SA8

From Data to Decisions: Decision-Makers' Perceptions of Artificial Intelligence-Powered Evidence Synthesis in HTA Submissions

Grace E. Fox,[1] Emanuele Arcà,[2] Yuemeng Hu,[3] Nita Santpurkar,[4] Oliver Walker,[5] Juliette Torres Ames,[5] Petra Nass[6]

[1]OPEN Health HEOR & Market Access, Bethesda, MD, USA, [2]OPEN Health HEOR & Market Access, Rotterdam, The Netherlands, [3]OPEN Health HEOR & Market Access, Shanghai, China, [4]OPEN Health HEOR & Market Access, Mumbai, India, [5]OPEN Health HEOR & Market Access, London, UK, [6]OPEN Health HEOR & Market Access, Berlin, Germany

INTRODUCTION

- Evidence synthesis is a cornerstone of any health technology assessment (HTA), with systematic literature reviews (SLRs) serving as an unbiased, reproducible process for identifying, collecting, and synthesizing results from multiple sources.[1]
- Since SLRs are highly labor-intensive, exploring avenues to integrate artificial intelligence (AI) with traditional methodologies is of interest.[2]
- However, it is unclear whether AI-supported SLRs will be accepted when used for HTA submissions.

OBJECTIVES

- The objective of this study was to engage with various HTA bodies to understand their perceptions regarding the use of AI-supported SLRs for HTA.

METHODS

Targeted literature review

- We conducted a targeted literature review (TLR) in Embase and MEDLINE (January 2019-October 2024) using terms for AI, natural language processing, and machine learning, combined with terms for HTA.
- Supplemental searches were conducted to identify policy documents and guidelines from HTA agency websites (January 2019-October 2024).

Stakeholder survey

- The results of the TLR informed development of a survey designed to gather insight into how HTA bodies regard the use of AI in SLRs.

RESULTS

Targeted literature review results

- The search of Embase and MEDLINE identified 354 records, and the grey literature search of HTA agency websites identified 7,560 records. Of these 7,914 records, 3 met the inclusion criteria.
- The majority of HTA bodies did not mention using AI for SLRs.
- Only the National Institute for Health and Care Excellence (NICE) in the United Kingdom[3,4] and the Institute for Quality and Efficiency in Healthcare (IQWiG) in Germany[5] provided recommendations on the use of AI for SLRs (see **Table 1**).

Table 1. HTA agency recommendations for the use of AI-supported SLRs

HTA AGENCY (COUNTRY)	RECOMMENDATIONS
• NICE (United Kingdom)	• AI should be employed to support, not replace, human efforts • AI use should be transparent and fully disclosed • ML and LLM may be used to create search strategies, automate study selection, automate study classification and visualization, and automate data extraction (noted as a less established use)
• IQWiG (Germany)	• Validated ML classifiers (e.g., RCT classifiers) may be used to create search strategies • ML may be evaluated and used to support study selection

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

Survey results

- A total of 17 participants completed the survey, including 10 individuals with roles within HTA bodies. This analysis focuses on these 10 participants, referred to as 'respondents' throughout this poster. The countries they represent are shown in **Figure 1**. On average, respondents had over 13 years of professional experience in HTA.

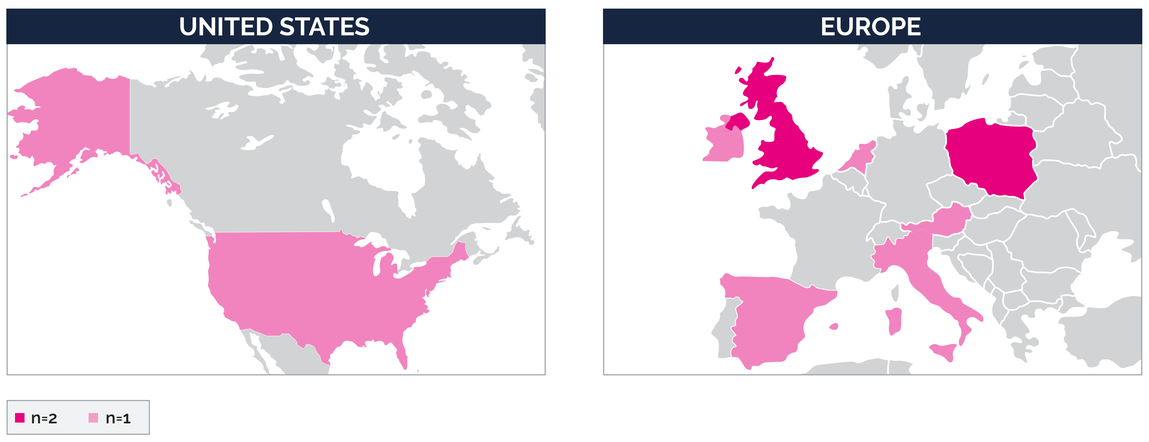

Figure 1. Geographical representation of HTA survey respondents

- The majority of respondents were somewhat familiar with AI in SLRs for HTA (n=6), but reported that their organizations were less familiar (n=6). Only two respondents reported that both they, and their organizations, were very familiar with AI.
- Most respondents had explored specific AI tools for SLRs, like **DistillerAI, LaserAI, Nested Knowledge, PittsAI, and Rayyan**. However, three respondents stated they were not using any AI tools.
- Key strengths, weaknesses, and areas of improvement regarding AI use in SLRs for HTA, reported by respondents, are summarized in **Figure 2.**
- Respondents agreed that **collaborative research to validate AI algorithms** for SLR was needed to validate and standardize its use in HTAs.

RESULTS (CONT)

STRENGTHS	WEAKNESSES	AREAS OF FOCUS
• Interest and experience in adoption • Increased efficiency • Reproducibility • Improved accuracy • Automated data extraction • Improved study selection process • Use in SLR updates and maintenance • Semantic search capabilities • Preference for study selection and data extraction by one human reviewer and one AI reviewer	• Lack of transparency for study selection • Excessive confidence/reliability in AI • Lack of human knowledge and training on AI use leading to low-quality results • AI inability in detecting study bias • Costs/investment	• Transparent methodologies and requirements • Rigorous validation processes and standards • Formal training on AI to understand the differences between AI, machine learning, and automation, and their limitations, and use in SLRs • Expert oversight • Good clinical practice guidelines

Figure 2. HTA respondents' perceptions of the strengths and weaknesses of AI use in SLRs for HTAs and areas for improvement for successful implementation

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

- **A majority agreed (n=4) or strongly agreed (n=4) that AI has the potential to improve the efficiency of SLRs in HTA.** However, perceptions of AI's potential to improve quality and accuracy was mixed, with half of respondents expressing neutrality. Only one respondent disagreed with both statements about AI improving efficiency and quality. AI as an assistant (not replacing a human) was seen as the most appropriate option.
- Transparency requirements were identified as the most important factor for regulating AI methodologies and ensuring high-quality outcomes. Six respondents agreed that AI platforms need to be **validated or certified by regulators and/or HTA bodies** to be used for SLRs.
- Respondents preferred AI tools developed and validated by **service providers (e.g., consultancies and AI platforms), regulatory agencies, HTA bodies and academia**. Manufacturers were ranked as the least relevant by five respondents (see **Figure 3**).

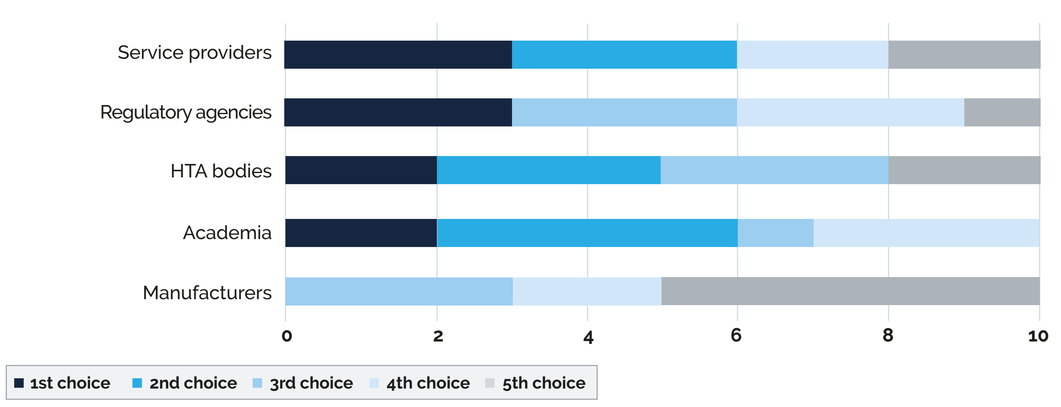

Figure 3. Respondent rankings of the relevance of stakeholder groups (from most to least relevant) in assuming primary responsibility for the development and validation of AI tools for SLRs in HTA

CONCLUSIONS

- Our findings suggest that AI-supported SLRs are likely to become an integral part of the HTA process.
- Most HTA bodies have neither rejected nor formally endorsed the use of AI for SLRs.
- The recommendations issued by NICE are more detailed than those issued by IQWiG and suggest that transparent use of AI could augment human efforts in multiple phases of the SLR process.
- The majority of HTA respondents were aware of, and already using, AI tools; however, scepticism remains on how much we can rely on AI tools for HTAs, with high emphasis on ensuring reliable, transparent, and unbiased results, in order to ensure effective and safe results are made available to patients. A collaborative approach across stakeholders will be required to address challenges and create best practices.

REFERENCES

1. Kim JSM, Pollock M, Kaunelis D, Weeks L. Guidance on review type selection for health technology assessments: key factors and considerations for deciding when to conduct a de novo systematic review, an update of a systematic review, or an overview of systematic reviews. *Syst Rev*. 2022;11(1):206. doi:10.1186/s13643-022-02071-7
2. Bolaños F, Salatino A, Osborne F, Motta E. Artificial intelligence for literature reviews: opportunities and challenges. *Artif Intell Rev*. 2024;57(10):259. doi:10.1007/s10462-024-10902-3
3. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. 2024. https://www.nice.org.uk/process/pmg20
4. National Institute for Health and Care Excellence. Use of AI in evidence generation: NICE position statement, V. 1. 2024. https://www.nice.org.uk/about/what-we-do/our-research-work/use-of-ai-in-evidence-generation--nice-position-statement
5. IQWiG. General methods, V. 7.0. 2023. https://www.iqwig.de/en/about-us/methods/methods-paper/

DISCLOSURES

This study was funded by OPEN Health. All authors are current or former employees of OPEN Health.

SCAN THE QR CODE TO PARTICIPATE IN OUR ONGOING SURVEY
Your responses will help us understand the experiences, perceptions, and concerns surrounding the integration of AI technologies in evidence synthesis for HTA

Presented at: ISPOR Europe 2024

Sponsored by OPEN Health's Scientific Office