Review Article

An AI-Human Collaborative Review of AI-Human Collaborative Reviews

Arindam Basu[1]

1. University of Canterbury, New Zealand

Generative artificial intelligence ("genAI") refers to applications of artificial intelligence used to generate content, including prose, poetry, scholarly documents, images, audio, and video files. A possible use case for genAI is the authoring of scholarly documents—including research reviews, primary research papers, and research proposals. GenAI is associated with risks of AI hallucination, where fake, spurious, and fraudulent materials are generated by genAI that pass off as authentic materials, leading to misinformation and ethical issues when it comes to research outputs. Given that genAI can be both convenient and harmful, the goal of this paper is to conduct a review of the state of the art of the balance in the use of AI tools. Three AI tools were used to develop this review: for search and expansion of search, and for data extraction. AI tools were not used for content authoring. The results of this review suggest that genAI tools, when combined with human authoring, can provide excellent exemplars of human-AI collaboration, particularly in improving the flow and quality of the output. At the same time, caveats and preventive frameworks must be put in place that can ensure transparency and foster responsible research conduct.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

Introduction

Emergence of generative AI as a collaborative tool

Artificial Intelligence ("AI") can mimic human-like processes (including learning, adapting, synthesising data, and self-correction) using software algorithms interacting with the environment [1]. Generative AI tools ("genAI") are those that can "generate" new information, including artefacts, text materials, images, music, and videos, using prompts that humans issue them; generalised pre-trained transformer

 doi.org/10.32388/3YMXXC

models are potentially useful for conducting automated literature reviews and knowledge discovery in biomedicine, for example, in drug discovery [2]. A paradigm of creativity has been ushered in since the inception of the conversational generative AI application, ChatGPT, in November 2022. In academia, generative AI can assist in idea generation, essay evaluation, storytelling, and providing feedback, even being considered a co-author in students' manuscripts [1].

As artificial generative pre-trained transformer-based applications continue to develop, this has opened up a metaphorical "Pandora's Box"; while on the one hand, genAI has the potential to level an uneven playing field by allowing users with limited English language skills to take advantage of automated spelling and grammar correction, thereby enabling them to express complex concepts in nuanced English to publish their research that, if specialised editing services were to be used, would have otherwise been time-consuming and expensive, there are equally concerns about AI generating erroneous and biased outputs, particularly when they remain unchecked, and issues around AI hallucinations. AI hallucination refers to the phenomenon where AI creates a convincing, contextually coherent, yet fabricated output in response to the user's input or previous context in response to a query. To mitigate the effect or impact of hallucination, it is essential that the prompts are contextually aware [3].

Given that generalised pre-trained transformer ("GPT") based applications can work both ways, in that scholars can use them to rapidly create scholarly documents, yet they have the inherent potential for serious errors, the goal of this paper is to review and critically examine both sides and identify the balance and, in particular, the state of knowledge as to how AI might best be used under the prevailing circumstances. To achieve this aim, this review has used AI tools as collaborating technology to answer the question, "How can AI and humans collaborate with each other for academic writing, what are the possible perils and pitfalls, and what are some recommended steps to mitigate such risks?" The next sections describe the methods of this process, and the subsequent sections lay out the results and discussion of the core issues.

Materials and Methods

The purpose was to conduct a narrative review of the extant literature to identify the current usage, limitations, and recommendations for the use of generative pre-trained transformer-based models for research on health and related sciences. The following query in the form of a prompt was presented to an AI model: "Identify the current status, benefits, drawbacks and recommended practices for using generalised pre-trained models for collaboration between humans and Artificial Intelligence". The AI model used for this purpose is referred to as AI2 Paperfinder, a free web-based tool found at https://paperfinder.allen.ai/

AI2 Paperfinder is a large language model-powered search system that iteratively builds a query from a prompt, as described in the following document:

> "When you (referring to the user) enter a query, you can watch as the system breaks down your query into relevant components, searches for papers, follows citations, evaluates for relevance, runs follow-up queries based on the results, and then presents not only the papers, but also short summaries of why the paper is relevant to your specific query."

(https://allenai.org/blog/paper-finder)

The list of studies was then expanded further using a second Artificial Intelligence application referred to as ResearchRabbit, a free web-based tool where one can use seed publications to generate further recommendations for publications that are related to the group of seed publications and visualise the inter-relationships among the individual publications. This process helps to support unstructured searching; this tool was used in conjunction with the Paperfinder tool to explore additional relevant publications related to the use of generative AI tools for scholarly literature outputs [4]. The tool can be found here: https://www.researchrabbit.ai/

Once the set of relevant publications was identified, the full text of the publications was retrieved. This was done in two ways: (1) where the ResearchRabbit webpage indicated the availability of the PDF, the PDF was downloaded to the computer, and (2) where a full text was not immediately available as indicated by ResearchRabbit, the full text was accessed using the link function of the ResearchRabbit application. The full text was then accessed for the paper.

The third tool used was Google NotebookLM. Google NotebookLM uses a conversational AI-powered tool that extracts information from PDF documents and enables "deep research" into a specific text or resource uploaded to the Google NotebookLM workspace [5]. Google NotebookLM was used to extract information from the PDFs to answer the following questions as prompts: (1) identify the authors of the publication; (2) summarise the main messages of the publication; and (3) list the conclusion from the publication or study paper. This information was verified with an independent review of the publication using these three objectives. In this case, information was abstracted based on the above questions. Google NotebookLM can be found at the following resource location: https://notebooklm.google.com/

These three tools (AI2 Paper Finder, ResearchRabbit, and Google NotebookLM) were combined to generate lists of references and citations and generate conversational insights based on questions and an independent review of the resources. The texts were then summarised to generate themes on three aspects to explore: (i) the benefits of how humans and AI can collaborate with each other to generate scholarly documents and produce knowledge; (ii) recommended best practices when humans collaborate with AI; (iii) the current status of AI in terms of co-authorship; (iv) limitations and caveats; and (v) recommended steps to make the best use of AI while steering clear of possible pitfalls.

The workflow described here for generating this review is free from the risk of producing fictitious or hallucinatory research publications for the following reasons. First, the list of publications is generated by the Ai2 paper finder tool, which is based on the Semantic Scholar database of over 200 million publication records – hence the source materials are actual publications. Second, ResearchRabbit is a reference filtration and recommender system that checks for associated document object identifiers at the time when the output from the Ai2 paper finder search is ingested into the system. This not only protects against false citations but also helps to locate full-text publications. Third, the PDF reader at Google NotebookLM was used for further processing, which only depended on the resources that were fed to the system, thus eliminating the risk of spurious, fake, or false publications. In generating this review of the current state of knowledge, both peer-reviewed articles and preprints were considered as knowledge objects, and both types of publications were critically appraised before summarisation. Finally, all uploaded materials were manually read, and data were abstracted and combined with insights generated from NotebookLM to verify the sources of information.

Results

Benefits of teaming up with AI to write papers

In the corpus of studies reviewed, there was a consensus that when humans and AI collaborate, the overall impact is greater than the sum of the parts. This is because generative AI is better at crafting English language following the rules of syntax and grammar, even though there is an associated risk of dullness; besides, generative AI can provide feedback and guidance in the process. Veach & Abualkibash (2023) tested how well ChatGPT (see https://chatgpt.com/) — a conversational generative pre-trained transformer-based AI model — can be used to write academic research papers. They tasked ChatGPT 3.5 Turbo with creating theoretical papers. Now, this does not quite qualify as an AI-assisted paper being written; rather, here, AI is itself the sole author of a paper, but there are important lessons. They reported that ChatGPT had a strong command of the English language, near-flawless spelling and grammar, and showed a consistent, authoritative, and logical tone [6]. Lin (2023) suggests that for humans, scholarly writing is often a "Sisyphean ordeal", in the sense of repetitive tasks and associated tedium. On the other hand, when humans collaborate with generalist large language models (LLMs), it leads to greater efficiency in academic writing, a joyful writing process, and what they term as cognitive offloading by freeing up mental resources from completing mechanical tasks. Lin identified five degrees of engagement with the AI in a collaborative writing process, starting with basic editing and using large language models as proofreaders; enhancing structural editing using paraphrasing prompts; harnessing the power of generative capacity to generate derivative or new content; and obtaining meaningful feedback in a "non-threatening", controllable environment of academic writing [7]. AbdElsalam & Abdel-Momen (2023) consider the paradigm-shifting potential of AI for scientific writing when humans and AI collaborate. Similar to Lin, they believe that tools like ChatGPT can assist scientists with content arrangement, draft generation, proofreading, extraction of relevant information from information-dense scientific papers, identification of research gaps, and even analysis of reviewer comments to prioritise revisions during the life cycle of content authoring and scholarly communications [8]. Nguyen *et al.* (2024) observed that doctoral students who engage in iterative, highly interactive processes with generative AI (GAI)-powered assisting tools tended to achieve better performance in their writing tasks compared to those students who use generative AI as supplementary sources of information and adopt a more linear style of writing [9].

Best practices of collaborating with AI to write academic documents

Humans engaging and collaborating with AI tools to craft academic documents need to maintain a level of vigilance so that the AI does not run away producing fictitious content and that prompts issued to the AI agents tend to result in differences in the outputs. Veach & Abualkibash (2023) recommend that the proper and thoughtful use of this technology is critical, and a need for humans to be vigilant, particularly with respect to AI-generated content for accuracy, and to maintain scientific integrity [6]. Tu *et al.* (2024) have recommended the use of effective prompt strategies that need to be carefully crafted and strategically formulated so that the researcher can retrieve information that is relevant to the task at hand. For conversational agents, prompts play a pivotal role in steering collaborative discussions, and therefore, clear, explicit prompts, often best crafted by the AI tools themselves, go a long way in ensuring that the resultant human-AI collaborative document can stimulate collective reflection and provoke thoughtful analysis of the research topics under consideration [10].

Can AI be treated as a co-author?

The answer to this question is a unanimous “no”. Lee (2023) examined the question “Can an artificial intelligence chatbot be the author of a scholarly article?” from two perspectives: copyright law and research ethics. From a legal perspective, an AI chatbot cannot be considered the author of a copyrighted work because it is not a human being. Then again, from the perspective of research ethics, the objection to AI chatbots being authors, at least from major scientific publications, is that AI chatbots cannot be held accountable for the work they produce [11]. Even though they are superior to search engines in technological advancement, their status is best relegated to the class of search agents.

Limitations and Caveats: Ethical issues, red teaming and hallucination

Several authors have highlighted issues around ethics, biases, the risk of generating “paper mills”, AI hallucinations, and “red teaming” that plague AI-generated content. Veach & Abualkibash (2023) noted that if humans were to “collaborate” with AI for authoring text, using a tool such as ChatGPT would result in formulaic text. In particular, ChatGPT’s rigid adherence to a formulaic structure to write the English language could strip away a human author’s natural tone and unique identity [6]. Koçak (2024) noted that when authors use AI for academic content authoring, they open themselves to a form of writing that is not free from its inherent bias, over which they have little control, misinformation, and plagiarism that are part of AI-generated content due to their designs. In turn, these interfere with scientific rigour and integrity [12].

Beyond these considerations, the extreme academic productivity of AI-human collaboration can also lead to the emergence of “paper mills” with unprecedented frequency. No doubt when we humans collaborate with AI, the issues around “staring at a blank canvas” or “writer’s block” are relatively less likely for us since a simple prompt such as “draft an outline of a research paper” will help an AI agent to create an outline and populate it with content that goes beyond “lorem ipsum”. Yet such convenience has its price, as an automated process such as this also poses a risk of giving rise to “paper mills”—fraudulent organisations can use this process to generate hundreds of “fake” or fraudulent publications and, in turn, sell them to unsuspecting academics, often early-career academics who need publications for career progression. This, in turn, threatens the trustworthiness of academic publications and indexing services [13].

Generalised pre-trained models are built on different text corpora, including WebText and OpenWebText [14]. When AI collaborates with humans, the machine learning algorithms, trained on such text corpora, build the models based on predicting the next text and thus control the flow of text on the basis of the input from humans; mimicking humans is a function of adding perplexity parameters in the form of “temperature” in the model, usually hidden from the user unless the user uses an application programming interface (“API”) to interact with the model. This becomes compounded when researchers use tools such as ChatGPT or other conversational explainable AI tools where the parameters are hidden and pre-specified, over which the user has no control. In turn, this introduces a “black box” and a level of opacity in the parameters that the AI uses when it collaborates with humans. While this is less of an issue with simple questions and answers with little bearing on daily life, when it comes to scientific healthcare knowledge domains, this lack of transparency can make a difference between life and death.

In particular, issues around AI hallucination, plagiarism, and unintentional or intentional “red teaming” are critical when one considers the adoption of AI as a collaborative tool for academic writing. Kacena *et al.* (2024) assessed the ability of ChatGPT to assist humans in writing credible, peer-reviewed, scientific review articles on a set of topics including Alzheimer’s disease, bone and neural regulation of fracture healing, and relating COVID-19 with musculoskeletal health; they used three sets of writing tasks where they instructed autonomous AI to generate the papers, humans unaided by AI to write research articles, and a collaboration of humans and AI (“AI-assisted”). They fact-checked for accuracy and plagiarism. They found that the AI-only approach resulted in up to 70% of the cited references being inaccurate. These inaccuracies included errors in the citation details (year, authors, title, journal), irrelevance of the text to the citation, or completely fabricated references, and that the AI-assisted approach resulted in a high likelihood of plagiarism in the initial drafts [15].

A well-documented flaw of ChatGPT 3.5 Turbo is its inability to provide accurate or truthful citations. It frequently generates “hypothetical references” that do not exist or are unrelated to the content, which risks the accidental spread of misinformation. Veach & Abualkibash (2023) have identified ethical and legal challenges, including the risk of accidental plagiarism; this is perhaps insurmountable as AI may derive from other works in its training data without proper recognition or citation [6]. There are also data privacy worries concerning the legal basis for gathering training data from the internet at large without explicit consent, which has led to ongoing investigations by governing bodies in multiple countries.

doi.org/10.32388/3YMXXC

Besides, AbdElsalam & Abdel-Momen (2023) identified a few key challenges, where they noted that over-reliance on AI could diminish human skill and intuition in scientific writing, that Large Language Models trained on general data may be less effective in specialised fields or languages with fewer resources, and that they can generate inaccurate or fabricated information, known as "hallucinations" [8].

AI hallucination, or AI's erroneous output, is a common and widespread problem. Among the other, perhaps less discussed aspects is that of "red teaming" and the algorithmic bias of large language models [14]. Red teaming refers to a process where individuals intentionally (or sometimes unintentionally) prompt or challenge an AI model, particularly large language models (LLMs), with text that aims to disrupt its algorithms. What's particularly concerning in the context of academic writing, as highlighted by Jains, is that researchers can *unknowingly* perform red teaming. This can happen if the researchers input large prompts with scientific data or text, or if they lack appropriate context and specificity in their queries to the chatbot. For instance, providing limited context or comparing two tangentially related subjects can inadvertently lead to red teaming [14]. Similarly, copying one's writing, which might include identifiable information, can also unintentionally feed the model context that biases its output, leading to context-induced bias. When an LLM is red-teamed, it leads to the generation of stereotypical or harmful content as the "human in the loop" has introduced the bias that the algorithm magnifies and reinforces. Unfortunately, biases cannot be corrected post hoc and need to be addressed at the time of the model's training. Hence, the problems of lack of transparency and bias are inter-related. As Lee (2023) has noted, a significant problem identified with AI chatbots like ChatGPT is that they cannot provide reliable sources for their writings and possess an "unfortunate ability to provide fake information in a convincing way" [11]. This means that any AI-generated text must be verified for authenticity by a human researcher, highlighting that they are not "ideal" research assistants. Doyal *et al.* (2023) state ethical concerns associated with the use of these AI tools, which include bias, misinformation, privacy, lack of transparency, job displacement, stifling creativity, plagiarism, authorship, and dependence [16].

Possible remedies and recommendations

In the face of these limitations, but also the advantages of AI-human collaborative writing processes, there is a need for review and adjustment of the relationships between humans and machines. Reza *et al.* (2025), for example, in "Co-Writing with AI, on Human Terms: Aligning Research with User Demands Across the Writing Process," have identified four overarching design strategies for AI writing support, which, in turn, emerged from their systematic review of 109 Human-Computer Interaction (HCI) papers published between 2018 and 2024 [17]. From a system designer's perspective, these strategies include structured guidance by an AI agent; AI as a guide to lead the human collaborator on a guided exploration of the topic being navigated; AI as an active co-writer or co-author as a partner and not taking over the role of writer; and finally, AI should serve as a provider of both qualitative and quantitative critical feedback in the revision process. Lee (2023) has cautioned that it is best to treat AI as a research tool and proceed with care, and even though chatbots are not suitable for "authorship" or to serve as co-authors, they advise treating them as research tools, as they caution that AI chatbots can be dangerous research assistants and suggest using them with "heavy" responsibilities on the part of the human collaborator who uses them [11]. Tu *et al.* (2024) suggest that authors who collaborate with AI in their writing should prioritise transparency by clearly disclosing the involvement of AI tools in the creation process [10]. They also noted a need for continual reflection, development, and familiarisation with the AI writing process to alleviate some of the shortcomings that we have discussed in this section. Doyal *et al.* (2023) noted the necessity of developing strategies to understand and address concerns around the detection of bias and misinformation, ensuring privacy and transparency; in the context of healthcare applications, they recommended the need for critical review by experts. Beyond these considerations, researchers and organisations have proposed frameworks that can overcome or at least address the current limitations of AI as human companions to write scholarly manuscripts [16].

Cho *et al.* (2023) have proposed Papercard as a tool for ensuring transparency in academic publications where authors have used AI for writing the manuscript. Papercard is a framework for human authors to declare transparently the use of Artificial Intelligence (AI) in their academic writing process and has four components: (1) a statement of declaration about machine assistance; (2) that authors transparently report the specific types of AI assistance used, ranging from generating key ideas and research questions to answering them, creating entire paper outlines, generating "original" content for sections, drafting, editing, proofreading, or merely offering advice; (3) that authors critically consider potential risks associated with using machine assistance, including the possibility of inaccurate content, harmful content, plagiarism, intellectual property issues, and accountability for misinformation, thereby declaring that they understand these risks and have taken appropriate steps to mitigate them; and (4) that authors specify the names, service providers, relevant dates, versions and specifications of the AI model(s) used, provide the prompts they used, and demonstrate their intellectual contributions [18]. In a sense, fulfilling all four requirements or even the suggestions of the framework are tall asks for many, if

doi.org/10.32388/3YMXXC

not most, authors, and many authors may not be familiar with these specifications, given the stage of AI adoption we are experiencing. The International Association of Scientific, Technical & Medical Publishers (STM) — a global trade association that provides services and support within the scholarly publishing ecosystem — released a "white paper" in 2023 where they articulated the principles for authors, editors, and reviewers of scholarly papers on the use of AI (see https://stm-assoc.org/new-white-paper-launch-generative-ai-in-scholarly-communications/). In the manuscript, they provided advisories for authors, editors, reviewers, and readers in terms of best practices, and what is allowed and not allowed. For authors, they recommend that GenAI can be used as a basic tool for refining, correcting, formatting, and editing texts and documents with disclosure and forbid the use of GenAI to create, alter, or manipulate original research data and results, such as images, blots, photographs, X-rays, and measurements, and that GenAI cannot be credited as an author of a published work. For editorial teams, they warn against using publicly available GenAI platforms for integrity checks in preference to human oversight; their advisory for "readers" is particularly notable, where they note that readers should not upload published manuscripts to publicly available GenAI platforms, as this material might be used in ways that violate copyright or contravene confidentiality/privacy requirements. In an era of an explosion of GenAI as a creative companion for humans and where we are arguing for the case of AI-human collaboration, these may seem restrictive, although judicious on the grounds of careful and deliberate best-case use of GenAI tools for scientific and scholarly publishing.

Discussion

In summary, the findings of this review suggest that, firstly, a human-AI collaboration in composing scholarly articles, grant applications, and reports can leverage the best of both worlds, where human creativity can meet with the infallible accuracy of a software tool following the heuristics of grammar and rules of best composition, and these tasks can be accomplished with unprecedented speed. Secondly, it must be noted that at the same time, as AI is not considered a human equivalent, an AI cannot be a co-author, regardless of its human-like ability to "think" and "suggest" improvements and changes in a manuscript that, were it to be a human being, would deem it worthy of authorship. Thirdly, it behoves the authors of the work to be mindful of the several shortcomings of GenAI systems, notably AI hallucinations and the role of "red teaming," where inadvertent or intentional misuse can lead to undesirable consequences as a result of the use of AI and serious biases. Therefore, best practices behoove that human agents must act in good faith, be transparent about their usage of AI, declare such usage, and refrain from using AI beyond simple editorial services.

Some of these may seem overly restrictive, given the spate of AI development and the multiple use cases of automation and knowledge generation where we are increasingly using AI. While AI hallucination is a reality that is best avoided, the judicious use of AI tools with curation can avoid and eliminate this possibility. For example, in the preparation of this manuscript, an AI tool was used, but the output from that AI tool was double-checked with other tools. The output from Ai2 paper finder was double-checked with researchrabbit when the resulting literature file was uploaded for ingestion, and this provided additional information in identifying those papers that did not have an accompanying document object identifier in them, so the software tracked the correct identifier and catalogued them. In developing this review, a chat-based document parser was also used with Google Notebooklm, but the information was verifiable with a close reading of the full text. Therefore, while caution is important in the context of AI-human collaboration, it is also important to build the necessary checks and balances into the design phase of the study to leverage the best features of AI that can be harnessed by humans with creativity and domain expertise. In conclusion,

The shortcoming of this review was its reliance on AI for the literature search. While this was a methodological choice to demonstrate and iteratively discover the potential of an AI-human collaboration at all phases of the research process, future studies need to utilise both AI with prompts and search engines. Second, the necessary papers from which the knowledge has been developed could have followed a more structured selection criterion, and that would facilitate the development of a more systematic review of the literature using AI tools. It is hoped that AI tools can be successfully used in future to build and nearly automate the process of systematic literature reviews. In its current application, over a short period, it has helped to generate an evidence base that can serve as a reliable talking point.

References

1. [a, b] *Jen, S. L., & Salam, A. R. H. (2024). A Systematic Review on The Use of Artificial Intelligence in Writing. International Journal of Academic Research in Progressive Education and Development. https://doi.org/10.6007/ijarped/v13-i1/20584*
2. [^] *Zhang, Y., Liu, C., Liu, M., Liu, T., Lin, H., Huang, C.-B., & Ning, L. (2023). Attention is all you need: utilizing attention in AI-enabled drug discovery. Briefings in Bioinformatics, 25(1). https://doi.org/10.1093/bib/bbad467*

3. ^ÖZER, M. (2024). IS ARTIFICAL INTELLIGENCE HALLUCINATING? Turkish Journal of Psychiatry. https://doi.org/10.5080/u27587
4. ^Cole, V., & Boutet, M. (2023). ResearchRabbit (product review). Journal of the Canadian Health Libraries Association / Journal de l'Association Des Bibliothèques de La Santé Du Canada, 44(2), 43–47. https://doi.org/10.29173/jchla29699
5. ^Tufino, E. (2025). NotebookLM: An LLM with RAG for active learning and collaborative tutoring. https://doi.org/10.48550/ARXIV.2504.09720
6. a, b, c, dVeach, A., & Abualkibash, M. (2023). Analysing Chatgpt's Potential Through the Lens of Creating Research Papers. International Journal of Computer Science & Information Technology (IJCSIT). https://doi.org/10.5121/ijcsit.2023.15405
7. ^Lin, Z. (2023). Techniques for supercharging academic writing with generative AI. Nature Biomedical Engineering. https://doi.org/10.1038/s41551-024-01185-8
8. a, bAbd-Elsalam, K. A., & Abdel-Momen, S. M. (2023). Artificial Intelligence's Development and Challenges in Scientific Writing. Egyptian Journal of Agricultural Research. https://doi.org/10.21608/ejar.2023.220363.1414
9. ^Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. Studies in Higher Education. https://doi.org/10.1080/03075079.2024.2323593
10. a, bTu, J., Hadan, H., Wang, D. M., Sgandurra, S. A., Mogavi, R. H., & Nacke, L. E. (2024). Augmenting the Author: Exploring the Potential of AI Collaboration in Academic Writing. arXiv.Org. https://doi.org/10.48550/arxiv.2404.16071
11. a, b, cLee, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? Null. https://doi.org/10.3352/jeehp.2022.20.6
12. ^Koçak, Z. (2024). Publication Ethics in the Era of Artificial Intelligence. Journal of Korean Medical Science. https://doi.org/10.3346/jkms.2024.39.e249
13. ^Parker, L., Boughton, S., Bero, L., & Byrne, J. A. (2024). Paper mill challenges: past, present, and future. Journal of Clinical Epidemiology, 176, 111549. https://doi.org/10.1016/j.jclinepi.2024.111549
14. a, b, cJain, R., & Jain, A. (2023). Generative AI in Writing Research Papers: A New Type of Algorithmic Bias and Uncertainty in Scholarly Work. arXiv.Org. https://doi.org/10.48550/arxiv.2312.10057
15. ^Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The Use of Artificial Intelligence in Writing Scientific Review Articles. Current Osteoporosis Reports. https://doi.org/10.1007/s11914-023-00852-0
16. a, bDoyal, A. S., Sender, D., Nanda, M., & Serrano, R. (2023). Chat GPT and Artificial Intelligence in Medical Writing: Concerns and Ethical Considerations. Cureus. https://doi.org/10.7759/cureus.43292

17. ^*Reza, M., Thomas-Mitchell, J., Dushniku, P., Laundry, N., Williams, J. J., & Kuzminykh, A. (2025). Co-Writing with AI, on Human Terms: Aligning Research with User Demands Across the Writing Process. arXiv.Org. https://doi.org/10.48550/arxiv.2504.12488*
18. ^*Cho, W. I., Cho, E. J., & Cho, K. (2023). PaperCard for Reporting Machine Assistance in Academic Writing. arXiv.Org. https://doi.org/10.48550/arxiv.2310.04824*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.