DOI: https://doi.org/10.2298/SARH240704053P
UDC: 001.891:004.8

PERSONAL VIEW ARTICLE / ЛИЧНИ СТАВ

AI in science – dusk or dawn?

Maša Petrović[1,2], Srđan Babić[1,2], Ana Dimitrijević[2], Branko Lozuk[1], Milovan Bojić[1]

[1]Dedinje Institute for Cardiovascular Diseases, Belgrade, Serbia;
[2]University of Belgrade, Faculty of Medicine, Belgrade, Serbia

SUMMARY
The peer review process remains a cornerstone of scientific integrity, ensuring that research findings are critically evaluated before entering the scientific record. With the growing integration of artificial intelligence (AI) and the widespread adoption of large language models (LLMs) such as ChatGPT, the research and publishing landscape is undergoing rapid transformation. While AI offers considerable advantages – enhancing efficiency in manuscript drafting, editing, and preliminary evaluation – it also introduces significant risks, particularly when used beyond its optimal scope. This viewpoint underscores the limitations of generative AI, including the phenomenon of "hallucinated" references and the inability to perform genuine critical thinking. These shortcomings raise serious concerns about the validity of scientific content when AI is used without appropriate human oversight. Emphasis is placed on preserving the human-centered nature of peer review, which is vital to safeguarding scientific credibility. In doing so, this article reinforces the necessity of evolving editorial and publishing policies, such as Elsevier's updated guidelines on the use of generative AI, to ensure responsible integration of these technologies into the research ecosystem.
Keywords: artificial intelligence; ChatGPT; large language models; peer review

The peer review process has long served as a cornerstone of scientific integrity, ensuring that manuscripts undergo rigorous evaluation by experts before publication. This system not only validates methodological soundness and scientific merit but also provides reassurance to clinicians and policymakers that published findings can be reliably integrated into evidence-based medical practice. However, despite its value, peer review is not without limitations – chiefly, the time-consuming nature of the process and the inherent risk of cognitive and personal biases [1].

With the advent of artificial intelligence (AI) and, more recently, large language models (LLMs) such as ChatGPT, there has been a growing temptation to streamline the scientific publishing pipeline. These tools offer appealing solutions to common barriers in scientific communication: drafting outlines, overcoming writer's block, performing rapid literature summarization, and even translating or proofreading manuscripts in record time [2, 3]. Yet, while the capabilities of AI are undeniably impressive, this raises a critical question: what are the limitations and implications of integrating generative AI into the publication workflow, particularly in the domain of peer review?

At present, many leading publishers, including Elsevier, Springer Nature, and JAMA Network, have established formal policies governing the use of generative AI in scientific writing and peer review [4–8]. These policies often emphasize transparency, discouraging unacknowledged AI authorship and warning against reliance on AI-generated content without human validation. The core concern underpinning these restrictions is the phenomenon known as "AI hallucination" – the generation of plausible-sounding but factually incorrect information [9, 10, 11].

This phenomenon poses a serious threat to the dissemination of accurate scientific knowledge. In medicine, where publications directly inform clinical guidelines and therapeutic decisions, the presence of fabricated facts or references can be detrimental. For example, ChatGPT may synthesize text that appears authoritative, complete with fabricated citations and erroneous data, despite having no access to real-time medical databases such as PubMed or updated literature past its training cut-off [9–16]. Even in newer, premium LLMs that are equipped with internet access, the generated references are frequently hallucinatory – fabricated altogether or inserted as placeholders with no meaningful connection to the supported claim. In some instances, the cited reference may be real but entirely unrelated to the content it is purported to substantiate, introducing a false sense of credibility and potentially misleading readers who do not perform manual verification.

Consequently, the uncritical use of such models risks introducing misinformation into the scientific corpus, potentially undermining clinical care and public trust [9, 17].

The peer review process is particularly vulnerable to this dynamic. While AI may be leveraged to assist in administrative triage (e.g., verifying submission completeness or adherence to formatting guidelines), its integration into substantive manuscript evaluation

Received • Примљено: July 4, 2024
Revised • Ревизија: June 27, 2025
Accepted • Прихваћено: July 7, 2025
Online first: July 9, 2025

Correspondence to:
Maša PETROVIĆ
University of Belgrade
Faculty of Medicine
Dr Subotića Starijeg 8
11000 Belgrade, Serbia
5rovicmasa@gmail.com

introduces the risk of dehumanizing a process built upon expert judgment and critical analysis. Peer review is not merely a procedural checkpoint but a cognitive exercise that demands synthesis, skepticism, contextualization, and the application of domain-specific expertise – capabilities that current AI lacks [1, 18].

LLMs such as ChatGPT generate responses based on statistical associations in training data, rather than through genuine comprehension or deductive reasoning. These models operate through token prediction, optimizing linguistic fluency rather than scientific validity [19]. In contrast, human reviewers draw upon a lifetime of experience, ethical reasoning, and real-world understanding of clinical implications – tools that no model, regardless of its complexity, can replicate. Thus, the substitution of human reviewers with AI compromises the foundational purpose of peer review and threatens the gatekeeping function that upholds scientific quality [20].

It is also critical to highlight that the over-standardization introduced by AI-driven manuscript screening or review can discourage novel or paradigm-shifting research. Homogenized feedback, patterned on previous outputs, may suppress the diversity of scientific thought and innovation. Moreover, inappropriate rejection of unconventional but methodologically sound work could prevent important advances from entering the academic discourse.

In light of these concerns, AI should be viewed as an augmentative – not autonomous – tool. It is well-positioned to assist authors in drafting, editing, or organizing manuscripts, and may be used for non-substantive tasks such as checking grammatical accuracy or enhancing language clarity [2, 3]. However, as manuscripts progress through submission and into review, reliance on AI should be consciously minimized to preserve the essential human elements of critique, reflection, and accountability.

In conclusion, while the incorporation of AI in medical research and publishing offers significant promise for increasing efficiency, its use must be bounded by ethical considerations and guided by firm human oversight. Misuse of generative AI risks undermining the reliability of the scientific literature, particularly in medicine, where lives may depend on the accuracy of published findings. Therefore, safeguarding peer review as a human-driven process remains paramount to maintaining the credibility, rigor, and ethical integrity of scientific discourse.

Ethics: This article was written in accordance with the ethical standards of the institutions and the journal.

Conflict of interest: None declared.

REFERENCES

1. Tennant JP, Dugan JM, Graziotin D, Jacques DC, Waldner F, Mietchen D, et al. A multi-disciplinary perspective on emergent and future innovations in peer review. F1000Res. 2017;6:1151. [DOI: 10.12688/f1000research.12037.3] [PMID: 29188015]
2. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. Am J Cancer Res. 2023;13(4):1148–54. [PMID: 37168339]
3. Blanchard F, Assefi M, Gatulle N, Constantin JM. ChatGPT in the world of medical research: from how it works to how to use it. Anaesth Crit Care Pain Med. 2023;42(3):101231. [DOI: 10.1016/j.accpm.2023.101231] [PMID: 37030395]
4. Elsevier. Generative AI Policy for Journals [Internet]. 2025. Available from: https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals.
5. Elsevier. The use of generative AI and AI-assisted technologies in the review process for Elsevier [Internet]. 2025. Available from: https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-the-review-process.
6. Elsevier. The use of generative AI and AI-assisted technologies in writing for Elsevier [Internet]. 2025. Available from: https://www.elsevier.com/about/policies-and-standards/the-use-of-generative-ai-and-ai-assisted-technologies-in-writing-for-elsevier.
7. Springer Nature. Editorial Policies [Internet]. 2025. Available from: https://www.springernature.com/gp/policies/editorial-policies.
8. Flanagin A, Bibbins-Domingo K, Berkwits M, Christiansen SL. Nonhuman "authors" and implications for the integrity of scientific publication and medical knowledge. JAMA. 2023;329(8):637–9. [DOI: 10.1001/jama.2023.1344] [PMID: 36719674]
9. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus. 2023;15(2):e35179. [DOI: 10.7759/cureus.35179] [PMID: 36811129]
10. Homolak J. Opportunities and risks of ChatGPT in medicine, science, and academic publishing: a modern Promethean dilemma. Croat Med J. 2023;64(1):1–3. [DOI: 10.3325/cmj.2023.64.1] [PMID: 36864812]
11. Jones N. AI hallucinations can't be stopped – but these techniques can limit their damage. Nature. 2025;637(8047):778–80. [DOI: 10.1038/d41586-025-00068-5] [PMID: 39838050]
12. Biswas S. ChatGPT and the future of medical writing. Radiology. 2023;307(2):e223312. [DOI: 10.1148/radiol.223312] [PMID: 36728748]
13. Temsah O, Khan SA, Chaiah Y, Senjab A, Alhasan K, Jamal A, et al. Overview of early ChatGPT's presence in medical literature: insights from a hybrid literature review by ChatGPT and human experts. Cureus. 2023;15(4):e37281. [DOI: 10.7759/cureus.37281] [PMID: 37038381]
14. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. J Med Syst. 2023;47(1):33. [DOI: 10.1007/s10916-023-01925-4] [PMID: 36869927]
15. Asgari E, Montaña-Brown N, Dubois M, Khalil S, Balloch J, Yeung JA, et al. A framework to assess clinical safety and hallucination rates of LLMs for medical text summarisation. NPJ Digit Med. 2025;8(1):274. [DOI: 10.1038/s41746-025-01670-7] [PMID: 40360677]
16. Roustan D, Bastardot F. The clinicians' guide to large language models: a general perspective with a focus on hallucinations. Interact J Med Res. 2025;14:e59823. [DOI: 10.2196/59823] [PMID: 39874574]
17. Rozencwajg S, Kantor E. Elevating scientific writing with ChatGPT: a guide for reviewers, editors and authors. Anaesth Crit Care Pain Med. 2023;42(3):101209. [DOI: 10.1016/j.accpm.2023.101209] [PMID: 36871626]
18. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet. 2009;374(9683):86–9. [DOI: 10.1016/S0140-6736(09)60329-9] [PMID: 19525005]
19. OpenAI. ChatGPT: Optimizing language models for dialogue [Internet]. 2022. Available from: https://openai.com/blog/chatgpt/.
20. Thorp HH. ChatGPT is fun, but not an author. Science. 2023;379(6630):313. [DOI: 10.1126/science.adg7879] [PMID: 36701446]

Вештачка интелигенција у науци – сумрак или зора?

Маша Петровић[1,2], Срђан Бабић[1,2], Ана Димитријевић[2], Бранко Лозук[1], Милован Бојић[1]

[1]Институт за кардиоваскуларне болести „Дедиње“, Београд, Србија;
[2]Универзитет у Београду, Медицински факултет, Београд, Србија

САЖЕТАК

Процес рецензије представља темељ научне валидности, осигуравајући да резултати истраживања буду критички процењени пре објављивања. Са све већом интеграцијом вештачке интелигенције и широком доступношћу великих језичких модела (*Large Language Models – LLMs*), попут *ChatGPT*-а, научноистраживачки и издавачки процес пролазе кроз значајне промене. Иако вештачка интелигенција доноси бројне предности – побољшање ефикасности у писању, уређивању и почетној евалуацији рукописа – њена примена изван тих оквира носи озбиљне ризике. Овај рад указује на ограничења генеративне вештачке интелигенције, укључујући појаву „халуцинираних“ референци и недостатак способности за критичко размишљање, што може угрозити научну поузданост када се ови алати користе без одговарајућег људског надзора. Посебан акценат стављен је на очување људске улоге у процесу рецензије као кључне карике у очувању кредибилитета науке. У том контексту, аутори подржавају потребу за ажурирањем издавачких политика, попут нових смерница издавачке куће *Elsevier* о употреби генеративне вештачке интелигенције, како би се осигурала одговорна и етичка интеграција ових технологија у научноистраживачки екосистем.

Кључне речи: вештачка интелигенција; *ChatGPT*; велики језички модели; рецензија