Iranian Journal of Basic Medical Sciences
ijbms.mums.ac.ir

Editorial

The credibility of scientific publication in the era of generative Artificial Intelligence (AI): Who is responsible for scientific integrity?

Leila Arabi [1], Ali Roohbakhsh [2], Bizhan Malaekeh-Nikouei [1], Bibi Sedigheh Fazly Bazzaz [3]*

[1] Nanotechnology Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran (Assistant Editor)
[2] Pharmaceutical Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran (Assistant Editor)
[3] Biotechnology Research Center, Pharmaceutical Technology Institute, Mashhad University of Medical Sciences, Mashhad, Iran (Editor–in–Chief)

►Please cite this article as:
Arabi L, Roohbakhsh A, Malaekeh-Nikouei B, Fazly Bazzaz BS. The credibility of scientific publication in the era of generative Artificial Intelligence (AI): Who is responsible for scientific integrity?. Iran J Basic Med Sci 2027; 30: 1-3. doi: https://dx.doi.org/10.22038/ijbms.2027.28729

Introduction

A new challenge to scientific knowledge

Scientific review articles are essential for advancing knowledge because they critically evaluate and interpret existing evidence. Their value extends beyond summarizing published findings; they also help identify knowledge gaps, resolve inconsistencies, and generate meaningful scientific insights. However, the rapid growth of generative artificial intelligence (AI), especially large language models (LLMs), raises concerns about the credibility and intellectual value of AI-assisted scientific reviews. In previous editorials in the Iranian Journal of Basic Medical Sciences (IJBMS), we explored AI's impact on academic writing and publication (1), how authors and reviewers can use AI legitimately, and how it can help detect plagiarism and data fabrication (2).

In this editorial, we discuss the credibility of scientific publications, particularly review articles, in the generative AI era. AI tools can quickly produce narrative reviews with coherent discussions and polished language. Although this improves research efficiency, it blurs the distinction between genuine scientific synthesis and automated text. The main concern is not inaccuracy, but that superficial manuscripts can appear credible and authoritative.

AI-generated reviews may reproduce existing information without adequately evaluating its validity. LLMs can generate reasonable but incorrect statements, misinterpret findings, and fabricate references—a phenomenon commonly described as AI hallucination (3). These limitations are particularly concerning for systematic reviews, where transparent literature searches, reproducible study selection, precise risk-of-bias assessment, and accurate data extraction are essential. AI assistance cannot replace these methodological approaches or researchers' critical judgment.

The emergence of AI-driven mass production reviews

An emerging concern is the accumulation of redundant, unreliable, insufficiently verified, or intellectually superficial scientific syntheses facilitated by generative AI. By reducing the time and effort required to produce review manuscripts, AI may amplify existing publication pressures and encourage the proliferation of articles that recycle established knowledge without meaningful intellectual contributions. This repetition can create a misleading impression of scientific consensus. Academic incentives, publication pressure, and the accessibility of generative AI may jointly encourage the "mass production" of low-value reviews.

Recognizing AI-generated academic writing: Warning signs and limitations

Another challenge concerns recognizing manuscripts that AI may have generated without adequate scholarly oversight. Figure 1, adopted from (4), shows the number of academic sentences modified by LLMs over time. Suspected characteristics include excessively polished but generic language, repetitive paragraph structures, superficial discussions of conflicting findings, and frequent expressions such as "rapidly evolving landscape," "multifaceted challenges," "it is crucial to note," and "pivotal role." In addition, a significant shift in the frequency of certain words, such as "realm", "pivotal", "meticulous", and "intricate", has been reported (4).

Figure 2, adapted from reference 5, presents the 100 adjectives most frequently overused by AI. Figure 3, also adapted from reference 5, presents a word cloud of the top 100 adjectives in LLM feedback.

Other warning signs include claims that lack specific evidence, repetitive conclusions, and references that do not support the statements. However, these characteristics alone

***Corresponding author:** Bibi Sedigheh Fazly Bazzaz. School of Pharmacy, Mashhad University of Medical Sciences, Mashhad, Iran. Tel: +98-51-31801100, Email: fazlis@mums.ac.ir

© 2027. This work is openly licensed via CC BY 4.0.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

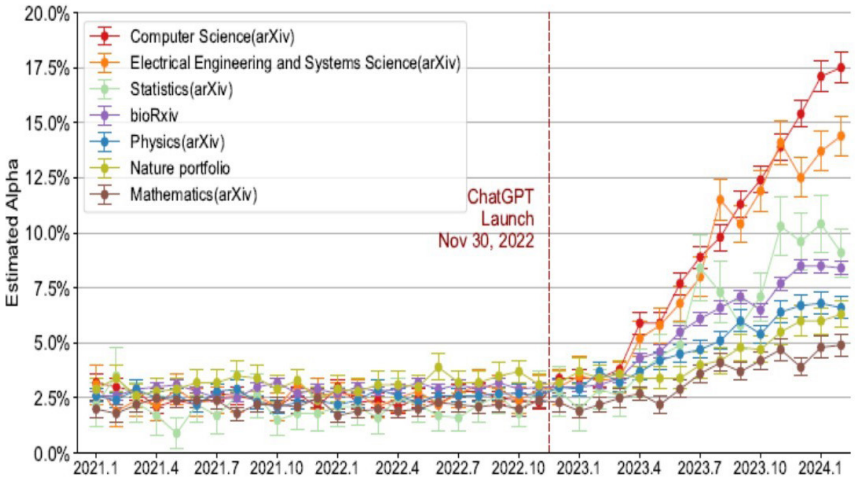


Figure 1. Number of sentences in academic writing modified by LLMs over time, adapted from (4)

cannot reliably indicate that AI authored a text. Human writers may show similar patterns, and AI-detection tools have demonstrated biases, especially against non-native English speakers (6). Therefore, editorial evaluations should focus on scientific content and reference accuracy rather than solely on linguistic style.

Implications for evidence-based medicine and clinical decision-making

The implications of this concern go beyond academic publishing. Review articles play a crucial role in informing clinical guidelines, health policies, and evidence-based decision-making. If unreliable AI-assisted syntheses are published, future studies may replicate their findings or include them in clinical recommendations, spreading misinformation. Additionally, an increasing number of low-value reviews may overwhelm researchers, clinicians, and policymakers, making it harder to identify trustworthy evidence. In this context, more publications do not necessarily mean more reliable scientific knowledge.

Rethinking authorship and scientific responsibility

AI can write a review, but who takes responsibility for its conclusions? This issue highlights the need for renewed attention to authorship and scientific responsibility. Authorship implies accountability for the accuracy and integrity of published work. AI systems cannot assume these responsibilities. The World Association of Medical Editors (WAME) emphasizes that chatbots should not be listed as authors and that human researchers remain accountable for AI-generated content (7). Recently, Elsevier revised its generative AI policies page to provide editors, book authors, and contributors with more explicit guidance on the acceptable

commendable	innovative	meticulous	intricate	notable
versatile	noteworthy	invaluable	pivotal	potent
fresh	ingenious	cogent	ongoing	tangible
profound	methodical	laudable	lucid	appreciable
fascinating	adaptable	admirable	refreshing	proficient
intriguing	thoughtful	credible	exceptional	digestible
prevalent	interpretative	remarkable	seamless	economical
proactive	interdisciplinary	sustainable	optimizable	comprehensive
vital	pragmatic	comprehensible	unique	fuller
authentic	foundational	distinctive	pertinent	valuable
invasive	speedy	inherent	considerable	holistic
insightful	operational	substantial	compelling	technological
beneficial	excellent	keen	cultural	unauthorized
strategic	expansive	prospective	vivid	consequential
manageable	unprecedented	inclusive	asymmetrical	cohesive
replicable	quicker	defensive	wider	imaginative
traditional	competent	contentious	widespread	environmental
instrumental	substantive	creative	academic	sizeable
extant	demonstrable	prudent	practicable	signatory
continental	unnoticed	automotive	minimalistic	intelligent

Figure 2. The 100 adjectives used disproportionately more often by AI, adapted from reference 5

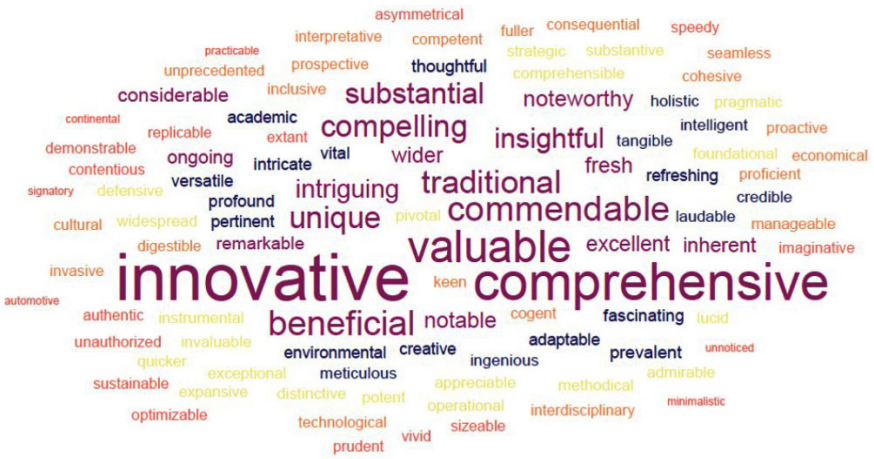


Figure 3. Word cloud of the 100 most frequently used adjectives in LLM feedback; font size indicates frequency. Adapted from reference 5

use of AI tools. Elsevier emphasizes that "Authors may use AI, but remain accountable" (8). Simply disclosing AI assistance is insufficient. The key issue is whether authors have thoroughly assessed the evidence, verified references, and applied independent scientific judgment. Addressing these concerns does not require banning AI in scientific writing. Properly validated AI tools can help organize literature, screen and extract data, and edit language. IJBMS asks authors to complete a declaration of AI tools. However, we acknowledge that journals should go beyond generic AI disclosure requirements and establish stronger expectations for verification, methodological transparency, and clear intellectual contributions.

Opportunities and challenges of generative AI for scientific publishing

Generative AI offers both opportunities and challenges for scientific publishing. The growing use of AI-driven evidence reminds us that producing more scientific articles does not necessarily mean producing better scientific knowledge. As machines become more capable of writing reviews, responsibility for ensuring those reviews are accurate, insightful, and worth publishing must remain firmly human.

While machines can help us write reviews more quickly, they should never replace the human judgment that makes those reviews valuable. We acknowledge that AI can genuinely enhance evidence synthesis when it is properly validated and supervised. The real concern is not AI assistance itself, but substituting scholarly judgment with unverified automated outputs. While these concerns are relevant, we still believe AI cannot replace the writers' viewpoints.

We recommend that editors and reviewers critically assess whether a review addresses significant unanswered questions, evaluates conflicting evidence effectively, and provides insights beyond existing publications. Scientific work's originality should be judged based on its contribution to knowledge. Journals should reconsider how they evaluate originality, intellectual contribution, methodological transparency, and the necessity of new review articles. A review that contributes little to new understanding should not be published.

References

1. Arabi L, Roohbakhsh A, Malaekeh-Nikouei B, Fazly Bazzaz BS. The impact of artificial intelligence (AI) in academic writing and publication: Iranian Journal of Basic Medical Sciences (IJBMS) policy. Iran J Basic Med Sci 2025; 28: 1-2
2. Arabi L, Roohbakhsh A, Malaekeh-Nikouei B, Fazly Bazzaz BS. Artificial intelligence (AI) in academic publishing: Legitimate use, plagiarism detection, and ethical challenges. Iran J Basic Med Sci 2026; 29: 1-2.
3. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: Implications in scientific writing. Cureus 2023;15:e35179
4. Liang W, Zhang Y, Wu Z, Lepp H, Ji W, Zhao X, Cao H, Liu S, He S, Huang Z, Yang D. Mapping the increasing use of LLMs in scientific papers. arXiv preprint 2024; 2404.01268
5. Liang W, Izzo Z, Zhang Y, Lepp H, Cao H, Zhao X, Chen L, Ye H, Liu S, Huang Z, McFarland DA. Monitoring AI-modified content at scale: A case study on the impact of ChatGPT on AI conference peer reviews. arXiv preprint 2024; 2403.07183
6. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. Patterns 2023; 4:100779
7. Zielinski C, Winker MA, Aggarwal R, Ferris LE, Heinemann M, Lapeña Jr JF, Pai SA, Citrome L, Alam M, Voight M, Habibzadeh F. Chatbots, generative AI, and scholarly manuscripts: WAME recommendations on chatbots and generative artificial intelligence in relation to scholarly publications. Colombia Médica 2023; 54:e1015868.
8. https://www.elsevier.com/about/policies-and-standards/publishing-ethics-books/generative-ai-policies