

# Preferences of Artificial Intelligence Use in Systematic Literature Reviews: A Discrete Choice Experiment

MSR130

Seye Abogunrin <sup>1</sup>, Bart P.H. Slob <sup>2</sup>, Marie Lane <sup>1</sup>, Sajad Emamipour <sup>2</sup>, Piotr Twardowski <sup>2</sup>, Cornelis Boersma <sup>2,3</sup>, Jurjen van der Schans <sup>2,3</sup>

<sup>1</sup> F. Hoffmann–La Roche, Basel, Switzerland. <sup>2</sup> Health–Ecore B.V., Zeist, The Netherlands. <sup>3</sup> Unit of Global Health, Department of Health Sciences, University Medical Center Groningen, Groningen, The Netherlands.

## Take Away

Workload reduction is not the only consideration for SLR reviewers when using AI tools.

## Background

Systematic literature reviews (SLRs) are essential for synthesizing research evidence and guiding informed decision-making. However, SLRs require significant resources and substantial efforts in terms of workload. Additionally, 15% of all SLR studies need update within a year, and up to 23% within 2 years <sup>[1]</sup>. In particular, the title and abstract screening phase is often tedious and thankless, requiring multiple experts to manually evaluate numerous titles and abstracts that may not meet the inclusion criteria <sup>[2]</sup>. The introduction of artificial intelligence (AI) tools can reduce this workload. However, there is a significant doubt within the community of reviewers about the actual utility and reliability of these newly emerging AI-platforms <sup>[3]</sup>. Since the use of AI in performing SLRs is a novel phenomenon, there is a research gap regarding the preferences of authors of SLRs towards these tools. Simply choosing the tool that offers the most dramatic workload reduction might not be the best strategy <sup>[4]</sup> and there may be a need to consider other factors, e.g., sensitivity of the tool, or the required user proficiency, while making such decisions.

## Aim

This study aims to understand SLR authors’ preferences regarding the purpose and use of AI tools in conducting SLRs. It focuses on the trade-offs between AI tool reliability, the (time) investment, and related benefits these tools provide.

## Methods

A discrete choice experiment was performed among professionals performing SLRs. Key attributes for AI tools were identified through a literature review and expert consultations. These attributes included the AI tool’s role in screening, required user proficiency, sensitivity, workload reduction, and the investment needed for training. Data were collected via an online survey, where participants provided background on their education and experience and completed 13-14 choice tasks featuring AI tools with varying attributes. Statistical analysis was performed using conditional multinomial logit. An additional analysis was performed by including the demographic characteristics (such as education, experience with SLR publication and familiarity with AI) as interaction variables.

Interested in this topic?

Please scan the QR code to check out our previous publication on the inter-reviewer reliability between AI and reviewers of SLRs.



## Results

The study received responses from 187 participants with diverse experience in performing SLRs and AI use. The familiarity with AI was generally low, with 55.6% of participants being (very) unfamiliar with AI. The most notable finding is a strong resistance to fully replacing human reviewers with AI ( $p < 0.001$ ). In contrast, intermediate proficiency in AI tools is positively associated with adoption ( $p = 0.030$ ). Similarly, workload reduction is also strongly linked to adoption ( $p < 0.001$ ). Interestingly, if expert proficiency is needed for the AI, authors with more research experience in years are less likely to adopt AI ( $p = 0.008$ ).

Figure 1 : Results showing the influence of the attributes on the choice

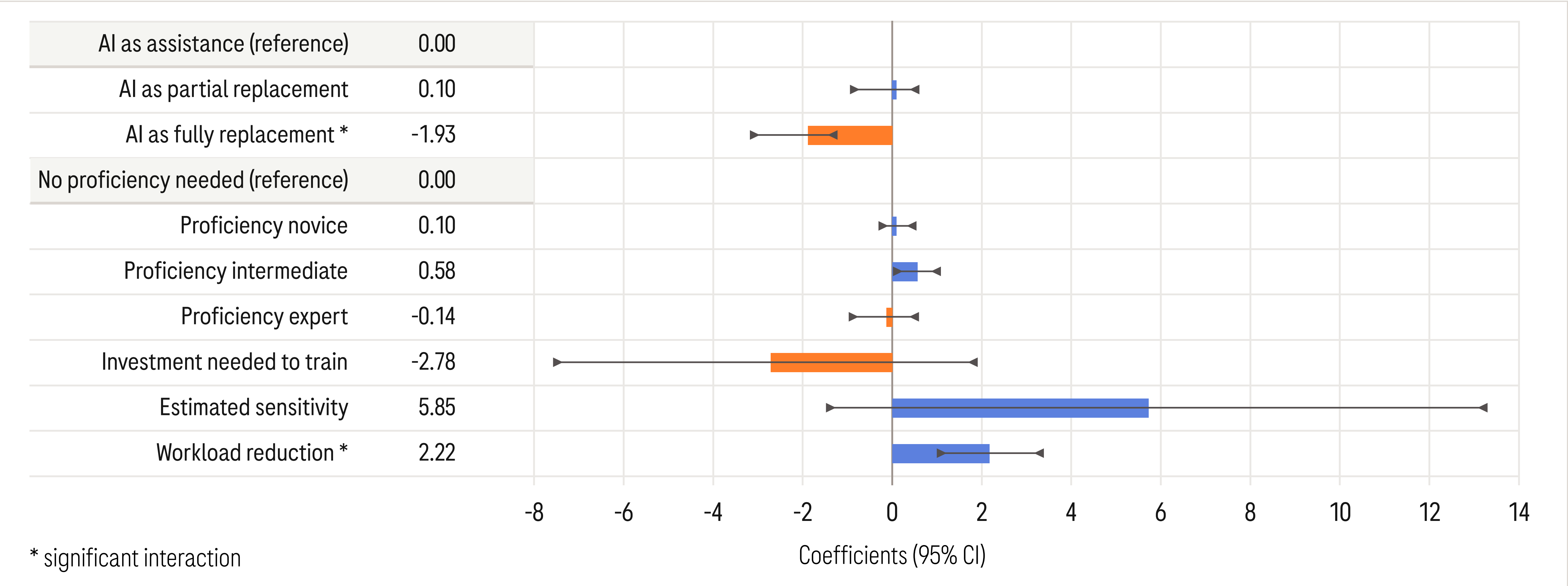
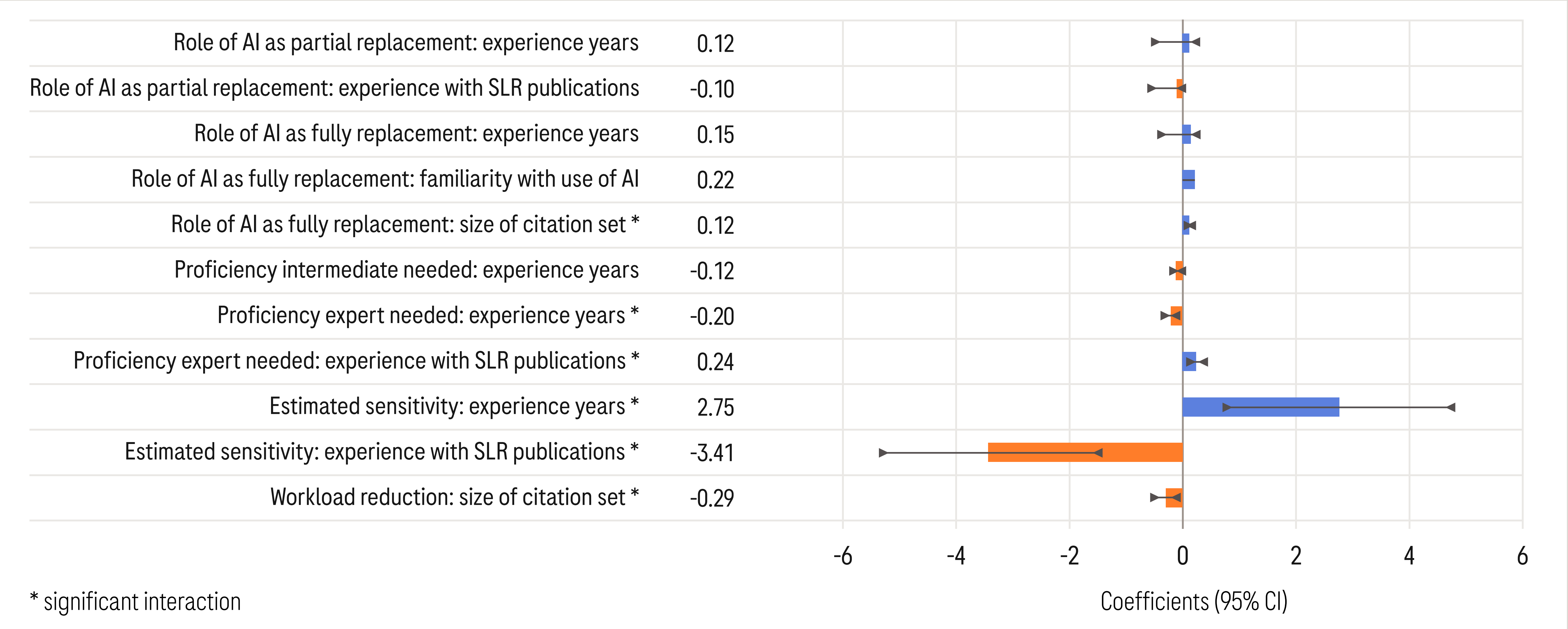


Figure 2 : Results showing the interactions of the reviewer characteristics on the attribute influence



## Strengths and limitations

To our knowledge, this is the first study to assess the preferences of SLR authors regarding AI tools in SLRs, considering in various criteria in the trade-off. We used literature and experts to construct the level of attributes and their alternatives, however, the choices constructed were hypothetical, considering the fast development of AI tools in SLRs.

## References

- Shojania KG, Sampson M, Ansari MT, Ji J, Doucette S, Moher D. How quickly do systematic reviews go out of date? A survival analysis. *Ann Intern Med*. 2007;147:224–33.
- Polanin JR, Pigott TD, Espelage DL, Grotperter JK. Best practice guidelines for abstract screening large-evidence systematic reviews and meta-analyses. *Res Synth Methods*. 2019;10:330–42.
- Blaizot A, Veettil SK, Saidoung P, Moreno-Garcia CF, Wiratunga N, Aceves-Martins M, et al. Using artificial intelligence methods for systematic review in health sciences: A systematic review. *Res Synth Methods*. 2022;13:353–62.
- Blaizot A, Veettil SK, Saidoung P, Moreno-Garcia CF, Wiratunga N, Aceves-Martins M, et al. Using artificial intelligence methods for systematic review in health sciences: A systematic review. *Res Synth Methods*. 2022;13:353–62.

## Conclusion

The key to AI adoption in SLRs is creating reliable, workload-reducing tools that assist rather than replace human reviewers, with moderate proficiency requirements and high sensitivity.