1	A Systematic Examination of Generative Artificial Intelligence (GAI) Usage Guidelines for
2	Scholarly Publishing in Medical Journals
3	Shuhui Yin, MA ¹ *, Peiyi Lu, PhD ² *, Zhuoran Xu, MSc ³ *, Zi Lian, EdD ⁴ , Chenfei Ye, PhD ^{5,6} ,
4	Chihua Li, DrPH ^{7,8}
5	*Equally contributed as first authors
6 7 8	 Applied Linguistics & Technology, Department of English, Iowa State University, IA, USA Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong SAR, China
9 10	3. Graduate Group in Genomics and Computational Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
11 12	 4. Center for Health Equity & Urban Science Education, Teachers College, Columbia University, New York, NY, USA
13 14	5. International Research Institute for Artificial Intelligence, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong, China
15 16 17	6. Pengcheng Laboratory, Shenzhen, Guangdong, China7. State Key Laboratory of Quality Research in Chinese Medicine, University of Macau, MacaoSAR China
17 18 19	8. Survey Research Center, Institute for Social Research, University of Michigan, MI, USA
20 21	Word count: 3,359
22	Corresponding authors
23 24	Chihua Li Institute for Social Research, University of Michigan, ML USA
25	426 Thompson St
26	Ann Arbor, MI 48104
27	Phone: +1 917-528-8696
28 29	<u>chihuali@umich.edu</u>
30	
31	Chenfei Ye
32	International Research Institute for Artificial Intelligence, Harbin Institute of Technology
33 34	(Shenzhen), Shenzhen, China Phone: +86 0755-27905894
57	

- 35 <u>chenfei.ye@foxmail.com</u>
- 36

37 Key points

- **Question** What are the provision and specificity of generative artificial intelligence (GAI) usage
- 39 guidelines for scholarly publishing in top and whole-spectrum medical journals and their
- 40 relationships with journal characteristics?
- 41 Findings Author guidelines were more abundant and specific in top journals than in whole-
- 42 spectrum journals. However, reviewer guidelines were extremely scarce in both groups of
- 43 journals. Journal ranking score was associated with both provision and specificity of GAI usage
- 44 guidelines in whole-spectrum journals while no significant relationship was found in top journals.
- 45 **Meaning** The lack of provision and specificity as well as the inconsistencies in existing
- 46 guidelines suggest that immediate attention is needed to guide GAI usage in scholarly publishing
- 47 and safeguard integrity and trust in medical research.

48

49

51 Abstract

52 Background

53 A thorough and in-depth examination of generative artificial intelligence (GAI) usage guidelines

54 in medical journals will inform potential gaps and promote proper GAI usage in scholarly

55 publishing. This study aims to examine the provision and specificity of GAI usage guidelines

56 and their relationships with journal characteristics.

57 Methods

58 From the SCImago Journal Rank (SJR) list for medicine in 2022, we selected 98 journals as top

59 journals to represent highly indexed journals and 144 as whole-spectrum sample journals to

⁶⁰ represent all medical journals. We examined their GAI usage guidelines for scholarly publishing

61 between December 2023 and January 2024.

62 **Results**

63 Compared to whole-spectrum sample journals, the top journals were more likely to provide

author guidelines (64.3% vs. 27.8%) and reviewer guidelines (11.2% vs. 0.0%) as well as refer to

external guidelines (85.7% vs 74.3%). Probit models showed that SJR score or region was not

associated with the provision of these guidelines among top journals. However, among whole-

67 spectrum sample journals, SJR score was positively associated with the provision of author

68 guidelines (0.85, 95% CI 0.49 to 1.25) and references to external guidelines (2.01, 95% CI 1.24

to 3.65). Liner models showed that SJR score was positively associated with the specificity level

of author and reviewer guidelines among whole-spectrum sample journals (1.21, 95% CI 0.72 to

1.70), and no such pattern was observed among top journals.

72 Conclusions

73 The provision of GAI usage guidelines is limited across medical journals, especially for reviewer

74 guidelines. The lack of specificity and consistency in existing guidelines highlights areas

75 deserving improvement. These findings suggest that immediate attention is needed to guide GAI

vage in scholarly publishing in medical journals.

77 Introduction

78 ChatGPT, a chatbot powered by generative artificial intelligence (GAI) through large language models, was released in November 2022. It can generate responses based on statistical 79 language patterns and is easily accessible to people without technical expertise.¹ Its multifaceted 80 capabilities and potential applications within diverse contexts have attracted billions of users and 81 visits.²⁻⁴ Consequently, many researchers have been actively exploring ChatGPT and other 82 similar tools' potential applications to scholarly publishing, including manuscript preparation 83 and peer review.⁵⁻⁸ To date, at least 1000 medical publications reported the use and impact of 84 such tools in scholarly publishing.⁹⁻¹¹ For manuscript preparation, researchers have used GAI 85 tools to conduct literature review, formulate study design, analyze data, and even interpret 86 results.^{12–15} For peer review, GAI has been used to summarize manuscript contents, review code, 87 check methods, and even draft comments and feedback.^{9,16} These practices highlight that such 88 tools have and will continue to transform scholarly publishing. 89

However, GAI is far from perfect and can lead to multiple concerns.^{17–19} Major concerns 90 regarding its role in scholarly publishing include the eligibility of these tools for authorship, the 91 92 risk of producing misleading or inaccurate information, breaches of data privacy and confidentiality, as well as challenges to integrity and originality.^{20–23} In response, many journals, 93 organizations, and publishers have started to provide GAI usage guidelines for authors and 94 reviewers and made continuous updates.^{24–26} For example, medical publication organizations, 95 96 and publishers, such as the International Committee of Medical Journal Editors (ICMJE) and Elsevier, have decided that GAI should not be listed as authors because they cannot be 97 accountable for the submitted work;^{27,28} some journals require authors to document details of 98 GAI usage and fact-check GAI-generated contents; and some journals strictly prohibit the use of 99 GAI in peer review.²⁹ These developments show the new challenges posed by GAI and the 100 scientific community's corresponding adaptations. 101

Despite a fast-growing increase in discussions on GAI usage in scientific research, few have examined GAI usage guidelines systematically. For example, one study examined GAI usage guidelines for authors based on the 100 largest publishers and top 100 highly ranked journals of different disciplines, and it found less than a quarter of these publishers and around 85% of the top journals provided author guidelines.³⁰ The other study examined author

disclosure requirements for GAI usage in 125 nursing journals and found less than 40% of them
 had explicit instructions.³¹ They offered valuable insights into the evolving ethical standards and
 practices of this rapidly changing field.

However, many other important aspects of GAI usage guidelines remain less explored or 110 unknown. First, few or no existing studies examined GAI guidelines for reviewers, ignoring the 111 equal importance of peer review in scholarly publishing. Second, previous studies either focused 112 on top-ranked journals or only on nursing journals, leading to an incomplete assessment of such 113 114 guidelines across medical journals. Third, little exploration has been made on the relationship between journal characteristics and GAI guidelines for authors and reviewers. To address these 115 gaps, our study focuses on medical journals and includes both top and whole-spectrum journals. 116 117 We aim to delineate the provision and specificity of GAI guidelines for authors and reviewers as 118 well as external guidelines referred by these journals. We will also examine potential 119 relationships between journal characteristics and the provision and specificity of different GAI 120 guidelines. These together will provide a thorough overview of current practices and contribute 121 to the ongoing development of GAI usage guidelines across medical journals.

122

123 Methods

124 Journal selection

Based on the SCImago Journal Rank (SJR), two lists of journals were selected from the journal ranking list for medicine in 2022: 'top' journals and whole-spectrum sample journals (**Figure 1**).³² Each journal's ranking was determined by its corresponding SJR score, a measure of the number of citations received by its articles that contextualizes the journal's prestige and popularity within the academic community.

The list of top journals was generated due to the observation that many leading journals had extensive discussions on how to use GAI or relevant tools in scholarly publishing, the inclusion of which may provide rich information on such guidelines.^{7,33} To obtain a list of top journals, we selected a representative journal with the highest SJR score from each journal family among the first 200 ranked journals. This is because journals within the same family

usually adopt identical GAI usage guidelines. For example, the Lancet and other 21 Lancet-affiliated journals were all among the first 200 ranked journals and adopted identical guidelines.

The list of whole-spectrum sample journals was generated to comprehensively examine the current state of GAI usage guidelines across medical journals. For this journal list, we randomly selected two journals from every 100 journals on the ranking list. Non-English journals were excluded, and journals of the same family were replaced. As a result, we included 98 journals for top journals and 144 journals for whole-spectrum sample journals.

142

143 **Data collection and extraction**

144 For each journal selected above, two members of our research team independently carried out a thorough examination of the guidelines, which were publicly available on each journal's 145 146 website. This was to identify any content related to the GAI usage. The following information was extracted to create two datasets for top journals and whole-spectrum sample journals 147 respectively: journal characteristics and three types of GAI usage guidelines (Table 1). Journal 148 characteristics included three items: journal name, SJR score, and region. Three types of GAI 149 150 usage guidelines included author guidelines, reviewer guidelines, and references to external 151 guidelines. All the data extraction was completed between December 2023 and January 2024. Any disagreements were resolved through group discussions. 152

153 For author guidelines, we identified eight requirements: usage permission, language editing, manuscript writing, data analysis and interpretation, image generating, fact-checking 154 155 requirement, usage documentation, and authorship eligibility. Usage permission refers to 156 whether GAI tools are allowed to be used in manuscript preparation. Language editing, manuscript writing, data analysis and interpretation, and image generating refer to whether GAI 157 158 tools are allowed to be used in each of these manuscript development steps. Fact-checking requirement refers to whether authors are required to check and verify AI-generated content. 159 160 Usage documentation refers to whether authors are required to disclose and document the use of GAI tools. Authorship eligibility refers to whether AI owns authorship and is allowed to be listed 161 162 as an author. Journals providing any of these eight requirements were marked as providing author guidelines. 163

For reviewer guidelines, we identified three requirements: usage permission, language editing, and usage documentation. Journals providing any of these three requirements were marked as providing reviewer guidelines. We further coded each of the 11 requirements for author and reviewer guidelines as 'Yes', 'No', or 'Not Reported'.

For each journal, we generated a specificity score for author guidelines (range: 0 to 8), reviewer guidelines (range: 0 to 3), and author and reviewer guidelines combined (range: 0 to 11), respectively. Those scores were derived from the above coding: a score of '1' was assigned for either 'Yes' or 'No' to each requirement; and a score of '0' for "Not Reported". A higher specificity score indicated a higher specificity level.

Many journals referred to external guidelines formulated by publishing organizations or publishers for authors and reviewers. For these external guidelines on GAI usage, we focused on standards formulated by four groups: the Committee on Publication Ethics (COPE), ICMJE, the World Association of Medical Editors (WAME), and publishers. We coded the reference to each of these four external guidelines as 'Yes' or 'Not Reported'. We did not examine the specificity score of external guidelines because they were not journal-specific.

179

180 Statistical analysis

We summarized and compared the characteristics and provision of different usage guidelines between top journals and whole-spectrum sample journals. We used Probit regression to examine relationships between journal characteristics and the provision of author or reviewer guidelines and references to external guidelines. We used linear regression to relationships between journal characteristics and the specificity score of these guidelines. We conducted all analyses using R 4.2. All data used for this study was provided in **Supplementary Tables 1 and 2**.

188

189 Results

Table 2 summarizes journal characteristics and the provision of different GAI usage
guidelines. The 98 top journals had a median SJR score of 4.2, and 95 of them (96.9%) were
based in either Northern America or Western Europe. The 144 whole-spectrum sample journals

193 had a median SJR score of 0.5, and they were more evenly distributed across the three regions. 194 The top journals were more likely to provide different types of GAI usage guidelines than whole-195 spectrum sample journals, including author guidelines (64.3% vs. 27.8%), reviewer guidelines (11.2% vs. 0.0%), and references to external guidelines (85.7% vs 74.3%). A detailed summary 196 197 of the provision and specificity of GAI usage guidelines is presented for top journals and wholespectrum sample journals in **Supplementary Table 3**, including eight requirements for authors, 198 199 three requirements for reviewers, and four frequently referred external guidelines. Compared to whole-spectrum sample journals, top journals had a higher specificity level of author guidelines, 200 201 reviewer guidelines, and the two combined.

Relationships between journal characteristics and the provision of GAI usage guidelines 202 among top journals differed from those among whole-spectrum sample journals (Table 3). 203 204 Among top journals, the provision of these guidelines generally showed no significant 205 differences comparing journals with varying SJR scores or those based in different regions. However, journals based in Western Europe were more likely to refer to external guidelines than 206 those based in Northern America (2.07, 95% CI: 1.07 to 4.26). Among whole-spectrum sample 207 journals, journals with a higher SJR score were more likely to provide author guidelines (0.85, 208 209 95% CI: 0.49 to 1.25) and refer to external guidelines (2.01, 95% CI: 1.24 to 3.65). Compared to journals based in Northern America, those in Western Europe had a similar likelihood of 210 211 providing author guidelines (0.02, 95% CI: -0.51 to 0.55), whereas journals in other regions were less likely to do so (-1.03, 95% CI: -1.67 to -0.43). Since no reviewer guidelines were provided 212 by whole-spectrum sample journals, no association estimates were obtained. 213

Figure 2 shows the specificity level of author and reviewer guidelines and detailed 214 215 requirements of different guidelines among top journals. The specificity score was determined by 216 how many of the 11 requirements of author and reviewer guidelines were provided. For author 217 guidelines, a consensus emerged on five requirements whenever provided: permitting GAI usage, allowing for language editing, requiring fact-checking, mandating usage documentation, and 218 219 prohibiting AI authorship. However, disagreements were observed about whether GAI can be used for writing manuscripts, analyzing and interpreting data, and generating images. For 220 221 reviewer guidelines, fewer journals provided requirements, and there were disagreements on permitting GAI usage and allowing for language editing. Among external guidelines, the ICMJE 222 223 guideline was the most frequently referred to. Details of external guidelines are summarized in

Supplementary Table 4. Among whole-spectrum sample journals, similar patterns of
 requirements were observed except that a lower proportion of journals provided specific
 requirements (Supplementary Figure 1).

Relationships between journal characteristics and specificity scores of author and reviewer guidelines were examined (**Supplementary Table 5**). Among top journals, the journal's SJR score was not associated with the specificity level of usage guidelines. In contrast, whole-spectrum sample journals with a higher SJR score exhibited a higher specificity level. Among both groups of journals, those based in Western Europe had a similar level of specificity as those in North America, whereas journals from other regions had a lower specificity level. Relationships for author guidelines showed similar patterns among top journals.

234

235 Discussion

236 Our study systematically examined GAI usage guidelines for scholarly publishing across medical journals, including author guidelines, reviewer guidelines, and references to external 237 guidelines. Around two-thirds of top journals provided author guidelines while less than one-238 third of whole-spectrum journals did so, highlighting differential acknowledgment and 239 240 integration of GAI practices and room for improvement across medical journals. The extremely low proportion of both top and whole-spectrum sample journals providing reviewer guidelines 241 showed their negligence in setting guidance on the use of such tools for reviewers. The high 242 243 proportion of both journal groups referred to external guidelines showed the critical role of these external standards in fostering ethical practices of GAI usage. We also identified journal 244 characteristics that were associated with the provision and specificity level of different guidelines. 245 246 Together, our results suggest the urgent need for explorations and improvements of GAI usage guidelines in scholarly publishing. 247

Among medical journals across the whole spectrum, only a small proportion of journals (28%) provided GAI usage guidelines for authors, highlighting a serious lack of alertness in guiding the proper use of such tools. Lack of proper guidance can lead to unethical practices or allow AI-generated manuscripts to infiltrate publications undetected, jeopardizing the quality, integrity, and transparency of research.^{34–36} Interestingly, the figure we observed is considerably lower than that reported by another study (86%) focusing on only top journals.³⁰ Moreover, the

254 provision of author guidelines among top medical journals (64%) is also lower in our study compared to that study. The main reason for the discrepancy in the provision of author guidelines 255 256 between our study and the other is the different journal selection methods. Our study selected 257 one representative journal from each journal family while the other study included multiple 258 journals of the same families. For example, among their 51 top medical journals, six were Nature family journals and three were JAMA family journals, which usually adopted the same 259 260 guidelines. This journal selection method can lead to biases in examining the provision of GAI usage guidelines. Indeed, we observed a similar proportion of journals providing author 261 guidelines among the first 200 SJR-ranked medical journals when we included multiple journals 262 of the same families. 263

264 The better provisions of author guidelines comparing top journals to whole-spectrum 265 journals can be attributed to several factors. They include more accessible advanced AI technology in Western Europe and Northern America, stringent regulatory environments, and 266 robust scholarly communication networks.^{37–39} Author guidelines provided by top journals are 267 also more specific than others, evidenced by two observations. First, top journals had a much 268 269 higher specificity level than whole-spectrum sample journals; second, among whole-spectrum 270 sample journals, journals with a higher SJR score tend to be more specific. This greater specificity level can help authors adhere to ethical standards and prevent potential misconduct, 271 272 including plagiarism or misrepresentation of AI-generated content as human work. Our findings 273 showed that both top and whole-spectrum journals have significant room for improvement.

Across journals, consensus has been achieved for five of the eight requirements examined 274 in author guidelines, and disagreements were observed for three requirements related to content 275 generating applications. The remarkable capacities of GAI justify its usage in scholarly writing, 276 especially for language editing.^{40–43} However, its risks threatening research integrity require 277 authors to conduct fact-checking and usage documentation but not to grant such tools 278 authorship.⁴⁴ These agreements show how the scientific community is embracing GAI in a 279 cautious way. For example, the emphasis on fact-checking and usage documentation is necessary 280 for preventing fabrication and potential biases and promoting transparency and trust.⁴⁵ 281

Notably, journals differ in instructions for usage documentation on what should bedisclosed. While some journals require authors to report any use of GAI, other journals allow

284 authors not to do so if such tools were solely used for language editing. This may be explained 285 by journals' different interpretations of GAI applications. Together with two other recent studies 286 showing inconsistencies in what to disclose of GAI usage, a need is highlighted surrounding standardization of GAI usage disclosure and documentation.^{30,31,46} More importantly, journals 287 288 held different and even opposite stances for content-generating applications, including manuscript writing, data analysis and interpretation, and image generating. This is due to concern 289 290 centered on plagiarism because it is challenging to trace and verify the originality of AIgenerated content.^{47,48} 291

As the first study that systematically examined GAI usage guidelines for reviewers, we 292 are surprised to find that an extremely small proportion of journals did so. For example, only 10% 293 294 of top journals and none of the whole-spectrum sample journals provided reviewer guidelines. 295 Given the reviewers' pivotal role in safeguarding the quality and integrity of scholarly work, this 296 negligence reflects a concerning gap in guiding the use of such tools in the peer review process. 297 The misuse of GAI in peer review can have multiple negative consequences, including the breach of private data and violation of intellectual property. These are important issues 298 considering the sensitive nature of health data and medical research.²² 299

Among the merely available reviewer guidelines identified, we observed disagreements on requirements regarding usage permission, language editing, and usage documentation across journals. We interpreted these disagreements as journals' different stances on how to protect the confidentiality of manuscripts.^{49,50} Some journals specified conditions in which GAI can be used. For example, they required reviewers to obtain permission from authors and editors, to confirm that manuscripts shall not be used as training data, and provide authors with the choice to opt for or against a GAI-assisted review process.^{26,51}

Around 90% of top journals and three quarters of whole-spectrum sample journals have referred to external guidelines on GAI usage formulated by ICMJE, COPE, WAWE, or publishers. This underscores a collective endeavor to establish ethical benchmarks for GAI usage in scholarly publishing. Notably, top journals based in Western Europe were more likely to refer to external standards than their Northern American counterparts. This may stem from Western Europe's pioneering regulations on GAI usage.^{52,53} Importantly, we identified conflicts between journal-specific guidelines and referred external guidelines for at least five of the top journals we

selected. For example, *Molecular Aspects of Medicine* prohibits content-generating applications,
 while its referred external guideline, WAME, permits such usage. These conflicts have also been
 reported by another study.³⁰ To avoid confusion, we call for attention to resolve such conflicts.

317 We identified three main aspects of GAI usage guidelines that need urgent improvements. First, it is important for journals to provide specific GAI usage guidelines for both authors and 318 reviewers rather than only refer to external guidelines. Journal-specific guidelines can better 319 320 ensure the relevance and effectiveness in guiding GAI usage in their unique scholar environment. 321 Second, details of the guidelines should be specified to avoid confusion, especially for contentgenerating applications. Our systematic evaluation of available GAI usage guidelines provides a 322 roadmap to clarify expectations on proper GAI practices. Third, disagreements identified across 323 324 guidelines warrant special attention. Although these disagreements highlight the continuously 325 evolving use of GAI, it is possible to establish a framework of proper practices and specify 326 requirements under different scenarios. For example, journals may consider standardizing usage 327 documentation by integrating questions into the manuscript submission system; ascertaining whether GAI was used, and if so, detailing specific purposes and ensuring accountability. 328 Besides, it is essential to regularly review, discuss, and update GAI usage guidelines. 329

Our study has several strengths. It pioneered a systematic and quantitative examination of 330 331 GAI usage guidelines across top and whole-spectrum medical journals, which provides an 332 overview of the current regulations in scholarly publishing. It also synthesized existing GAI 333 usage guidelines for both authors and reviewers and established a systematic checklist. We 334 further provided recommendations aimed at improving existing GAI usage guidelines for scholarly publishing and promoting the proper use of GAI tools. While our study focused on 335 medical journals, we believe our findings and recommendations can be applied to journals in 336 other disciplines. 337

Our study also has limitations. First, for whole-spectrum sample journals, we sampled 2% of the 7,174 medical journals with an SJR score and included only English journals. Despite this, our use of stratified random sampling enables the representation of all medical journals listed in the SJR database and allows for timely communication of our important findings. Second, the scarce provision of reviewer guidelines observed in our findings might be attributed to the potential non-public accessibility of reviewer guidelines, which we think is uncommon. Third,

- our analysis did not include guidelines for editors, despite their relevance and importance, due to
- their scarce availability across the journals we examined. In the future, we would like to
- collaborate with different journals to conduct such examinations.
- 347

348 Conclusions

- 349 The provision of GAI usage guidelines is limited across medical journals, especially for
- 350 reviewer guidelines. The lack of specificity and inconsistencies in existing guidelines highlight
- areas deserving improvement. Immediate attention is needed to guide GAI usage in scholarly
- 352 publishing and safeguard the integrity and trust of medical research.

- 353 Corresponding Author: Chenfei Ye, Chihua Li.
- **Author Contributions:** Li had full access to all of the data in the study and took responsibility
- for the integrity of the data and the accuracy of the data analysis.
- 356 *Concept and design:* Yin, Li.
- 357 *Acquisition, analysis, or interpretation of data:* All authors.
- 358 *Drafting of the manuscript:* Yin, Li.
- 359 *Critical review of the manuscript for important intellectual content:* All authors.
- 360 *Statistical analysis:* All authors.
- 361 *Obtained funding:* Ye, Li.
- 362 *Administrative, technical, or material support:* Yin, Li.
- 363 *Supervision:* Li.
- 364 **Conflict of Interest Disclosures:** None.
- **Funding/Support:** Chenfei Ye is supported by grants from the National Natural Science
- Foundation of P.R. China (62106113); Guangdong Basic and Applied Basic Research
- Foundation (2023A1515010792), Basic Research Foundation of Shenzhen Science and
- Technology Stable Support Program (GXWD 20231129121139001).
- **Role of the Funder/Sponsor:** The funders had no role in the design and conduct of the study;
- 370 collection, management, analysis, and interpretation of the data; preparation, review, or approval
- of the manuscript; and decision to submit the manuscript for publication.
- 372 **Data Sharing Statement:** All data relevant to the study are included in the article or included as
- 373 supplementary materials. Full data and analytical code are available from the corresponding
- author at chihuali@umich.edu.

375 **References**

- 1. OpenAI. Introducing ChatGPT. Accessed December 1, 2023. https://openai.com/blog/chatgpt
- Hern A. AI bot ChatGPT stuns academics with essay-writing skills and usability. *The Guardian*. Accessed December 4, 2022.
- https://www.theguardian.com/technology/2022/dec/04/ai-bot-chatgpt-stuns-academics-with essay-writing-skills-and-usability
- 381 3. Stokel-Walker C. AI bot ChatGPT writes smart essays should professors worry? *Nature*.
 382 Accessed December 9, 2022. doi:10.1038/d41586-022-04397-7
- 4. Editorial. ChatGPT: friend or foe? *Lancet Digit Health*. 2023;5(3):e102. doi:10.1016/S2589 7500(23)00023-7
- 5. 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. doi:10.1001/jama.2023.1344
- 6. Watkins R. Guidance for researchers and peer-reviewers on the ethical use of Large Language
 Models (LLMs) in scientific research workflows. *AI Ethics*. Accessed May 16, 2023.
 doi:10.1007/s43681-023-00294-5
- 391 7. Biswas S. ChatGPT and the future of medical writing. *Radiology*. 2023;307(2):e223312.
 392 doi:10.1148/radiol.223312
- 8. Hosseini M, Horbach SPJM. 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(1):4. doi:10.1186/s41073-023-00133-5
- 396 9. Van Noorden R, Perkel JM. AI and science: what 1,600 researchers think. *Nature*.
 397 2023;621(7980):672-675. doi:10.1038/d41586-023-02980-0
- Thirunavukarasu AJ, Ting DSJ, Elangovan K, et al. Large language models in medicine. *Nat Med.* 2023; 29(8): 1930-1940. doi:10.1038/s41591-023-02448-8
- 400 11. AI will transform science now researchers must tame it. *Nature*. 2023;621(7980):658-658.
 401 doi:10.1038/d41586-023-02988-6
- 402 12. Editorial. Prepare for truly useful large language models. *Nat Biomed Eng.* 2023;7(2):85-86.
 403 doi:10.1038/s41551-023-01012-6
- 404 13. Editorial. Not a generative AI–generated Editorial. *Nat Cancer*. 2023;4(2):151-152.
 405 doi:10.1038/s43018-023-00531-0
- 406 14. Owens B. How Nature readers are using ChatGPT. *Nature*. 2023;615(7950):20-20.
 407 doi:10.1038/d41586-023-00500-8

- 408 15. Messeri L, Crockett MJ. Artificial intelligence and illusions of understanding in scientific
 409 research. *Nature*. 2024;627(8002):49-58. doi:10.1038/s41586-024-07146-0
- 410 16. Donker T. The dangers of using large language models for peer review. *Lancet Infect Dis.*411 2023;23(7):781. doi:10.1016/S1473-3099(23)00290-6
- 412 17. Zheng H, Zhan H. ChatGPT in Scientific Writing: A Cautionary Tale. *Am J Med*.
 413 2023;136(8):725-726.e6. doi:10.1016/j.amjmed.2023.02.011
- 414 18. Hwang SI, Lim JS, Lee RW, et al. Is ChatGPT a "Fire of Prometheus" for Non-Native
 415 English-Speaking Researchers in Academic Writing? *Korean J Radiol*. 2023;24(10):952.
 416 doi:10.3348/kjr.2023.0773
- 417 19. Van Dis EAM, Bollen J, Zuidema W, Van Rooij R, Bockting CL. ChatGPT: five priorities
 418 for research. *Nature*. 2023;614(7947):224-226. doi:10.1038/d41586-023-00288-7
- 20. Thorp HH. ChatGPT is fun, but not an author. *Science*. 2023;379(6630):313.
 doi:10.1126/science.adg7879
- 421 21. Stokel-Walker C. ChatGPT listed as author on research papers: many scientists disapprove.
 422 *Nature*. 2023;613(7945):620-621. doi:10.1038/d41586-023-00107-z
- 423 22. National Institutes of Health. The use of generative artificial intelligence technologies is
 424 prohibited for the NIH peer review process. Accessed December 21, 2023.
 425 https://grants.nih.gov/grants/guide/notice-files/NOT-OD-23-149.html
- 426 23. Liverpool L. AI intensifies fight against 'paper mills' that churn out fake research. *Nature*.
 427 2023;618(7964):222-223.
- 428 24. Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for authors, peer reviewers,
 429 and editors on use of AI, language models, and chatbots. *JAMA*. 2023;330(8):702-703.
 430 doi:10.1001/jama.2023.12500
- 431 25. Committee on Publication Ethics. Authorship and AI tools. Accessed December 8, 2023.
 432 https://publicationethics.org/cope-position-statements/ai-author
- 26. Leung TI, De Azevedo Cardoso T, Mavragani A, Eysenbach G. Best practices for using AI
 tools as an author, peer reviewer, or editor. *J Med Internet Res.* 2023;25:e51584.
- 435 doi:10.2196/51584
- 436 27. International Committee of Medical Journal Editors. Defining the role of authors and437 contributors. Accessed December 8, 2023.
- 438 https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-
- 439 of-authors-and-contributors.html
- 440 28. Elsevier. Publishing ethics. Accessed December 8, 2023.
- 441 https://www.elsevier.com/about/policies-and-standards/publishing-ethics#0-publishing-ethics

- 442 29. The Lancet Peer Review. Accessed December 10, 2023. https://www.thelancet.com/peer 443 review
- 30. Ganjavi C, Eppler MB, Pekcan A, Biedermann B, Abreu A, Collins GS, Gill IS, Cacciamani
 GE. Publishers' and journals' instructions to authors on use of generative artificial intelligence
 in academic and scientific publishing: bibliometric analysis. *BMJ*. 2024;384:e077192.
- 447 doi:10.1136/bmj-2023-077192
- 31. Tang A, Li K, Kwok KO, Cao L, Luong S, Tam W. The importance of transparency:
 declaring the use of generative artificial intelligence (AI) in academic writing. *J Nurs Scholarsh.* 2024;56(2):314-318. doi:10.1111/jnu.12938
- 451 32. SciMago. Journal and country rank. Accessed December 1, 2023.
 452 https://www.scimagojr.com/journalrank.php
- 453 33. Editorial. Tools such as ChatGPT threaten transparent science; here are our ground rules for
 454 their use. *Nature*. 2023;613(7945):612. doi:10.1038/d41586-023-00191-1
- 455 34. Prillaman M. "ChatGPT detector" catches AI-generated papers with unprecedented accuracy.
 456 *Nature*. Accessed November 06, 2023. doi:10.1038/d41586-023-03479-4
- 457 35. Misra DP, Chandwar K. ChatGPT, artificial intelligence and scientific writing: what authors,
 458 peer reviewers and editors should know. *J R Coll Physicians Edinb*. 2023;53(2):90-93.
 459 doi:10.1177/14782715231181023
- 36. Conroy G. Scientific sleuths spot dishonest ChatGPT use in papers. *Nature*. Accessed
 September 8, 2023. doi:10.1038/d41586-023-02477-w
- 37. Roche C, Wall PJ, Lewis D. Ethics and diversity in artificial intelligence policies, strategies
 and initiatives. *AI Ethics*. 2023;3(4):1095-1115. doi:10.1007/s43681-022-00218-9
- 38. Fjeld J, Achten N, Hilligoss H, Nagy A, Srikumar M. Principled artificial intelligence:
 mapping consensus in ethical and rights-based approaches to principles for AI. *Berkman Klein Cent Res Publ.* Accessed January 15, 2020. doi:10.2139/ssrn.3518482
- 39. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell*.
 2019;1(9):389-399. doi:10.1038/s42256-019-0088-2
- 40. Berdejo-Espinola V, Amano T. AI tools can improve equity in science. *Science*. 2023 Mar 10;379(6636):991. doi: 10.1126/science.adg9714.
- 41. Yan M, Cerri GG, Moraes FY. ChatGPT and medicine: how AI language models are shaping
 the future and health related careers. *Nat Biotechnol*. 2023;41(11):1657-1658.
 doi:10.1038/s41587-023-02011-3
- 474 42. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific
 475 review articles. *Am J Cancer Res.* 2023;13(4):1148-1154.

43. Li H, Moon JT, Purkayastha S, Celi LA, Trivedi H, Gichoya JW. Ethics of large language
models in medicine and medical research. *Lancet Digit Health*. 2023;5(6):e333-e335.
doi:10.1016/S2589-7500(23)00083-3

- 44. Bell S. The write algorithm: promoting responsible artificial intelligence usage and
 accountability in academic writing. *BMC Med*. 2023;21(1):334. doi:10.1186/s12916-02303039-7
- 482 45. Vintzileos AM, Chavez MR, Romero R. A role for artificial intelligence chatbots in the
 483 writing of scientific articles. *Am J Obstet Gynecol*. 2023;229(2):89-90.
 484 doi:10.1016/j.ajog.2023.03.040
- 46. Beall J. Predatory publishers are corrupting open access. *Nature*. 2012;489(7415):179-179.
 doi:10.1038/489179a
- 487 47. Lund BD, Wang T, Mannuru NR, Nie B, Shimray S, Wang Z. ChatGPT and a new academic
 488 reality: artificial intelligence written research papers and the ethics of the large language
 489 models in scholarly publishing. *J Assoc Inf Sci Technol*. 2023;74(5):570-581.
 490 doi:10.1002/asi.24750
- 491 48. Dehouche N. Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3).
- 492 *Ethics Sci Environ Polit*. 2021;21:17-23.
- 493 49. Hosseini M, Horbach SPJM. Fighting reviewer fatigue or amplifying bias? considerations
 494 and recommendations for use of ChatGPT and other large language models in scholarly peer
 495 review. *Res Integr Peer Rev.* 2023;8(1):4. doi:10.1186/s41073-023-00133-5
- 50. Al-Hawawreh M, Aljuhani A, Jararweh Y. Chatgpt for cybersecurity: practical applications,
 challenges, and future directions. *Clust Comput.* 2023;26(6):3421-3436. doi:10.1007/s10586023-04124-5
- 499 51. Flanagin A, Kendall-Taylor J, Bibbins-Domingo K. Guidance for Authors, Peer Reviewers,
 500 and Editors on Use of AI, Language Models, and Chatbots. *JAMA*. 2023;330(8):702.
 501 doi:10.1001/jama.2023.12500
- 502 52. Li Z. Why the European AI Act transparency obligation is insufficient. *Nat Mach Intell*.
 503 2023;5(6):559-560. doi:10.1038/s42256-023-00672-y
- 53. Roberts H, Cowls J, Hine E, et al. Governing artificial intelligence in China and the
 European Union: Comparing aims and promoting ethical outcomes. *Inf Soc.* 2023;39(2):79-97.
 doi:10.1080/01972243.2022.2124565

Figure 1. Flowchart of journal selection



Footnote: * Journals from the same family most likely adopted identical guidelines for GAI usage. # Non-English journals were excluded; Journals from the same family were replaced.



Figure 2. Specificity level of author and reviewer guidelines and requirements of different GAI usage guidelines among top journals

Footnote: The specificity score displays how many requirements were provided for author and reviewer guidelines combined (range: 0-11).

	Journal characteristics	Three types of GAI usage guidelines				
Items		Author guidelines (Y/N/NR)	Reviewer guidelines (Y/N/NR)	References to external guidelines (Y/NR)		
1	Journal name	Usage permission	Usage permission	COPE		
2	SJR score	Language editing	Language editing	ICMJE		
3	Region	Manuscript writing	Usage documentation	WAME		
4		Data analysis and interpretation		Publishers		
5		Image generating				
6		Fact-checking requirement				
7		Usage documentation				
8		Authorship eligibility				

Table 1. Journal characteristics and three types of generative artificial intelligence (GAI) usage guidelines extracted

Footnote: Abbreviations: Y, Yes; N, No; NR, Not Reported;

COPE, Committee on Publication Ethics; ICMJE, International Committee of Medical Journal Editors; WAME, World Association of Medical Editors.

	Top journals (n = 98)	Whole-spectrum sample journals (n = 144)	P value [#]	
Journal characteristics				
SJR score, median (IQR)	4.2 (3.6, 6.7)	0.5 (0.2, 0.9)	< 0.01	
Region, n (%)				
Northern America	50 (51.0)	39 (27.1)		
Western Europe	45 (45.9)	51 (35.4)	< 0.01	
Other regions*	3 (3.1)	54 (37.5)		
Provision of different usage guidelines				
Author guidelines, n (%)	63 (64.3)	40 (27.8)	< 0.01	
Reviewer guidelines, n (%)	11 (11.2)	0 (0.0)	< 0.01	
References to external guidelines, n (%)	84 (85.7)	107 (74.3)	< 0.05	

Table 2. Journal characteristics and provision of different GAI usage guidelines

Footnote: *Other regions included Africa, Asiatic region, Eastern Europe, Latin America, Middle East, and Pacific region.

Differences between the two groups of journals were analyzed using Wilcoxon Rank Sum test or Chi-Square tests.

	Top journals			Whole-spectrum sample journals				
	Journal characteristics	Coefficient	95% CI	P value	Coefficient	95% CI	P value	
		Provision of any author guidelines						
Model 1	SJR score	-0.01	(-0.04, 0.02)	0.58	0.85	(0.49, 1.25)	< 0.01	
	Northern America	Ref	-	-	Ref	-	-	
Model 2	Western Europe	0.02	(-0.50, 0.54)	0.95	0.02	(-0.51, 0.55)	0.94	
	Other regions	-5.62	(NA, 22.74)	0.98	-1.03	(-1.67, -0.43)	< 0.01	
			Provision of any reviewer guidelines					
Model 1	SJR score	0.00	(-0.06, 0.03)	1.00	-	-	-	
	Northern America	Ref	-	-	-	-	-	
Model 2	Western Europe	-0.05	(-0.72, 0.62)	0.89	-	-	-	
	Other regions	-4.22	(NA, 41.32)	0.99	-	-	-	
		Any reference to external guidelines						
Model 1	SJR score	1.00	(0.97, 1.06)	0.85	2.01	(1.24, 3.65)	< 0.05	
	Northern America	Ref	-	-	Ref	-	-	
Model 2	Western Europe	2.07	(1.07, 4.26)	< 0.05	1.53	(0.87, 2.70)	0.14	
	Other regions	101.92	(0.00, NA)	0.99	1.24	(0.72, 2.14)	0.44	

Table 3. Probit regression analysis of the relationship between journal characteristics and the provision of different GAI usage guidelines

Footnote: NA (not available) indicates that the 95% CI hits infinity due to small sample size. No association estimates were obtained for the provision of reviewer guidelines in whole-spectrum sample journals because none of these journals provided reviewer guidelines.