Do health technology assessment (HTA) bodies recommend the conduct and submission of artificial intelligence-based literature reviews (AILRs)?

HTA194

View all of Parexel's posters at ISPOR Europe

Gautamjeet Singh Mangat (Mangat GS[1]), Nameeta Pilkhwal (Pilkhwal N[1]), Sugandh Sharma (Sharma S[2]), Rito Bergemann (Bergemann R[3])

[1]Parexel, Mohali, India; [2]Parexel, Chandigarh, India; [3]Parexel, Switzerland

Background

Existing scenario: A systematic review aims to provide a complete, exhaustive summary of current literature relevant to a research question with an objective and transparent approach. These are considered the best quality evidence and are essential for any evidence-based decision including HTA submissions. However, conducting systematic reviews for clinical decision-making is time-consuming, labor-intensive and may take anywhere between 6 months to a year depending on the objective. With the focus on getting novel medications faster to the patients, conducting systematic reviews sometimes becomes challenging. Thus, researchers are exploring automation such as artificial intelligence (AI) to improve the turn-around time of systematic reviews while ensuring that the quality of the evidence is not impacted.

The potential of AI in literature reviews: AI can support the entire SLR workflow from search strategy development to reporting. The key steps where AI is being explored and implemented are:

1) Screening of citations captured in the searches: AI tools after rigorous training provide a probability score for the inclusion/exclusion of a citation based on pre-defined PICOS criteria. The current tools also work on full-text PDFs and highlight relevant sections/phrases to support their recommendation.
2) Extraction of relevant data from citations: Many AI tools support automatic extraction of key data such as study design, patient population details, etc. from the full text of the citations.
3) Reporting of evidence: AI can support the development of brief summaries based on pre-defined templates by compilation of evidence into textual, tabular, and graphical formats.
4) Updating the SLR: The AI tools play a crucial role in living SLRs by allowing automatic updates for the search.

However, it is important to have human intervention at each step of the SLR process conducted by AI tools to ensure high quality. It is also important to have regulations for AI's ethical use and ensure reproducibility.

The role of HTA: To ensure AI-based literature reviews (AILRs) meet the gold standard of evidence, governments and health authorities must efficiently regulate AI. However, the very definition of AI in health is still the subject of discussion, debate, and negotiation among researchers and government authorities. Our research seeks to understand the available guidance on AILRs by HTA agencies and the road ahead for AILRs for HTA submissions.

Methods

We reviewed methodological guidance focusing on the use of AI and its acceptance by the following HTA bodies: National Institute for Health and Care Excellence (NICE) for England, Scottish Medicines Consortium (SMC) for Scotland, the National Centre for Pharmacoeconomics (NCPE) for Ireland, Haute Autorité de santé (HAS) for France, Gemeinsamer Bundesausschuss/Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (G-BA/IQWiG) for Germany, Canadian Agency for Drugs and Technologies in Health (CADTH) for Canada, and Pharmaceutical Benefits Advisory Committee (PBAC) for Australia.

Results

We did not identify any clear recommendations regarding the application of AILRs submitted as a part of the evidence package for reimbursement.

- NICE(1) recommended a **priority screening technique** that uses a machine learning (ML) algorithm to enhance screening efficiency under its guideline manual. This can be used to identify a higher proportion of relevant papers earlier in the screening process or to set a cut-off for manual screening.
- NICE also guides setting thresholds, considering the following:
  - The number of references identified so far through the search and how this identification rate has changed over the review (for example, how many candidate papers were found in each 1,000 screened)
  - The overall number of studies expected, which may be based on a previous version of the guideline (if it is an update), published systematic reviews, or the experience of the guideline committee
  - The ratio of relevant/irrelevant records found at the random sampling stage (if undertaken) before priority screening
- SMC(2) refers readers to NICE methodologies.
- NCPE(3) acknowledged the future of systematic reviewing via ML algorithms in their HRB-CICER (Health Research Board-Collaboration in Ireland for Clinical Effectiveness Reviews) report but did not mention anything substantial to guide users for acceptable usage of AI.
- IQWiG(4) recommended using ML-validated classifiers for identifying RCTs (Randomized controlled trials) under bibliographic searches in general methods.

Recommendation of key HTA bodies on AILRs

HTA Bodies	AI Guidance Maturity	Details
NICE – National Institute for Health and Care Excellence	◑	Recommends priority screening technique with the use of ML
Scottish Medicines Consortium	◔	Directs readers to NICE methodologies; no further information available
NCPE – National Centre for Pharmacoeconomics, Ireland	○	Acknowledges potential of AI; No further information available
HAS – Haute Autorité de Santé	○	No relevant information is available
Gemeinsamer Bundesausschuss	◔	Under development
CADTH Evidence Driven.	○	No relevant information is available
PBS – The Pharmaceutical Benefits Scheme	○	No relevant information is available

General guidance: Some HTAs refer to Cochrane guidance as their primary source, which is evaluating these algorithms to improve the efficiency of systematic review production through different initiatives, e.g., Cochrane RCT classifier, Transform project, and hybrid models, Screen4Me. We also identified an initiative where six HTA agencies are collaborating on various topics to bring better healthcare for people, including AI. This initiative might also provide guidance for adopting AI in the coming years.

Conclusion

- AILRs has the potential to make literature reviews more efficient. However, any use of AI must be augmented with manual review to ensure high quality output as none of the existing AI tools provide 100% accuracy of the analysis.
- Use of AI tools, even with limited application, need to be regulated by HTA bodies. This research did not identify any explicit guidance on AILRs across HTAs with only NICE making some recommendations for using AI, specifically for screening. There is much that remains to be done by HTA bodies to develop guidance for the responsible use of AI in combination with human review.

References

1) NICE (2023). Developing NICE guidelines: the manual. Process and methods [PMG20]. Last updated 02 August 2023
2) SMC (2023). Policies & publications. SMC guides and reports. Last updated 2023.
3) NCPE (2022). Update processes for guidelines – Systematic review (CICER: Collaboration in Ireland for Clinical Effectiveness Reviews). Last updated February 2022.
4) IQWiG (2022). General Methods (Version 6.1). Last updated January 2022.