



## Cochrane's focus is on responsible use of AI in systematic reviews

Guidance for responsible AI use in evidence synthesis



## Focus on responsible use of AI in systematic reviews

- Encourage studies within reviews and research into AI tools
- Endorse public sharing of this research
- Define expected standards for evaluation and routes for endorsement in Cochrane processes



MRC Hub for Trials Methodology Research

Home Aims of the Network Executive of the Network Documentation Research Projects

The Northern Ireland Hub for Trials Methodology Research

Received: 23 October 2023 | Accepted: 4 January 2024  
DOI: 10.1002/tem.15041

METHODS ARTICLE

UNIVERSITY SITES / THE NORTHERN IRELAND NETWORK FOR TRIALS METHODOLOGY RESEARCH

### SWAT/SWAR Information

Studies Within a Trial (SWAT) and Studies Within a Review (SWAR)

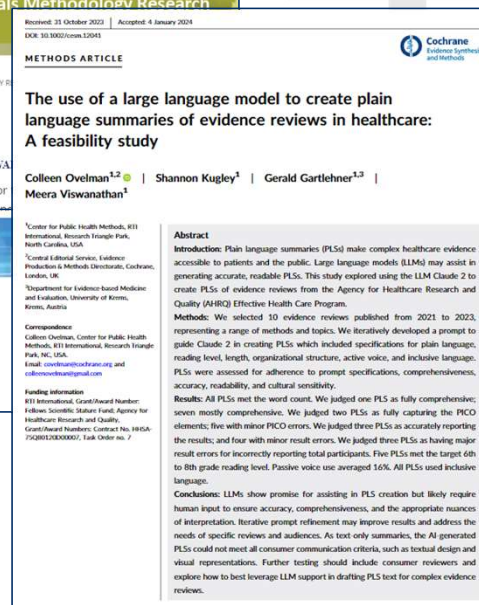
The SWAT and SWAR programme is identifying methods for

### COCHRANE EVIDENCE SYNTHESIS AND METHODS

Edited By: Michael Brown (Michigan State University, USA)

JOURNAL METRICS >

Online ISSN: 2832-9023



Received: 23 October 2023 | Accepted: 4 January 2024  
DOI: 10.1002/tem.15041

METHODS ARTICLE

### The use of a large language model to create plain language summaries of evidence reviews in healthcare: A feasibility study

Colleen Ovelman<sup>1,2</sup> | Shannon Kugley<sup>1</sup> | Gerald Gartlehner<sup>1,3</sup> | Meera Viswanathan<sup>1</sup>

<sup>1</sup>Center for Public Health Methods, RTI International, Research Triangle Park, North Carolina, USA  
<sup>2</sup>Central Editorial Services, Evidence Production & Methods Directorate, Cochrane, London, UK  
<sup>3</sup>Department for Evidence-Based Medicine and Evaluation, University of KwaZulu-Natal, Durban, South Africa

Correspondence: Colleen Ovelman, Center for Public Health Methods, RTI International, Research Triangle Park, NC, USA. Email: [covelman@cochrane.org](mailto:covelman@cochrane.org) and [covelman@rti.org](mailto:covelman@rti.org)

Funding information: RTI International, Grant/Award Number: F30CE000801; National Institutes of Health, Grant/Award Number: 1R01CE000801; National Institutes of Health, Grant/Award Number: 1R01CE000801; National Institutes of Health, Grant/Award Number: 1R01CE000801

#### Abstract

**Introduction:** Plain language summaries (PLSs) make complex healthcare evidence accessible to patients and the public. Large language models (LLMs) may assist in generating accurate, readable PLSs. This study explored using the LLM Claude 2 to create PLSs of evidence reviews from the Agency for Healthcare Research and Quality (AHRQ) Effective Health Care Program.

**Methods:** We selected 10 evidence reviews published from 2021 to 2023, representing a range of methods and topics. We iteratively developed a prompt to guide Claude 2 in creating PLSs which included specifications for plain language, reading level, length, organizational structure, active voice, and inclusive language. PLSs were assessed for adherence to prompt specifications, comprehensiveness, accuracy, readability, and cultural sensitivity.

**Results:** All PLSs met the word count. We judged one PLS as fully comprehensive; seven mostly comprehensive. We judged two PLSs as fully capturing the PCO elements; five with minor PCO errors. We judged three PLSs as accurately reporting the results; and four with minor result errors. We judged three PLSs as having major result errors for incorrectly reporting total participants. Five PLSs met the target 6th to 8th grade reading level. Passive voice use averaged 16%. All PLSs used inclusive language.

**Conclusions:** LLMs show promise for assisting in PLS creation but likely require human input to ensure accuracy, comprehensiveness, and the appropriate nuances of interpretation. Iterative prompt refinement may improve results and address the needs of specific reviews and audiences. As text-only summaries, the AI-generated PLSs could not meet all consumer communication criteria, such as textual design and visual representations. Further testing should include consumer reviewers and explore how to best leverage LLM support in drafting PLS text for complex evidence reviews.

## Use of AI in writing is permitted, but must be transparent

- AI tools cannot be credited as authors
- Authors bear full responsibility for the article's accuracy and validity
- A transparent and detailed description of the AI tools used, and content generated, is mandatory



## Summary

- Cochrane is committed to ensuring the responsible use of AI in its systematic reviews
- Generative AI opens lots of opportunities and challenges
- Looking forward to collective efforts to define new standards





**Thank you and any questions?**