

This editorial examines the underlying factors contributing to the non-disclosure of AI assistance in scientific writing and proposes a practical tiered transparency framework to address legitimate researcher concerns while maintaining scholarly integrity.

## The Evidence of Widespread but Hidden AI Use

AI tools have rapidly transformed academic writing practices; however, usage patterns demonstrate significant gaps between adoption and disclosure. Recent analysis of AI conference peer reviews detected substantial language model modifications in 6.5% to 16.9% of submissions, indicating widespread integration beyond simple grammar correction.<sup>5</sup> The detection methods revealed that certain AI-generated phrases like “commendable”, “meticulous”, and “intricate” showed dramatic frequency increases, with some adjectives appearing 9.8 to 34.7 times more often in recent reviews compared to historical baselines, suggesting systematic but undisclosed AI integration across academic venues.<sup>5</sup> Contemporary research indicates the widespread adoption of AI tools across academic disciplines. Nevertheless, formal disclosure remains sporadic, creating what researchers describe as a “hidden ecosystem” of AI assistance.<sup>6</sup> Adding to this evidence, a bibliometric review of 1,998 radiology manuscripts published in Elsevier journals found that only 34 papers (1.7%) acknowledged any involvement of LLMs (large language models) - AI systems trained on massive text datasets to process and generate human-like language, primarily for minor language edits.<sup>7</sup> No upward trend was observed in 2024, and a clear predominance of disclosures was from non-English-speaking institutions. This pattern suggests that while researchers recognize the utility of AI assistance, they may hesitate to acknowledge its use publicly. This lack of acknowledgement could be due to multiple interconnected factors, including concerns about academic legitimacy, fear of bias from editors and reviewers, and uncertainty about appropriate disclosure standards.<sup>8</sup>

## The Productivity-Disclosure Dilemma

Experimental evidence suggests that generative AI can enhance professional writing productivity by 40% in terms of time savings and improve quality ratings by 18%.<sup>9</sup> However, these productivity gains come alongside concerning usage patterns that may discourage disclosure.<sup>9</sup> Research shows that 33% of professionals submitted AI output without editing, while 53% made only superficial modifications lasting 3.3 minutes on average, suggesting that writers essentially replace their own work rather than using AI collaboratively.<sup>9</sup> This complete substitution of human effort may create reluctance to disclose AI usage, as researchers might feel that such dependence undermines their perceptions of scholarly contributions. A study found that workers who benefited most from AI assistance were those who initially performed poorly, indicating that AI may serve as a compensatory rather than an enhancement mechanism.<sup>9</sup>

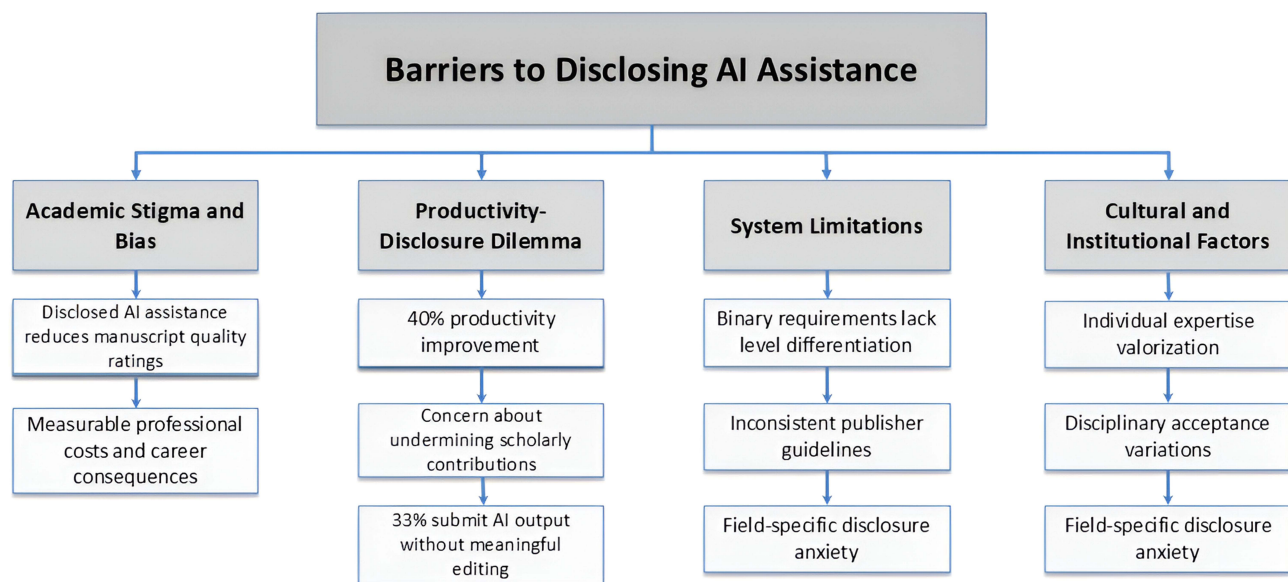
More troubling is the recognition that sophisticated AI models may become less reliable in ways that human supervisors cannot easily detect, making transparent disclosure increasingly necessary for maintaining scientific integrity.<sup>10</sup> This creates a fundamental paradox: while AI assistance becomes more powerful and prevalent, disclosure practices remain inconsistent, creating an uneven playing field where some researchers quietly benefit from undisclosed AI support while others transparently acknowledge their usage. The problem deepens because these systems produce increasingly sophisticated outputs that may contain subtle errors or biases undetectable through casual human review, turning disclosure from an ethical choice into a practical necessity for maintaining research quality.<sup>11</sup>

## The Academic Stigma Against AI Disclosure

Empirical research validates that legitimate concerns about academic acceptance compound these disclosure challenges. Li et al experimental investigation demonstrated that revealing AI assistance systematically reduces perceived manuscript quality, with readers rating AI-assisted argumentative essays 0.32 points lower (on a 5-point scale) than identical content presented without disclosure.<sup>12</sup> The stigma appears particularly pronounced when AI contributes substantial content generation rather than merely providing editing support, with disclosed AI-generated drafts suffering disproportionate quality penalties compared to AI-edited manuscripts.<sup>12</sup> These findings suggest rational foundations for the reluctance of researchers to disclose, as transparency may genuinely disadvantage their work during evaluation processes. Current evidence suggests that transparency in AI use remains limited, partly due to ethical concerns and cultural resistance.<sup>13,14</sup> Rajpurohit and Dobhada highlight key issues such as bias, plagiarism, and accountability, noting that a lack of clear standards complicates responsible disclosure.<sup>14</sup> Nensa further observes that reluctance to disclose is particularly rooted in traditional academic environments where originality and individual expertise are highly valued, a dynamic that may create a form of “disclosure anxiety” affecting reporting practices across disciplines.<sup>13</sup>

A study revealed that individuals with higher writing confidence showed significantly greater bias against disclosed AI assistance.<sup>12</sup> High-confidence evaluators consistently downgraded AI-disclosed work, while those with less writing confidence barely changed their ratings.<sup>12</sup> This pattern suggests potential disciplinary variations, where fields emphasizing traditional writing skills may exhibit stronger stigmatization of AI assistance. Such evidence indicates that under-disclosure may represent a strategic adaptation to biased evaluation systems rather than simply an ethical lapse. However, this creates a concerning feedback loop: the fear of stigmatization perpetuates non-disclosure, which maintains uninformed, negative attitudes toward AI assistance, ultimately undermining both transparency and the fair evaluation of scholarly contributions. Furthermore, the disclosed AI assistance reduced the likelihood of manuscripts achieving top rankings by significant margins, demonstrating that transparency carries measurable professional costs.<sup>12</sup> The evidence points to four interconnected categories of deterrents (Figure 1). These obstacles collectively explain why researchers avoid transparency despite growing publisher requirements.

Institutional and disciplinary cultures significantly influence disclosure behavior, with fields that maintain strong traditions of methodological transparency showing a greater acceptance of AI disclosure, while others view such acknowledgment as problematic.<sup>4</sup> The transparency gap between AI usage and disclosure practices has become particularly evident in recent empirical investigations. A comprehensive survey of 5,229 researchers revealed that while 28% acknowledged using AI for manuscript editing, the majority failed to disclose this assistance when submitting their work.<sup>2</sup> The survey exposed striking variations in disclosure attitudes across different applications. Although 90%



**Figure 1** Barriers to AI disclosure in scientific writing. Hierarchical flowchart depicting four main barrier categories branching from a central concept. Each category contains supporting evidence.

considered AI editing ethically acceptable, researchers disagreed about disclosure requirements, with only 35% supporting the use without disclosure and 55% favoring some form of acknowledgment.<sup>2</sup> More concerning, researchers showed significant confusion about what constitutes disclosure-worthy AI assistance, with many treating sophisticated language model interactions as equivalent to basic grammar checking. This pattern extended across multiple AI applications, with approximately 8% reporting first draft generation, and similar rates for translation services; yet, disclosure remained systematically underreported. The survey found that when researchers did employ AI tools, “they more often than not said they had not disclosed it at the time”.<sup>2</sup> Perhaps more telling, 65% of respondents claimed never to have used AI in any academic scenario, a figure that appears incongruent with objective detection studies.<sup>2</sup> This disconnect suggests either a widespread misunderstanding of what constitutes AI usage or deliberate concealment, both of which threaten the transparency principles essential for scientific reproducibility.<sup>15</sup> These disclosure patterns reflect deeper concerns about how AI-assisted work is perceived and evaluated.

## Limitations of Current Disclosure Systems

The underdisclosure phenomenon may stem from several factors: uncertainty about what constitutes disclosure-worthy AI use, concerns about editorial bias against AI-assisted manuscripts, or simple oversight in increasingly routine AI interactions. However, the persistence of this gap represents more than a compliance issue. It creates an asymmetric landscape where transparent researchers potentially face disadvantages compared to those who benefit from undisclosed AI support, ultimately undermining the collaborative nature of scientific discourse. Moreover, current disclosure requirements across academic journals show significant inconsistencies and limitations that may inadvertently discourage transparency or create confusion among authors. This disconnect suggests that existing binary disclosure systems, requiring authors to simply state whether AI was used, lack the specificity necessary for meaningful transparency. Moreover, current detection methods remain unreliable, with false-positive rates that can potentially damage authors’ reputations.<sup>16,17</sup>

The reluctance to disclose AI use becomes more understandable when considering the current limitations of detection technologies. The fundamental limitation lies in treating all AI assistance as equivalent. Current guidelines typically require disclosure of any AI use beyond basic grammar checking, but fail to differentiate between substantial content generation and minor editing assistance. For instance, an author using generative AI to refine sentence clarity faces the same disclosure requirements as one generating entire sections from prompts. This approach creates several problems:

authors may avoid helpful editing tools to circumvent the stigma of disclosure, reviewers cannot assess the appropriate level of AI contribution, and readers lack context for evaluating the human intellectual input. Additionally, the study on AI-modified content at scale revealed that disclosure requirements vary dramatically across publishers, with some requiring detailed prompt disclosure, while others accept simple acknowledgments.<sup>5</sup> Recent evidence suggests that non-disclosure often stems from uncertainty rather than deliberate concealment.

## A Tiered Transparency Framework

To address these limitations, we propose a tiered disclosure framework that categorizes assistance into four distinct levels.

Current publisher guidelines from the ICMJE (International Committee of Medical Journal Editors) and major publishers require the disclosure of generative AI use in manuscript preparation, including the tool used, its purpose, and oversight responsibilities.<sup>18</sup> While these policies maintain full author accountability, they lack a structured framework for distinguishing between different levels of AI assistance. Most guidelines treat all AI use beyond basic grammar checking as equivalent, failing to differentiate between minor language enhancement and substantial content generation.

Our proposed tiered framework addresses this limitation by categorizing AI assistance according to intellectual contribution (Table 1):

**Level 1 (No Disclosure):** Basic technical assistance equivalent to conventional editing tools, proofreading, grammar correction, translation, and formatting, that present minimal integrity risks.

**Level 2 (Simple Acknowledgment):** Linguistic enhancement preserving original intellectual contributions, editing for clarity, flow, and style, requiring only brief acknowledgment: “AI tools assisted with language enhancement.”

**Level 3 (Detailed Disclosure):** Substantial content involvement, section drafting, brainstorming, literature synthesis, requiring comprehensive reporting of AI model, oversight procedures, and contribution ratios.

**Level 4 (Comprehensive Disclosure):** Primary material generation encompasses AI-assisted creation of core intellectual contributions, including hypothesis formulation, research question development, theoretical framework construction, and substantial interpretation of results or conclusions that constitute the manuscript’s central scholarly contribution. This level requires rigorous documentation of prompts, model versions, verification methods, and explicit delineation of human contributions. A recent example demonstrates that such comprehensive disclosure is compatible with high-impact research, as shown by Virtual Lab studies that transparently documented AI agent contributions to nanobody design while achieving experimental validation.<sup>19</sup>

This approach establishes clear boundaries, reducing disclosure uncertainty while enabling appropriate assessment of human intellectual input. Implementation requires collaboration between publishers and developers to create standardized assessment tools and incorporate measures for detecting undisclosed use, while protecting authors from false accusations.

**Table 1** Proposed Four-Level AI Disclosure Framework for Scientific Writing

Level	Disclosure Type	AI Assistance Category	Requirements	Example Statement
Level 1	No Disclosure	Basic technical assistance equivalent to conventional editing tools	Proofreading, grammar correction, translation, and formatting that present minimal integrity risks	Not required.
Level 2	Simple Acknowledgment	Linguistic enhancement preserving original intellectual contributions	Editing for clarity, flow, and style	AI tools assisted with language enhancement.
Level 3	Detailed Disclosure	Substantial content involvement	Section drafting, brainstorming, and literature synthesis	GAI (model) assisted with literature synthesis and initial drafting (estimated 30%). All content was verified and refined by the author.
Level 4	Comprehensive Disclosure	Primary material generation	Extensive content creation requiring human oversight	GAI (model) provided substantial assistance with hypothesis, formulation, framework development, and analysis. Full documentation available in supplementary materials.

Importantly, author accountability remains the cornerstone principle across all levels, ensuring that researchers remain fully responsible for the accuracy, integrity, and scholarly merit of their work, regardless of the assistance received.

In parallel, it has been emphasized that generative AI tools, much like statistical software before them, should be treated as instruments that enhance rather than replace human expertise.<sup>13</sup> It is argued that embracing transparency while maintaining rigorous oversight ensures AI integration advances scientific integrity rather than undermines it.

In summary, the transparency paradox surrounding AI disclosure in scientific writing demands immediate attention from the academic community. It appears that current approaches, which focus solely on mandatory disclosure, overlook legitimate concerns of researchers about evaluation bias and career consequences. The proposed tiered framework offers a pragmatic path forward, acknowledging varying levels of AI assistance while maintaining scholarly integrity. As AI tools become increasingly sophisticated and ubiquitous, the academic community must move beyond simplistic binary disclosure requirements toward tiered transparency standards that support innovation while preserving trust in scientific literature. Future research should prioritize understanding the psychological/cultural and institutional barriers that prevent researchers from disclosing their use of writing assistance, including academic stigma, career concerns, peer pressure, and institutional cultures that discourage transparency. Investigations examining how evaluator bias varies across disciplines, career stages, and cultural contexts will be essential for developing targeted interventions to normalize disclosure practices and reduce professional penalties associated with transparency. Complementing this foundational research, empirical studies should evaluate the effectiveness of structured disclosure systems across different academic disciplines, measuring their impact on research quality assessment and author compliance rates, while implementation research focusing on practical adoption strategies, author training programs, and publisher policies will be vital for translating these insights into meaningful improvements in scientific transparency.

## Data Sharing Statement

Data sharing is not applicable as no new data was generated for this work.

## Author Contributions

Ahmed S. BaHammmam: Conceptualization, Writing – original draft preparation, Writing – review and editing, and Supervision.

The author gives final approval of the version to be published; has agreed on the journal to which the article has been submitted; and has agreed to be accountable for all aspects of the work.

## Funding

The Strategic Technologies Program of the National Plan for Sciences and Technology and Innovation in the Kingdom of Saudi Arabia, Riyadh, Saudi Arabia (MED511-02-08).

## Disclosure

The author reports no conflicts of interest in this work. Grammarly assisted with grammar correction during the preparation of this paper.

---

## References

1. Salman HA, Ahmad MA, Ibrahim R, Mahmood J. Systematic analysis of generative AI tools integration in academic research and peer review. *Online J Commun Media Technol.* 2025;15(1):e202502. doi:10.30935/ojcm/15832
2. Kwon D. Is it OK for AI to write science papers? Nature survey shows researchers are split. *Nature.* 2025;641(8063):574–578. doi:10.1038/d41586-025-01463-8
3. Bahammam AS, Trabelsi K, Pandi-Perumal SR, Jahrami H. Adapting to the Impact of Artificial Intelligence in Scientific Writing: balancing Benefits and Drawbacks while Developing Policies and Regulations. *J Nat Sci Med.* 2023;6(3):152–158. doi:10.4103/jnsm.jnsm\_89\_23
4. BaHammmam AS. Peer Review in the Artificial Intelligence Era: a Call for Developing Responsible Integration Guidelines. *Nat Sci Sleep.* 2025;17:159–164. doi:10.2147/NSS.S513872
5. Liang W, Izzo Z, Zhang Y, et al. Monitoring AI-modified content at scale: a case study on the impact of ChatGPT on AI conference peer reviews. Proceedings of the 41st International Conference on Machine Learning. Vienna, Austria. 2024.

6. Fiorillo L. Confronting the demonization of AI writing: reevaluating its role in upholding scientific integrity. *Oral Oncol Rep.* 2024;12:100685. doi:10.1016/j.oor.2024.100685
7. Jonah Barrett D, Heng R, Perchik JD. Documenting Disclosure: limited Reporting of Generative AI Usage in Radiology Research Manuscripts. *Acad Radiol.* 2025. doi:10.1016/j.acra.2025.06.057
8. Vitente A, Lazaro R, Escuadra C, Regino J, Rotor E. Editorial: the Use of Artificial Intelligence (AI)-Assisted Technologies in Scientific Discourse. *Philippine J Phys Ther.* 2023;2(1):1–3. doi:10.46409/002.HNUY6271
9. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. *Science.* 2023;381(6654):187–192. doi:10.1126/science.adh2586
10. Gopali S, Siami-Namini S, Abri F, Namin AS. The performance of the LSTM-based code generated by Large Language Models (LLMs) in forecasting time series data. *Nat Language Processing J.* 2024;9:100120. doi:10.1016/j.nlp.2024.100120
11. Hosseini M, Horbach SPJM. Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res Integrity Peer Rev.* 2023;8(1):4. doi:10.1186/s41073-023-00133-5
12. Li Z, Liang C, Peng J, Yin M. *How Does the Disclosure of AI Assistance Affect the Perceptions of Writing?* Miami, Florida, USA; 2024.
13. Nensa F. Embracing generative AI: a necessary evolution in professional writing. *Eur J Radiol Artif Intell.* 2025;1:100001. doi:10.1016/j.ejrai.2024.100001
14. Rajpurohit L, Dobhada S. Scientific writing and its ethical considerations using AI tools. *Oral Oncol Rep.* 2024;9:100196. doi:10.1016/j.oor.2024.100196
15. Terry PE. I Used AI in the Drafting of this Editorial. How Should I Reference AI's Contribution? *Am J Health Promot.* 2025;39(4):561–563.
16. Giray L, Sevnarayan K, Ranjbaran Madiseh F. Beyond Policing: AI Writing Detection Tools, Trust, Academic Integrity, and Their Implications for College Writing. *Int Reference Serv Quart.* 2025;29(1):83–116. doi:10.1080/10875301.2024.2437174
17. Bordalejo B, Pafumi D, Onuh F, Khalid AKMI, Pearce MS, O'Donnell DP. "Scarlet Cloak and the Forest Adventure": a preliminary study of the impact of AI on commonly used writing tools. *Int J Educ Technol Higher Educ.* 2025;22(1):6. doi:10.1186/s41239-025-00505-5
18. International Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. 2025. Available from: <https://www.icmje.org/icmje-recommendations.pdf>. Accessed Sep 26, 2025.
19. Swanson K, Wu W, Bulaong NL, Pak JE, Zou J. The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies. *Nature.* 2025. doi:10.1038/s41586-025-09442-9

## Nature and Science of Sleep

### Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>

**Dovepress**  
Taylor & Francis Group