Navigating the complexities of artificial intelligence in scientific writing: a dual perspective

Gabriel Levin ,[1] Sabrina Piedimonte,[2] Behrouz Zand [3]

[1]Gynecologic Oncology, McGill University, Montreal, Quebec, Canada
[2]Department of Obstetrics and Gynecology, University of Montreal, Montreal, Quebec, Canada
[3]Houston Methodist Hospital, Houston, Texas, USA

Correspondence to
Dr Behrouz Zand; bzand@houstonmethodist.org

Received 12 May 2024
Accepted 28 June 2024
Published Online First 7 August 2024

To cite: Levin G, Piedimonte S, Zand B. *Int J Gynecol Cancer* 2024;**34**:1495–1498.

Artificial intelligence (AI) is rapidly being integrated into many aspects of patient care in oncology,[1] and presents unique opportunities to improve predictions in all aspects including cancer screening, diagnosis, treatment, surgical outcomes, recurrence, and survival.[2] Beyond the advancements in utilizing AI in clinical oncology, AI is also widely used by scholars and researchers.[3] Recently, the publication of scientific medical literature has seen a significant rise, with an annual growth rate of approximately 10%.[4] However, this rise driven by the competitive pressure of 'publish or perish' in academia may lead to the potential of authorship misconduct and misrepresentation.[5] The recent introduction of generative artificial intelligence (GAI) carries the potential both to revolutionize scientific writing and to be ethically misused.[6]

Current editorial and peer-review systems are already strained by the overwhelming influx of publications and may not be able to stand guard against the potential misuse of GAI in scientific writing. This begs the question: Can a machine effectively navigate the vast expanse of academic knowledge better than a scholar? This question underscores the transformative impact of GIA in scientific writing. A recent study showed that GAI-generated abstracts went undetected by this very journal's reviewers of various expertise levels – underscoring the importance of discussing the upsides and downsides of the integration of GAI into scientific and academic practices.[7] This editorial explores the intricate balance of risks and benefits introduced by GAI in scientific writing, aiming to illuminate how this technology may influence the dissemination of medical knowledge.

The scholarly GAI technologies are revolutionizing the traditional academic publishing process.[8] AI in academic publishing has transformative potential beyond assisting individual authors. It may also streamline traditional editorial processes such as manuscript handling, peer review, and formatting. By automating these processes, AI frees editors to focus on nuanced decision-making. Moreover, AI optimizes reviewer selection and enhances search result sorting in databases like PubMed, enhancing research accessibility.[9] AI-powered summarizers and translators make complex articles more accessible, benefiting scholars of all levels, including those hindered by language barriers.[10 11] In oncology, AI expedites literature reviews, meta-analyses, and systematic reviews bolstering research efficiency and accuracy.[12]

Tools such as Scholarcy and Scite offer concise summaries, highlight key studies, and identify areas for further investigation.[13 14] Text mining tools like Scopus and Web of Science aid in trend analysis and identifying emerging research themes.[15] AI also facilitates predicting future trends and efficiently matches manuscripts with suitable journals, optimizing the submission process for increased publication success.[16] Additionally, AI's capacity for analyzing big data accelerates discovery by uncovering hidden trends and patterns.[17–19] However, this efficiency comes with trade-offs. The reliance on AI for critical tasks such as peer review may compromise the depth of scrutiny, potentially leading to the acceptance of subpar research if not reviewed by human experts.[20] Furthermore, AI could miss nuanced arguments or context-specific details often caught by human editors, and it could not replace the nuanced judgment and experiential insights of clinicians and patients.[21] Another significant concern is the potential for bias in GAI content, as these models may inadvertently reflect or amplify existing biases in the data they were trained on.[22] Overreliance on AI algorithms, devoid of human input or oversight, could amplify unintended 'hallucinations'.[23–25] These hallucinations are text outputs that appear coherent and plausible but are factually incorrect or nonsensical. For example, if you asked me a very detailed question about the most effective treatment for advanced stage ovarian cancer, a large language model (LLM) might string together information to provide what seems like a convincing answer based on a clinical trial. But those details may not be factual, and it may even have made up that clinical trial. The LLM may have 'hallucinated' this based on patterns in its training data without having a true source to back the patterns up.

The concern is that as AI systems become more advanced, these hallucinations could become harder to spot. If humans put too much faith in AI outputs

without verification, it could lead to the spread of misinformation or flawed decisions based on inaccurate information which can be exceedingly dangerous in medicine.

While AI is invaluable for analyzing structured data, its application to unstructured textual data without supervision poses challenges. It is imperative to prevent AI-driven research in oncology from incorporating hallucinations or misinformation, such as misinterpreting genetic variations or biasing outputs based on population characteristics. These risks may arise during data collection, preparation, sampling, model development, and evaluation.[24]

The integration of AI also raises profound ethical questions. The misuse of GAI tools to fabricate or manipulate data could undermine the credibility of academic work.[26] Their ability to produce text indistinguishable from human-written content simplifies the manuscript writing process, rendering it effortless – further straining the overloaded publishing system.[7 27] Privacy issues, particularly data used to train AI models, are also of significant concern.[28 29] For instance, the Royal Free NHS Trust granted Google's DeepMind access to over a million NHS patient records without patient consent, sparking significant privacy debates and legal challenges.[30] The opaque nature of AI algorithms, often referred to as a 'black box,' exacerbates privacy challenges, especially in handling personal medical information. This issue is particularly sensitive in oncology, where genetic data with broader familial implications could be compromised.

Despite efforts to anonymize data, re-identifying individuals remains a concern, necessitating the exploration of alternative approaches such as generative data models that produce synthetic data.[28] Additionally, as AI technologies become increasingly integrated into healthcare settings like gynecologic oncology, transparency and accountability are crucial. Clinicians and patients must understand AI systems' functioning, decision-making processes, and accountability to build trust and address concerns regarding algorithmic bias or errors.

To navigate these challenges effectively, rigorous oversight and explicit contractual agreements are essential when private AI companies handle patient data.[31] Upholding patient rights and preventing data misuse are paramount as AI continues to advance and reshape medical practice.

Significantly, there is a concerning potential for AI technologies to worsen healthcare disparities. Take for example GPT-4, the multimodal LLM created by OpenAI. A recent study conducted a comprehensive analysis of GPT-4's performance in various medical tasks. It revealed troubling racial and gender biases deeply embedded within its functions.[32] These biases impact critical aspects of medical practice, including education, diagnostic reasoning, treatment recommendations, and patient assessments. GPT-4's tendencies to exaggerate disease prevalence differences among demographic groups, perpetuate stereotypes, and amplify harmful biases are particularly troubling. Moreover, there is a natural tendency for AI benefits to favor more privileged populations over those in rural or underserved areas. To promote fairness in gynecologic oncology care, it is imperative to ensure equitable access to AI tools and address potential biases in AI algorithms. By addressing these disparities, we can work toward a more just and inclusive healthcare system.

As AI increasingly integrates into academia, it warrants a revaluation of traditional roles within academic publishing. With machines assuming more tasks, the responsibilities of human reviewers and editors require redefinition. Academic work may be welcomed by reviewers or editors in tasks that can be performed efficiently and quickly by AI algorithms. Routine tasks like language quality screening, plagiarism checks, adherence to journal guidelines, and reference recommendation by keyword identification can be efficiently handled by AI algorithms, potentially sparing reviewers, and editors from these mundane duties. AI tools have also been suggested to identify AI-written scientific text.[33 34] However, reviewers using AI must ensure they maintain high-quality peer-review standards. The ease of performing a complete peer review through GAI modules may be incredibly attractive for reviewers, who are already overburdened and function on a voluntary basis.[35] This may potentially lead to erroneous decisions by editorial boards. Despite guidelines from organizations such as the International Committee of Medical Journal Editors,[36] the consistent implementation of AI standards across scholarly publishing remains a challenge.

We stand at the brink of an era where hybrid AI-human scientific writing is now commonplace, blurring the distinctions between the two. Our aim should be to move forward to an era where human productivity and innovation thrive, language barriers dissolve, and humans retain responsibility and accountability for generated content. This vision aligns with AI's role in clinical settings, enhancing predictions and personalized care while preserving the invaluable human elements of touch, emotion, and medical judgment.

The solution is not to completely stop AI but rather to maximize its potential benefits while maintaining ethical principles. Developing comprehensive standards and guidelines for AI in publishing is crucial. These guidelines should address the quality of outputs and ensure transparency and accountability in AI processes. Publishers and journals have already suggested integrating AI usage guidelines and recommendations.[37–39] These revolve around maintaining standards of accountability, transparency, and confidentiality. The future should not see AI replacing humans but augmenting them. A hybrid model, where AI handles routine tasks, allowing humans to focus on complex decision-making and creative processes, might be ideal. Such collaboration could enhance both efficiency and quality in academic publishing. AI for detecting plagiarism is a precise example of how a hybrid model should work. Publishers use AI for efficiently detecting plagiarism, saving time for editorial teams and reviewers, and providing the editorial team with accurate data to suggest recommendations that require a human touch and common sense.

It is important to use AI responsibly when generating content for scientific writing. Lack of human oversight could potentially lead to the inclusion of inaccurate or misleading information. A recent study[40] highlighted this risk when an AI-generated paragraph was unintentionally included in the discussion section of a published article. The AI-generated text, which stated, 'In summary, the management of bilateral iatrogenic I'm very sorry, but I don't have access to real-time information or patient-specific data, as I am an AI language model. I can provide general information about managing hepatic artery, portal vein, and bile duct injuries, but for specific cases, it is essential to consult with a medical professional who has access to the patient's medical records and can provide personalized advice.

It is recommended to discuss the case with a hepatobiliary surgeon, or a multidisciplinary team experienced in managing complex liver injuries,' was missed by both the authors and the journal reviewers, highlighting the need for vigilance and human oversight when using AI in scientific writing. This article was subsequently retracted by the journal as the authors did not get patient informed consent. They also violated the journal's policy as they used generative AI in their writing without disclosure.

AI undeniably offers substantial benefits to academic publishing, many of which are already being realized by publishers, journals, and editorial teams, as well as authors. A new era of enhanced efficiency, greater accessibility, and the ability to uncover new insights is currently unfolding. Yet, it also poses significant risks that could undermine the integrity and quality of scientific writing. It is essential for the academic community to actively participate in shaping the trajectory of AI integration. By engaging in informed discussions and policymaking, stakeholders can ensure that AI is used as a beneficial tool. The emergence of special task forces and workshops within oncology societies aimed at addressing the integration of AI in scientific writing may be of value. As we stand on the brink of potentially the most significant transformation in academic publishing, let us navigate this new era with a sharp vision that leverages the strengths of AI but remains steadfastly guided by human wisdom and ethical vigilance.

X Behrouz Zand @ZandBehrouz

Contributors BZ: Led the conceptualization and design of the manuscript, performed part of the literature review and interpretation of findings, and was responsible for the final proofreading and editing of the manuscript. Approved the final version to be published. GL: Conducted a significant portion of the literature review and was primarily responsible for the interpretation of findings. Contributed substantially to writing the sections on literature review findings and their interpretations in the manuscript. Reviewed and approved the final version of the manuscript, ensuring accuracy and integrity. SP: Provided expertise in artificial intelligence, contributed to specific sections of the manuscript dealing with AI applications, and reviewed the manuscript for technical accuracy. Approved the final manuscript and ensured the technical details were correctly represented. We utilized AI for proofreading and assisting with the literature review. AI-based tools enhanced grammatical accuracy and readability for proofreading, ensuring the manuscript adhered to high linguistic standards. This was critical for clearly communicating complex ideas. Additionally, we employed natural language processing and machine learning algorithms for some of the literature reviews. These tools streamlined the review process by efficiently identifying and summarizing relevant studies, which allowed us to integrate current and significant findings into our editorial.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer reviewed.

ORCID iDs
Gabriel Levin http://orcid.org/0000-0003-1282-5379
Behrouz Zand http://orcid.org/0000-0002-8327-3535

REFERENCES

1 Kolla L, Parikh RB. Uses and limitations of artificial intelligence for oncology. *Cancer* 2024;130:2101–7.
2 Lotter W, Hassett MJ, Schultz N, *et al*. Artificial intelligence in oncology: current landscape, challenges, and future directions. *Cancer Discov* 2024;14:711–26.
3 Bagenal J. Generative artificial intelligence and scientific publishing: urgent questions, difficult answers. *Lancet* 2024;403:1118–20.
4 Zhao X, Jiang H, Yin J, *et al*. Changing trends in clinical research literature on PubMed database from 1991 to 2020. *Eur J Med Res* 2022;27:95.
5 Kearney M, Downing M, Gignac EA. Research integrity and academic medicine: the pressure to publish and research misconduct. *J Osteopath Med* 2024;124:187–94.
6 Prillaman M. Is ChatGPT making scientists hyper-productive? The highs and lows of using AI. *Nat New Biol* 2024;627:16–7.
7 Levin G, Pareja R, Viveros-Carreño D, *et al*. Association of reviewer experience with discriminating human-written versus ChatGPT-written abstracts. *Int J Gynecol Cancer* 2024;34:669–74.
8 Vincent JL. How artificial intelligence will affect the future of medical publishing. *Crit Care* 2023;27:271.
9 Kiester L, Turp C. Artificial intelligence behind the scenes: PubMed's best match algorithm. *J Med Libr Assoc* 2022;110:15–22.
10 Noel-Storr A, Dooley G, Affengruber L, *et al*. Citation screening using crowdsourcing and machine learning produced accurate results: evaluation of Cochrane's modified Screen4Me service. *J Clin Epidemiol* 2021;130:23–31.
11 Hamel C, Kelly SE, Thavorn K, *et al*. An evaluation of DistillerSR's machine learning-based prioritization tool for title/abstract screening - impact on reviewer-relevant outcomes. *BMC Med Res Methodol* 2020;20:256.
12 Yao X, Kumar MV, Su E, *et al*. Evaluating the efficacy of artificial intelligence tools for the automation of systematic reviews in cancer research: a systematic review. *Cancer Epidemiol* 2024;88:102511.
13 Scholarcy-knowledge made simple. Scholarcy, Available: https://www.scholarcy.com/
14 Brody S. Scite. *J Med Libr Assoc* 2021;109:707–10.
15 Spasić I, Livsey J, Keane JA, *et al*. Text mining of cancer-related information: review of current status and future directions. *Int J Med Inform* 2014;83:605–23.
16 Kousha K, Thelwall M. Artificial intelligence to support publishing and peer review: a summary and review. *Learn Publ* 2024;37:4–12.
17 Khan ZF, Alotaibi SR. Applications of artificial intelligence and big data analytics in m-Health: a healthcare system perspective. *J Healthc Eng* 2020;2020:8894694.
18 Nayarisseri A, Khandelwal R, Tanwar P, *et al*. Artificial intelligence, big data and machine learning approaches in precision medicine & drug discovery. *Curr Drug Targets* 2021;22:631–55.
19 Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med* 2023;388:1201–8.
20 Donker T. The dangers of using large language models for peer review. *Lancet Infect Dis* 2023;23:781.
21 Buriak JM, Hersam MC, Kamat PV. Can ChatGPT and other AI bots serve as peer reviewers? *ACS Energy Lett* 2024;9:191–2.
22 Hastings J. Preventing harm from non-conscious bias in medical generative AI. *Lancet Digit Health* 2024;6:e2–3.
23 Hatem R, Simmons B, Thornton JE. A call to address AI "hallucinations" and how healthcare professionals can mitigate their risks. *Cureus* 2023;15:e44720.
24 Vokinger KN, Feuerriegel S, Kesselheim AS. Mitigating bias in machine learning for medicine. *Commun Med (Lond)* 2021;1:25.
25 Mittermaier M, Raza MM, Kvedar JC. Bias in AI-based models for medical applications: challenges and mitigation strategies. *NPJ Digit Med* 2023;6:113.
26 Rahimi F, Talebi Bezmin Abadi A. ChatGPT and publication ethics. *Arch Med Res* 2023;54:272–4.
27 Levin G, Meyer R, Yasmeen A, *et al*. ChatGPT-written OBGYN abstracts fool practitioners. *Am J Obstet Gynecol MFM* 2023;100993.
28 Murdoch B. Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC Med Ethics* 2021;22:122.
29 Sweeney SM, Hamadeh HK, Abrams N, *et al*. Challenges to using big data in cancer. *Cancer Res* 2023;83:1175–82.
30 Lomas N. 2024. Available: https://techcrunch.com/
31 Powles J, Hodson H. Google DeepMind and healthcare in an age of algorithms. *Health Technol (Berl)* 2017;7:351–67.
32 Zack T, Lehman E, Suzgun M, *et al*. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *Lancet Digit Health* 2024;6:e12–22.
33 Levin G, Meyer R, Guigue P-A, *et al*. It takes one to know one-Machine learning for identifying OBGYN abstracts written by ChatGPT. *Int J Gynaecol Obstet* 2024;165:1257–60.

34 Levin G, Meyer R, Kadoch E, *et al*. Identifying ChatGPT-written OBGYN abstracts using a simple tool. *Am J Obstet Gynecol MFM* 2023;5:100936.
35 Hosseini M, Horbach S. Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res Integr Peer Rev* 2023;8:4.
36 (ICMJE) TICoMJE. Up-dated ICMJE recommendations. January 2024.
37 Leung TI, de Azevedo Cardoso T, Mavragani A, *et al*. Best practices for using AI tools as an author, peer reviewer, or editor. *J Med Internet Res* 2023;25:e51584.
38 Stanbrook MB, Weinhold M, Kelsall D. A new policy on the use of artificial intelligence tools for manuscripts submitted to *CMAJ*. *CMAJ* 2023;195:E958–9.
39 Ganjavi C, Eppler MB, Pekcan A, *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.
40 Raneem Bader AI, Alnees M, Adler N, *et al*. Successful management of an Iatrogenic portal vein and hepatic artery injury in a 4-month-old female patient: a case report. *Radiol Case Rep* 2024;19:2106–11.