Artificial Intelligence-Enabled "Research Assistants" in HEOR – Hire or Fire?

Laura Dormer
Co-Founder & Editorial Director
Becaris Publishing, UK
laura@becaris.com

AI in scholarly publishing

Authors

- Writing
- Journal selection

Editors/Reviewers/ Publishers

- Article screening & quality control
- Alternate journal advice
- Peer review
- Copy editing & reference checking
- Metadata creation/enrichment

Readers/Publishers

- Summarization
- Collections
- Post-publication quality/impact evaluation
- Accessibility

Authors

Writing

- Language and grammar checking
- Improving readability
- Summarizing content to create abstracts

Journal selection

- Upload a title/abstract/keywords into a platform to receive journal recommendations
- Publisher- and non-publisher-specific tools
- General chatbot tools are also starting to be used in this way

Plain Language Summaries

Journal of Comparative Effectiveness Research

Entered into ChatGPT:

Can you re-write this abstract in plain language?:

Aim: Network meta-analyses (NMAs) increasingly feature time-varying hazards to account for non-proportional hazards between different drug classes. This paper outlines an algorithm for selecting clinically plausible fractional polynomial NMA models. **Methods:** The NMA of four immune checkpoint inhibitors (ICIs) + tyrosine kinase inhibitors (TKIs) and one TKI therapy for renal cell carcinoma (RCC) served as case study. Overall survival (OS) and progression free survival (PFS) data were reconstructed from the literature, 46 models were fitted. The algorithm entailed a-priori face validity criteria for survival and hazards, based on clinical expert input, and predictive accuracy against trial data. Selected models were compared with statistically best-fitting models. **Results:** Three valid PFS and two OS models were identified. All models overestimated PFS, the OS model featured crossing ICI + TKI versus TKI curves as per expert opinion. Conventionally selected models showed implausible survival. **Conclusion:** The selection algorithm considering face validity, predictive accuracy, and expert opinion improved the clinical plausibility of first-line RCC survival models.

Petersohn S, McGregor B, Klijn SL *et al*. Challenges in conducting fractional polynomial and standard parametric network meta-analyses of immune checkpoint inhibitors for first-line advanced renal cell carcinoma. *J. Comp. Eff. Res.* 12(8), e230004 (2023).

ChatGPT response:

Objective: The goal of this study was to develop a method for choosing appropriate mathematical models that describe the effects of different treatments on cancer survival rates.

Methods: We looked at the outcomes of treatments for a type of kidney cancer. We studied the survival rates of patients who received immune checkpoint inhibitors (ICIs) along with tyrosine kinase inhibitors (TKIs), and those who only received TKIs. We gathered data from various sources and used statistical techniques to create models. We considered factors like how long patients survived and how long they remained free of disease progression. We used expert opinions and data from clinical trials to decide which models were the most accurate and reasonable.

Results: We identified three models that accurately represented disease progression and two models that accurately represented overall survival. However, all models tended to predict longer disease-free periods than what was observed in reality. The model for overall survival showed that patients treated with a combination of ICIs and TKIs had different survival patterns compared to those treated with only TKIs, which matches what experts thought. In contrast, the models selected using traditional methods did not make sense when compared to real-world outcomes.

Conclusion: Our approach of selecting models based on expert opinions, real-world data, and how well they predicted outcomes improved the reliability of the models for predicting survival in patients with this type of kidney cancer when treated with different therapies.

Editors, Reviewers & Publishers

Article screening & quality control

- Technical checks
- Methodology and statistics checks
- Plagiarism detection
- Image manipulation

Alternate journal advice

- For sound, but out-of-scope articles

Editors, Reviewers & Publishers

Peer review

- Identifying appropriate peer reviewers
- Conducting peer review
- Assessing/improving the quality of peer review

Copy editing & reference checking

- Language, grammar, readability, summarizing
- Tools to check reliability of references

Metadata creation/enrichment

- Content discoverability
- Content recommendations

Readers & Publishers

Summarization

- Extracting key facts from articles
- Build your library
- Based on large databases of articles

Collections

- Categorized content for easy navigation

Quality/impact evaluation

- Identify impactful papers
- ‘Smart Citations’
- Research integrity checking

Accessibility

- Alt text for images
- Audio captions
- Language translation
- Accessibility standards

What are the risks?

Lack of transparency

- Undisclosed use of generative AI by authors
- IP and privacy infringement

Confidentiality issues

- Secure storage of uploaded data
- Access to content

Hallucinations

- Made up, but plausible sounding information

Bias

- Trained on biased data

Papermills

- Text generation
- Image fabrication

What do the guidelines say?

	ICMJE	WAME
Authorship	• Chatbots cannot be listed as authors	
Use when authoring	• Use of AI-assisted technologies should be disclosed to Editors • Authors are responsible for the submitted material	
	• AI-generated material cannot be referenced as a primary source	
Use when peer reviewing	• Reviewers must maintain confidentiality of the manuscripts they are reviewing • Use of AI-assisted technologies should be disclosed to Editors • Reviewers are responsible for the output of AI-assisted technologies	
Use during the editorial workflow		• Use of AI-assisted technologies should be disclosed to Authors • Editors are responsible for the output of AI-assisted technologies • Editors need and should have access to tools to detect AI use

What do journal policies say?

- No to authorship
- Use must be acknowledged
- Authors, reviewers and editors are responsible for any work completed using AI tools

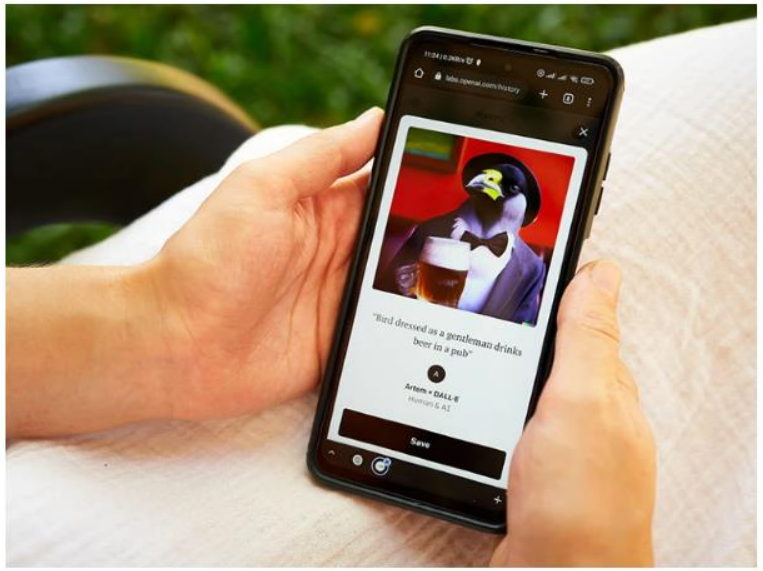
Artificial intelligence (AI). Text generated from AI, machine learning, or similar algorithmic tools cannot be used in papers published in *Science* journals, nor can the accompanying figures, images, or graphics be the products of such tools, without explicit permission from the editors. In addition, an AI program cannot be an author of a *Science* journal paper. A violation of this policy constitutes scientific misconduct.

Artificial intelligence (AI). Use of large language models and other generative AI tools is not allowed. The reviewer is responsible for writing their own review and should not solicit input from colleagues without first obtaining permission from the editor.

Science

Why *Nature* will not allow the use of generative AI in images and video

Saying 'no' to this kind of visual content is a question of research integrity, consent, privacy and intellectual-property protection.



Nature will not publish imagery created wholly or partly using generative AI. Credit: Artem Medvediev/Alamy

Hire or fire?

Hire – with care!

There is a need for:

- AI governance and detection to catch up with tools
- Cross-publisher collaboration to develop AI detection tools
- Training and education of researchers and editors on best practice for using AI tools
- Human judgement!