Editorial

Giuseppe Lippi*, Karl J. Lackner, Bohuslav Melichar, Shiyang Pan, Peter Schlattmann, Ronda Greaves, Philippe Gillery and Mario Plebani

Challenging the dogma: why reviewers should be allowed to use AI tools

https://doi.org/10.1515/cclm-2025-1180

Keywords: artificial intelligence; peer-review; scientific writing

Significant advancements in artificial intelligence (AI) have facilitated the development of a vast array of AI-driven tools for research and academic writing over the past decade. The emergence of these innovative and disruptive technologies has sparked considerable interest within the scientific community, catalyzing the rapid expansion of potential applications [1]. Concurrently, their integration into academic workflows has triggered ongoing discussions regarding the appropriate scope, ethical considerations, and methodologies for their deployment [2].

In December 2022, *Nature* published the first article raising concerns about the use of AI tools, specifically ChatGPT and other forms of generative artificial intelligence (GAI) in academic writing [3]. Since then, an increasing number of journals and publishers have updated their editorial policies and author guidelines, aiming to establish clearer protocols for the disclosure of GAI usage in scholarly work. However, as noted by Ganjavi and colleagues [4], the lack of standardized policies, coupled with the current variability in editorial guidance, has given rise to substantial concerns. Importantly, the implications of GAI technologies extend beyond manuscript preparation, with significant potential effects on the peer-review process and other aspects of scholarly publishing.

Peer-review has long been regarded as the cornerstone of scientific integrity, yet it is increasingly challenged by increasing submission volumes and falling reviewer availability, both in *Clinical Chemistry and Laboratory Medicine* and across the wider landscape of scientific publishing [5]. To this end, the recent letter by Anna Carobene raises pressing issues surrounding the use of AI in peer-review [6]. While her balanced assessment deserves recognition for clarifying policy considerations, several of its premises require deeper analysis. Specifically, the insistence on restricting reviewers' use of AI rests on questionable distinctions that, if left unchallenged, risk undermining both the efficiency and the quality of peer-review.

The central contention here is straightforward: there is no essential reason to prohibit reviewers from using AI tools to generate preliminary impressions of a manuscript's strengths and weaknesses. Indeed, many scientists already use such tools in their own writing process before submission, treating AI as a "first-pass filter" to improve clarity, highlight inconsistencies, or catch obvious errors in their manuscripts. Extending this practice to reviewers is not a violation of scholarly norms, but a logical continuation, while disallowing it creates an inconsistency that may be counterproductive, impossible to govern and even illogical, for reasonable reasons. Instead, what we should be doing is embracing AI under controlled limits (to be clearly defined) (Figure 1).

One of the recurring themes in Carobene's letter is the asymmetry between authors and reviewers [6]. While authors may use AI under conditions of disclosure, reviewers are constrained by stricter prohibitions, often justified by concerns over confidentiality. The underlying assumption seems to be that authorship is a creative activity where AI can be tolerated, whereas reviewing is a critical act where AI

***Corresponding author: Giuseppe Lippi**, Section of Clinical Biochemistry, University of Verona, Ospedale Policlinico G.B. Rossi, Piazzale LA Scuro 10, 34134, Verona, Italy, E-mail: giuseppe.lippi@univr.it. https://orcid.org/0000-0001-9523-9054

Karl J. Lackner, Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

Bohuslav Melichar, Department of Oncology, Palacky University Medical School and University Hospital, Olomouc, Czech Republic

Shiyang Pan, Department of Laboratory Medicine, Affiliated Hospital of Nanjing Medical University, Nanjing, China

Peter Schlattmann, Institute for Medical Statistics, Computer and Data Sciences (IMSID), University Hospital Jena, Jena, Germany

Ronda Greaves, Biochemical Genetics, Victorian Clinical Genetics Services Ltd, Bundoora, VIC, Australia

Philippe Gillery, Laboratory of Pediatric Biology and Research, University Hospital of Reims, Reims, France

Mario Plebani, Medicine-DIMED, University of Padova, Padova, Italy. https://orcid.org/0000-0002-0270-1711

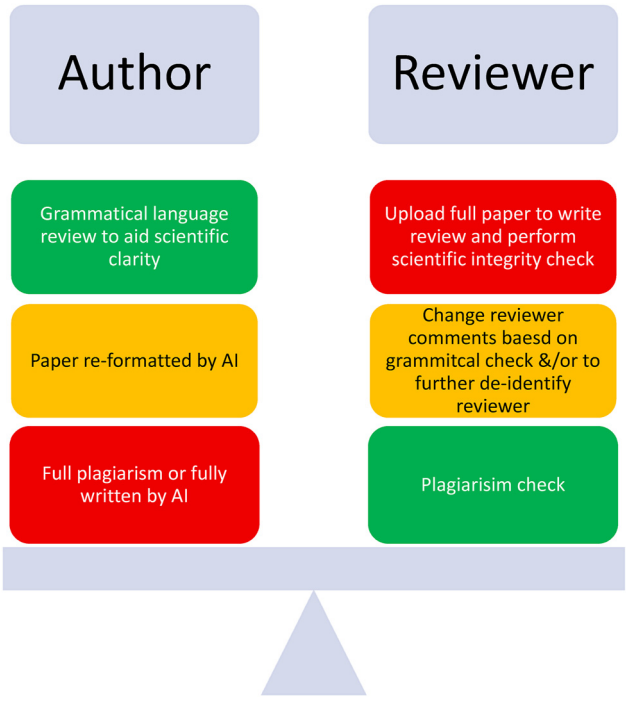

Figure 1: Defining the acceptable influence of artificial intelligence (AI) should be the goal. In this traffic light diagram, examples of what we are asking AI to do for both authors and reviewers are provided, and this is compared to how we accept AI for authors.

poses unacceptable risks. Nonetheless, closer examination reveals that this apparent distinction diminishes under critical scrutiny. Both authorship and peer-review involve the structured evaluation of information, the synthesis of insights, and the clear articulation of findings. In both circumstances, AI has utility as a supportive tool that helps identifying overlooked issues, improve precision of language, and accelerate otherwise tedious tasks. To deny reviewers access to such tools while allowing authors to use them seems illogical, and ultimately detrimental to the quality of peer-review itself.

A legitimate concern, however, is the protection of confidentiality when manuscripts are processed through public AI models, where data may be used to train future models or potentially be misappropriated for unintended or malicious purposes. This risk should not be underestimated, as uploading unpublished manuscripts to open platforms may constitute a violation of the confidentiality agreement between journals and authors. It requires reviewers to disclose any instances in which AI has been used as an auxiliary tool in the course of their assessment. Yet, the existence of risk does not justify a total ban. The history of peer-review shows that ethical obligations are never enforced by banning all possible misuses, but by setting norms and infrastructures that enable responsible practice. Just as reviewers are entrusted with manuscripts under the understanding that they must not circulate them to third parties, in the same way the use of AI can be regulated through clear boundaries. The solution lies in secure infrastructures, like publisher-provided AI tools embedded within submission platforms, which will ensure that manuscript data never leaves controlled environments. For example, the first and corresponding author of this editorial uploaded in ChatGPT the PDF of a personal article under review in a scientific journal, asking the AI how the confidentiality of the data in the file could be safeguarded. ChatGPT responded: "*The confidentiality of the data in your uploaded PDF is protected because it is only used within this conversation, not shared with other users, and not used to train OpenAI's models. Your file is stored temporarily in a secure environment, encrypted in transit and at rest, and automatically deleted after the session unless you enable memory in your settings. Only you and the model in this chat can access its contents, and any outputs or analyses are generated solely at your request, ensuring the information remains private and secure*". Based on this example, it seems that current implementations provide reasonable safeguards for confidentiality, as long as reviewers do not upload manuscripts to open/public instances. Moreover, many practical applications of AI in manuscript review (e.g., analyzing notes, applying methodological checklists, or summarizing key arguments) can be performed without disclosing the full text of the manuscript and thus carry minimal confidentiality risk. Consequently, characterizing all potential applications of AI in peer-review as violations of trust constitutes an oversimplification and fails to capture the heterogeneity of potential practices.

The issue of transparency deserves similar scrutiny. In principle, disclosure of AI use is desirable. But in practice, current policies mandating disclosure while simultaneously prohibiting AI create a perverse dynamic. Reviewers who perceive legitimate value in AI are faced with a dilemma, i.e., either refraining from its use, and thus potentially compromising the quality of their reports, or employing it clandestinely, without transparent disclosure. This is not a trivial issue. When journal policies are misaligned with the realities of scientific practice, they foster widespread non-compliance and erode trust. A more effective approach would not entail the intensification of prohibitive measures, but rather aligning policy with current practice. That is, permitting AI under clear guidelines, requiring disclosure and placing the burden of accountability on reviewers to critically evaluate its outputs. In this respect, parity with authors is not optional but essential.

Another frequent objection is that AI might erode reviewer accountability. Large language models (LLMs) are

indeed prone to factual inaccuracies and even to so-called "hallucinations" [7]. Nevertheless, this vulnerability does not absolve reviewers of responsibility, but it just heightens it. Just as the use of statistical software does not excuse a scientist from misinterpreting results, the use of AI does not excuse a reviewer from flawed judgments. Accountability lies in application of expertise, not in the choice of tools. Denying the use of AI on the basis of accountability implies a double standard: it presumes that reviewers forfeit responsibility when assisted by AI, while authors are presumed to retain it under analogous circumstances. An important issue that may arise with widespread use of AI in peer review is the potential for "homogenized" evaluations. Excessive reliance on AI could reduce the diversity of reviewer perspectives, potentially reducing the richness and quality that characterize high-quality peer-review.

Carobene's letter also raises the issue of reviewer compensation, noting its potential benefits, but warning of distorted incentives [6]. These warnings are well-founded. Financial compensation risks attracting reviewers motivated primarily by reward rather than by scholarly duty, while also being economically unsustainable for many publishers and journals, especially those smaller or society-led. Yet, this is precisely why AI has such transformative potential. It must be intended not as a replacement for human judgment, but as a mechanism to make reviewing more efficient and less burdensome. By automating routine tasks such as checking adherence to reporting standards, identifying missing references, or summarizing the manuscript's structure, AI frees reviewers to concentrate on the highest-value aspects (i.e., assessing methodological rigor, conceptual innovation, and interpretive soundness).

The process of peer-review has never been static, but has adapted to each wave of technological progress, from handwritten reports to typewriters, from postal submission to electronic platforms, from simple plagiarism checks to sophisticated similarity-detection software. AI represents the next stage in this evolution. Opposing its use out of apprehension replicates the anxieties historically directed at computers and automated statistical tools. What seemed threatening then has now become indispensable. The same may be true of AI in peer-review. Rather than entrenching outdated prohibitions, we should embrace AI responsibly, regulate it wisely, and recognize it for what it is: not a threat to scholarly standards, but a necessary ally in sustaining them.

As AI continues to advance rapidly, the central question is no longer whether it will be integrated into peer review, but rather how its use can be effectively managed and harnessed to ensure fairness to authors while enhancing the efficiency and value of the editorial process. Our view is increasingly echoed by the editorial boards of other major scientific journals, including JAMA. In their August 28, 2025 comment entitled "Artificial Intelligence in Peer Review" [8], the editors of *JAMA* reported early efforts to consider how AI could be used to support peer review in ways that go beyond simply speeding up the process, emphasizing the potential of these tools to assist reviewers, reduce bias and improve the transparency of editorial decisions, while at the same time warning against uncritical reliance on algorithms. The focus was placed on ensuring that any use of AI preserves fairness, integrity, and scientific rigor. This measured approach suggests that major journals now view AI not only as a technical aid, but also as part of a broader effort to reinforce trust and uphold quality in the peer-review process.

Table 1: Policy framework for the use of artificial intelligence (AI) in the peer-review of scientific articles.

Principle	Description
Secure infrastructure	Journals should provide their own AI tools, ensuring manuscripts are never exposed to external sources.
Flexible disclosure	Reviewers should not be penalized for using AI, but should be invited to disclose briefly how it was applied (e.g., "AI used to check grammar and summarize reviewer notes").
Human primacy	Reviewers must affirm that all judgments remain their own, with AI serving only as an assistant.
Parity with authors	Reviewer policies should mirror author policies, closing the asymmetry that currently undermines trust.

AI, artificial intelligence.

Taken together, these considerations lead to a simple conclusion. AI should be permitted in peer-review, not clandestinely, but openly, under safeguards that protect confidentiality and preserve human judgment. The policy framework proposed by Carobene [6], emphasizing confidentiality, transparency, and scope limits, remains valuable, but must be refined. A workable framework should include four key principles, i.e., secure infrastructure, flexible disclosure, human primacy, and parity with authors, as detailed in Table 1.

Research ethics: Not applicable.
Informed consent: Not applicable.
Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.
Use of Large Language Models, AI and Machine Learning Tools: The authors wish to disclose that ChatGPT 3.5 was used for enhancing the clarity and coherence of the manuscript writing. The tool was only used for language refinement purposes, ensuring the text was clear and coherent without altering the scientific content or generating any new text.

Conflict of interest: The authors state no conflict of interest.
Research funding: None declared.
Data availability: Not applicable.

References

1. Khan MK, Ferdous J, Mourshed G, Hossain SB. Use of artificial intelligence in scientific writing. Mymensingh Med J 2025;34:592–7.
2. Cheng A, Calhoun A, Reedy G. Artificial intelligence-assisted academic writing: recommendations for ethical use. Adv Simul (Lond) 2025;10:22.
3. Stokel-Walker C. AI bot ChatGPT writes smart essays – should professors worry? Nature 2022. https://doi.org/10.1038/d41586-022-04397-7 [Epub ahead of print].
4. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, et al. Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis. BMJ 2024;384:e077192.
5. Lippi G. How do I peer-review a scientific article? A personal perspective. Ann Transl Med 2018;6:68.
6. Carobene A. Permitting disclosed AI assistance in peer review: parity, confidentiality, and recognition. Clin Chem Lab Med 2025;63: e255–7.
7. Hatem R, Simmons B, Thornton JE. A call to address AI "hallucinations" and how healthcare professionals can mitigate their risks. Cureus 2023; 15:e44720.
8. Perlis RH, Christakis DA, Bressler NM, Öngür D, Kendall-Taylor J, Flanagin A, et al. Artificial intelligence in peer review. JAMA 2025;28. https://doi.org/10.1001/jama.2025.15827 [Epub ahead of print].