OPEN HEALTH

Smarter SLRs or Risky Shortcuts? Perspectives From Global HTA Stakeholders

Grace E. Fox (first author);
Juliette Torres Ames;
Nita Santpurkar; Jed Avissar;
Emanuele Arcà (senior author)

www.openhealthgroup.com

Smarter SLRs or Risky Shortcuts? Perspectives From Global HTA Stakeholders

Grace E. Fox (first author); Juliette Torres Ames; Nita Santpurkar; Jed Avissar; Emanuele Arcà (senior author)

Introduction

Systematic literature reviews (SLRs) are essential to health technology assessments (HTAs), offering a structured, transparent, and reproducible method for synthesizing evidence. However, conducting SLRs demands substantial time and resources. Advancements in artificial intelligence (AI) have the potential to enhance the efficiency of SLRs; however, it remains unclear how HTA agencies regard the use of AI in SLRs.

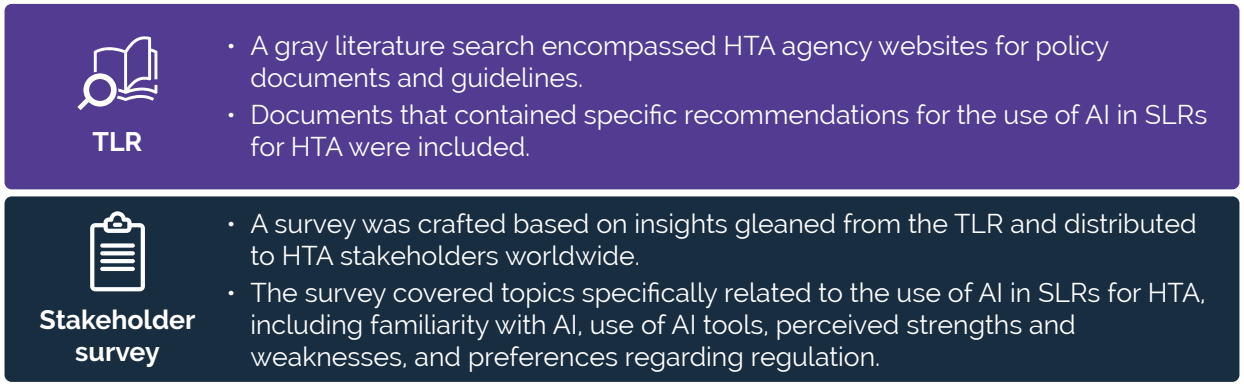
Objective

This study set out to investigate the views of HTA decision-makers about employing AI in the conduct of SLRs.

Methods

To investigate how HTA decision-makers view the application of AI in SLRs, a mixed-methods strategy was used that integrated a targeted literature review (TLR) with a stakeholder survey (**Figure 1**). The TLR sought to uncover current guidance, policy documents, and research related to AI's role in evidence synthesis for HTA submissions. Findings from the TLR were utilized to shape a survey aimed at professionals within HTA agencies. This method facilitated the gathering of both contextual insights from the literature and practical perspectives from decision-makers engaged in HTA activities. The current white paper presents the most recent update of the TLR and the latest data cut of the survey, extending and elaborating on previously disseminated findings.[1-5]

Figure 1: Mixed-methods strategy



Abbreviations: AI, artificial intelligence; HTA, health technology assessment; SLR, systematic literature review.

Results

TLR results

At the time of the TLR, only three HTA agencies — CDA-AMC (Canada's Drug Agency, Canada), NICE (National Institute for Health and Care Excellence, UK), and IQWiG (Institute for Quality and Efficiency in Health Care, Germany) — provided specific recommendations for AI usage in SLRs (**Table 1**). Guidance from CDA-AMC and NICE was deemed to be fairly comprehensive, with an emphasis on employing AI as a tool to support human reviewers, not replace them. These guidance documents stress the importance of transparency and outline several potential applications of AI, including developing search strategies, automating the selection and classification of studies, visualizing data, and, to a lesser extent, performing data extraction. CDA-AMC documents also note that the agency has developed a tool for evaluating AI-based search technologies. In contrast, IQWiG guidance at survey time was more limited. IQWiG permits the use of validated machine learning classifiers to develop search strategies and assist in study selection tasks. No formal position statements were identified from other HTA agencies.

Table 1: HTA agency guidelines on the adoption of AI-supported SLRs

HTA agency (country)	Recommendations
CDA-AMC (Canada)[6]	• ML methods and LLMs **may support:** – Generating search strategies – Automating study classification (e.g., by study design) – Screening titles/abstracts and full texts – Visualizing search results • LLMs (less established) **may be used** to: – Automate data extraction from quantitative and qualitative studies • Submitters are encouraged to follow emerging best practices (e.g., from Cochrane and GIN). • CDA-AMC has developed an **evaluation tool** for assessing AI-based search technologies.
NICE (United Kingdom)[7,8]	• AI should be used to **support**, not replace, human reviewers in the SLR process. • Use of AI must be **transparent and fully disclosed** in HTA submissions. • ML and LLMs **may be used** to: – Develop search strategies – Automate study selection and classification – Assist with data visualization – (Less established) Automate data extraction
IQWiG (Germany)[9]	• Validated ML classifiers (e.g., for identifying RCTs) **may be used** to develop search strategies. • ML tools **may support** study selection processes.

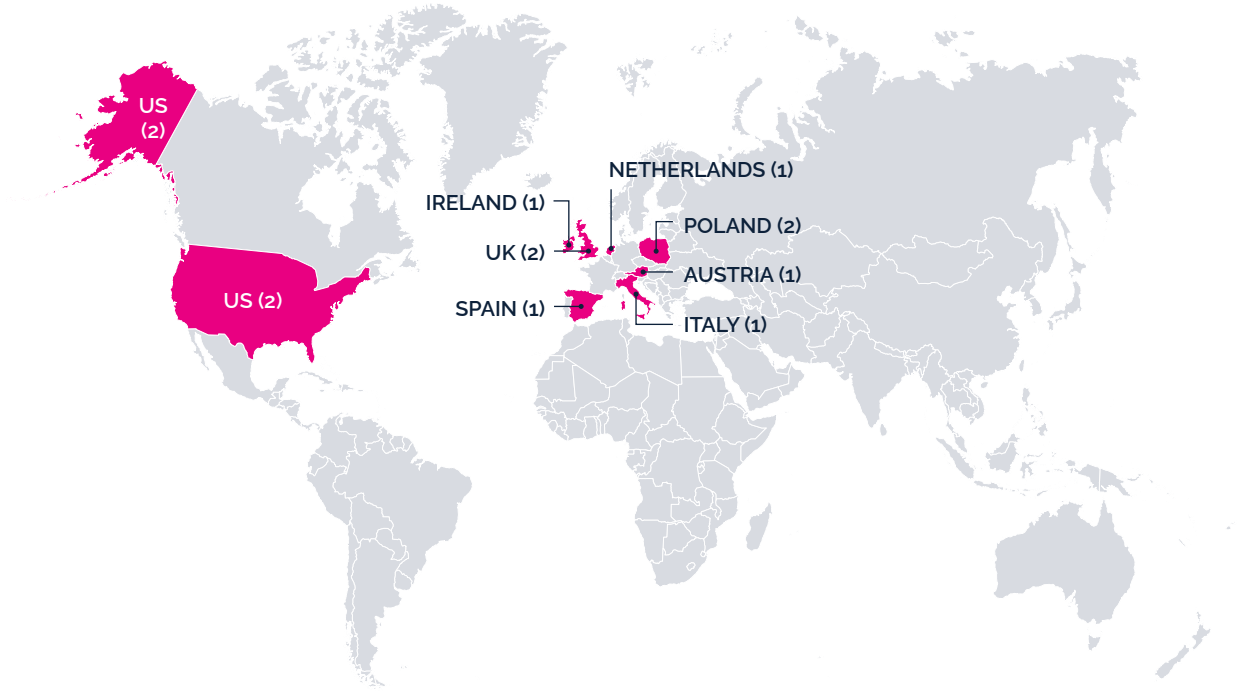
Abbreviations: AI, artificial intelligence; CDA-AMC, Canada's Drug Agency; GIN, Guidelines International Network; HTA, health technology assessment; IQWiG, Institute for Quality and Efficiency in Health Care; LLM, large language model; ML, machine learning; NICE, National Institute for Health and Care Excellence; RCT, randomized controlled trial; SLR, systematic literature review.

Survey results

Demographics

Eleven HTA stakeholders, referred to as "respondents" in this white paper, represented eight different nations, with two respondents each from Poland, the UK, and the US, and one respondent each from Austria, Ireland, Italy, the Netherlands, and Spain. **Figure 2** illustrates the countries represented. On average, these respondents had more than 14 years of professional experience in HTA.

Figure 2: Geographic distribution of respondents



Theme - Familiarity

Respondents were asked to characterize their level of familiarity with the use of AI in the creation of SLRs. Most respondents (n=6) reported being somewhat familiar with how AI can be used in the creation of SLRs, and a few (n=3) claimed to be very familiar with such use. In contrast, a small number (n=2) of respondents reported little to no familiarity with AI in this context, suggesting that adoption is still in the early stages or limited.

Respondents were also asked how they perceived their organization's level of familiarity with AI use in SLRs. Most respondents (n=6) judged their organization's level of familiarity to be similar to their own. The remaining respondents (n=5) ascribed lower familiarity to their organizations compared with their own.

Theme - Use of AI tools by respondents

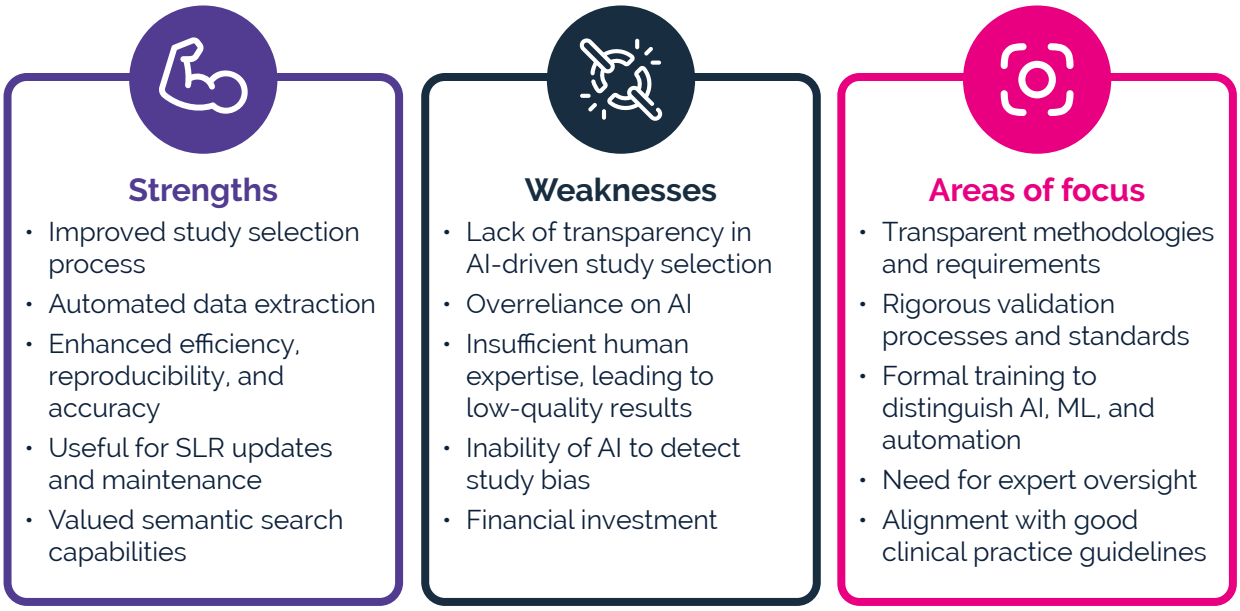
Most respondents (n=7) had explored how AI tools can be used for conducting SLRs, with tools such as DistillerSR, Laser AI, Nested Knowledge, and Rayyan being commonly cited. One respondent mentioned the use of large language models (LLMs) such as GPT-4 and Claude 3. Additionally, one respondent noted that their organization had started exploring Microsoft Copilot in the context of SLRs.

Three respondents reported that they had not evaluated the use of AI tools for SLRs in HTA. Another respondent characterized their organization as still in the initial phases of assessing AI tools.

Theme - Strengths, weaknesses, and areas for improvement

Respondents highlighted several strengths, weaknesses, and areas for improvement regarding AI use in SLRs for HTA (**Figure 3**). There was agreement on the need for collaborative research to validate and standardize AI algorithms used in this setting. The majority of respondents either agreed (n=4) or strongly agreed (n=5) that AI has the potential to improve the efficiency of SLRs. However, opinions on AI's ability to enhance quality and accuracy were more divided, with more than half of the respondents expressing neutrality (n=6). Only one respondent expressed disagreement with both efficiency- and quality-related statements. Increased efficiency was ranked highest by all respondents (n=11) when they were asked about the benefits of incorporating AI into SLRs for HTAs. Respondents ranked reproducibility (n=6) and improved accuracy (n=5) second. Respondents were least confident in AI's potential to reduce bias in SLRs.

Figure 3: Perceptions of HTA respondents on the strengths and weaknesses of AI use in SLRs and areas of focus for implementation



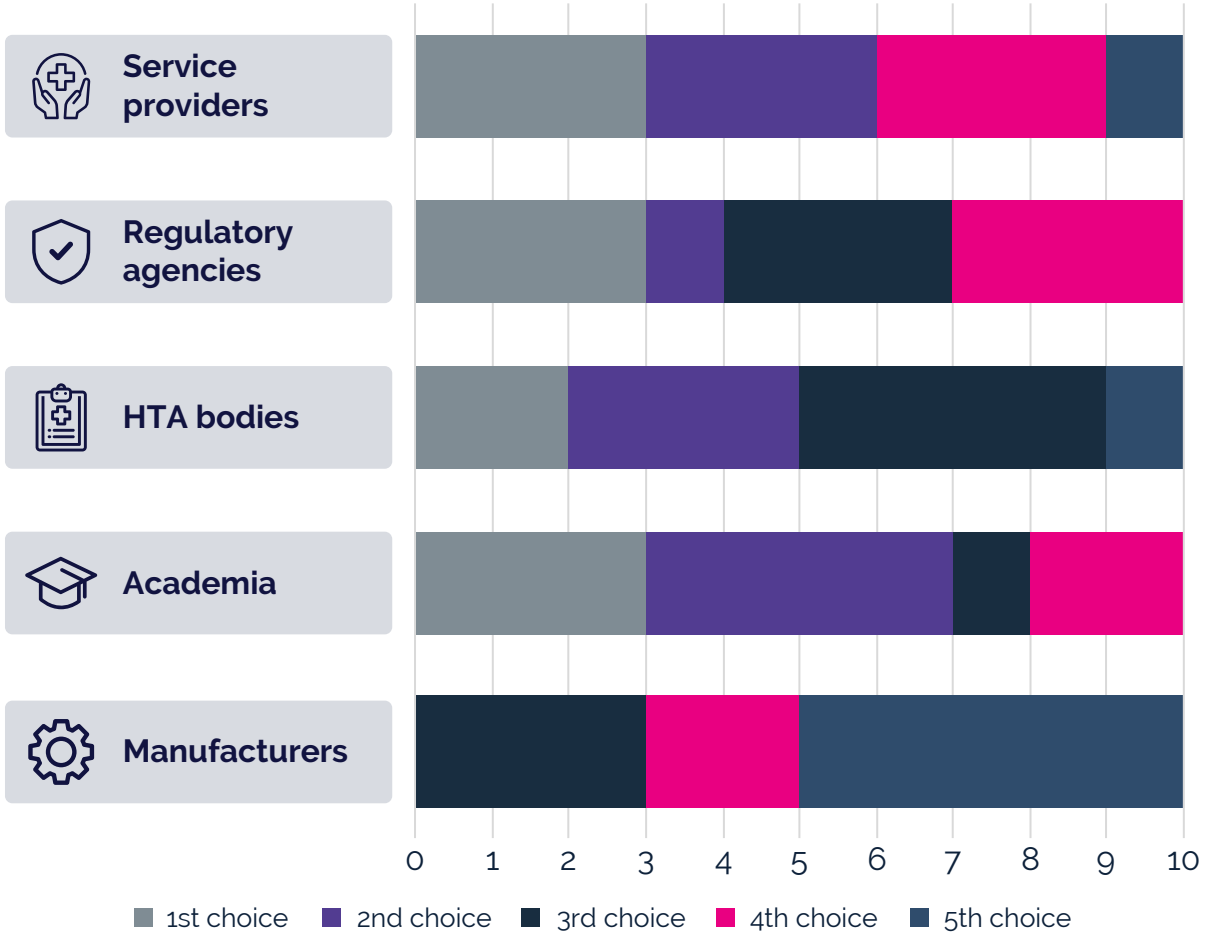
Abbreviations: AI, artificial intelligence; ML, machine learning; SLR, systematic literature review.

Theme - Regulatory needs

Transparency requirements were identified as the most critical factor for ensuring high-quality AI-supported SLRs. Seven respondents agreed (n=4) or strongly agreed (n=3) that AI platforms should be validated or certified by regulators or HTA bodies prior to use in SLRs submitted for HTA.

When asked about preferences for AI tool developers, respondents favored tools created and validated by service providers (e.g., consultancies and AI platforms), regulatory agencies, HTA bodies, and academia. AI tools developed by manufacturers were viewed less favorably, with six respondents ranking them as the least preferred option (**Figure 4**).

Figure 4: Respondent rankings of the relevance of stakeholder groups (from most to least relevant) in assuming primary responsibility for the development and validation of AI tools for SLRs in HTA



Abbreviations: AI, artificial intelligence; HTA, health technology assessment; SLR, systematic literature review.

Theme - Implementation

One compelling reason HTA agencies would appreciate collaboration with academia and other HTA stakeholders is to gain access to AI tools (n=6), with budgetary constraints a likely motivator. Respondents generally showed a strong preference for support in the form of technical expertise, knowledge sharing, and networking opportunities.

Theme - Timeline

Four respondents said they expected that the use of AI in SLRs for HTAs would be validated and standardized within the next 2-3 years. Three respondents reported that they believed validation and standardization could occur much sooner, within the next year.

Conclusions

This research indicates that AI-assisted SLRs are expected to play a vital role in the HTA process.

TLR results

- Most HTA agencies have neither opposed nor officially supported the implementation of AI for SLRs.
- The recommendations issued by CDA-AMC and NICE are more detailed than those from IQWiG and suggest that the transparent use of AI could enhance human efforts throughout various stages of the SLR process.

Survey results

- Most respondents reported being familiar with AI in SLRs for HTAs, and many are already using AI tools.
- While the efficiency benefits of AI were recognized, concerns about transparency, bias, and overreliance remained.
- Respondents emphasized the need for regulatory validation of AI tools, and there was a priority for support through access to AI tools and expertise.
- Many expect AI standardization to occur within the next 2-3 years and noted that overcoming challenges while establishing best practices will require collaboration among stakeholders.

Acknowledgements

- The authors would like to thank Petra Nass, Oliver Walker, Yuemeng Hu, and Christina DuVernay for their contributions to this work. This study was funded through the OPEN Health Scientific Office buyout fund.

References

1. Fox GE, Arcà E, Hu Y, Santpurkar N, Walker O, Torres Ames J, Nass P. From Data to Decisions: Decision-Makers' Perceptions of Artificial Intelligence-Powered Evidence Synthesis in HTA Submissions. ISPOR Europe 2024. Available at: https://www.valueinhealthjournal.com/article/S1098-3015(24)05952-7/fulltext (Accessed July 2025)
2. Fox GE, Torres Ames J, Santpurkar N, Arcà E, Nass P. Artificial Intelligence–Supported Evidence Synthesis for Reimbursement Decision-Making. AMCP Annual Meeting 2025. Available at: https://amcpannual.org/sites/default/files/2025-04/JMCP%20Supplement%20-%20AMCP%202025.pdf (Accessed July 2025)
3. Arcà E, Torres Ames J, Avissar J, Santpurkar N, Fox GE. Unlocking AI's Potential in Pricing and Reimbursement: Insights Across Global Healthcare Archetypes. ISPOR Annual Meeting 2025. Available at: https://www.ispor.org/heor-resources/presentations-database/presentation-cti/ispor-2025/poster-session-3/unlocking-ai-s-potential-in-pricing-and-reimbursement-insights-across-global-healthcare-archetypes (Accessed July 2025)
4. Fox GE, Santpurkar N, Torres Ames J, Avissar J, Arcà E. AI Tools for Literature Reviews: Are Current Guidelines Meeting the Needs of Researchers? ISPOR Annual Meeting 2025. Available at: https://www.ispor.org/heor-resources/presentations-database/presentation-cti/ispor-2025/ai-assisted-literature-reviews-requirements-and-advances/ai-tools-for-literature-reviews-are-current-guidelines-meeting-the-needs-of-researchers (Accessed July 2025)
5. Fox GE, Torres Ames J, Santpurkar N, Nass P, Arcà E. Opportunities and Challenges: HTA Body Perspectives on the Integration of Artificial Intelligence in Evidence Synthesis. HTAi Annual Meeting 2025. (Accessed July 2025)
6. CDA-AMC. Canada's Drug Agency Position Statement on the Use of Artificial Intelligence in the Generation and Reporting of Evidence. 2025. Available at: https://www.cda-amc.ca/sites/default/files/MG%20Methods/Position_Statement_AI_Renumbered.pdf (Accessed July 2025)
7. NICE. Developing NICE Guidelines: The Manual. 2024. Available at: https://www.nice.org.uk/process/pmg20 (Accessed July 2025)
8. NICE. Use of AI in Evidence Generation: NICE Position Statement, V. 1. 2024. Available at: https://www.nice.org.uk/about/what-we-do/our-research-work/use-of-ai-in-evidence-generation--nice-position-statement (Accessed July 2025)
9. IQWiG. General Methods, V. 7.0. 2023. Available at: https://www.iqwig.de/en/about-us/methods/methods-paper/ (Accessed July 2025)

QR code for survey

Please scan the QR code to participate in our ongoing survey.

OPEN HEALTH

www.openhealthgroup.com