A review of guidelines and checklists for the use of artificial intelligence/machine learning (AI/ML) in evidence generation: current landscape and recommendations

Raju Gautam¹. Saeed Anwar². Radha Sharma³. Tushar Srivastava¹

¹ConnectHEOR, London, United Kingdom; ²ConnectHEOR, Delhi, India; ³ConnectHEOR, Edmonton, Alberta, Canada

Email: tushar.srivastava@connectheor.com

BACKGROUND

- Artificial intelligence (AI) and machine learning (ML) have been rapidly assimilated into accessible and powerful tools that are transforming several sectors and functions, including evidence generation (EG) for healthcare decision making.^{1,2}
- The use of AI/ML for EG has increased recently due to its ability to streamline the resource-intensive EG processes.³
- Technology Assessment (HTA) bodies ✤ Health evaluate new health technologies to determine their safety, effectiveness, and economic value. This involves comparing the latest technology with existing alternatives to see if it works similarly, better, or worse. Essential evidence for HTA is typically generated through randomized controlled trials real-world evidence (RWE), economic (RCTs), evaluation, and systematic reviews, which remain the cornerstone of evidence-based medicine.^{4,5}
- The guideline and checklist on AI/ML use in EG are important to ensure that AI/ML enhances the robustness and fairness of EG in the HTA process.
- ✤ AIM: To review the guidelines and checklists available on the use of AI/ML in EG and outline their recommendations.

METHODS

- A pragmatic literature review was conducted using PubMed and Google Scholar from January 2019 to January 2025 to identify guidelines and checklists on the use of AI/ML in EG for healthcare decision making. A desk search was done in April 2025 to identify any additional literature.
- The websites of key HTA agencies (i.e., NICE, SMC, NCPE, HAS, IQWiG, TLV, CDA and PBAC) were also searched for the recommendation of the use of AI/ML for EG.



Austral



Checklist

SPIRIT-AI¹¹

STARD-AI¹⁵

TA ody	Guidance on the Use of Al/ML in EG	Recommendations
ICE ⁶	Yes	 The NICE position statement highlights the use of ML and LLMs to development of economic models. NLP can mine electronic health world data into structured formats. AI should be used to assist, not replace, human involvement in de Organizations and authors to "clearly declare its (AI) use, explain the structure of the
MC ¹⁶	No	NA
CPE ^{17,18}	No	NA
AS ^{19,20}	No	NA
WiG ²¹	Yes	The IQWIG General Methods 2023 states: "Machine learning approa selection" and "Validated classifiers from machine learning may be up
LV ²²	No	NA
DA ^{23,26}	Yes	 CDA recommendations align with NICE guidance on the use of AI/MI Al Definition: Updated to reflect the Canadian standard – "AI enal recognizing patterns in data." Ethical Considerations: AI in health must promote well-being, res Legislation: AI use must follow Canada's Voluntary Code and the applicable laws, including data protection and ethical standards.
BAC ^{24,25}	No	NA

Models; ML: Machine Learning; NA: Not Applicable; NLP: Natural Language Processing; NMA: Network meta-analysis; RCTs: Randomized Controlled Trials; RWE: Real-World Evidence

Poster presented at ISPOR in Montreal, Canada, May 13–16, 2025



ConnectHEOR

RESULTS

Five guidelines and six checklists were identified on AI/ML use in EG, including one (CDA-AMC) identified in April 2025 update.

These guidelines were NICE Position Statement, NICE Real World Evidence Framework, CDA-AMC Position Statement, Cochrane 2024 guidance on AI in evidence synthesis, and US FDA guidelines on AI/ML in drug development.

The checklists included CONSORT-AI, SPIRIT-AI, CHEERS-AI, DECIDE-AI, TRIPOD-AI and STARD-AI.

These guidelines recommend the use of AI/ML models across various EG categories. Examples include the use of LLMs for study selection, automated data extraction, and ML-based classifiers to identify RCTs from titles/abstracts for systematic reviews. LLMs can also be used to automate the creation, calibration, reporting, and adaptation of HE models. Additionally, NLP can transform unstructured realworld data into structured formats for RWE analysis. NLP can also mine EHRs and support the identification of eligible clinical trial participants or the reporting of adverse events for clinical evidence.

Among eight HTA agencies, only two NICE (UK) and CDA-AMC (Canada) has developed AI/ML EG guidelines. The IQWiG (Germany) stated the use of AI/ML in EG in their methodological guidelines.

DISCUSSION

This review found a few guidelines and checklists. Most of the HTA bodies do not have recommendation on the use of AI/ML in EG.

Specifically, only NICE and CDA-AMC has proper recommendations regarding AI/ML use in EG which most HTA bodies lack.

There are several options to explore the use of AI/ML for the various steps in EG which can save time and resources. Nevertheless, proper guidelines are needed regarding its use.

CONCLUSION

The emergence of guidelines and checklists on AI/ML use in EG is the right step in standardizing and ensuring quality in use of AI/ML. However, it must also address key issues like transparency and responsibility to ensure fair, robust evidence generation in HTA.

Alheiaily, A., et al. (2024). Future Healthcare Journal, 6(2), 94-Khosravi, P., et al. (2024). Frontiers in Artificial Intelligence Abogunrin, S., et al. (2025), Frontiers in Pharmacology, 16, Article 145424 World Health Organization (WHO), (n.d.), Health technology assessme O'Reilly D et al (2021) In Methods in Molecular Biology (Vol. 2249 pp. 429–454) Springe National Institute for Health and Care Excellence. Use of AI in evidence generation: NICE position statement. 2024 National Institute for Health and Care Excellence (NICE). NICE real-world evidence framework. 2022. Noel-Storr A, et al. Artificial intelligence (AI) technologies in Cochrane. May 9, 2024 US Food and Drug Administration. 2023 - revised 2025. Using artificial intelligence & machine learning in the development of drug & biological products: Discussion paper and request for feedback Liu X. Rivera SC. et al. BMJ 2020:370:m3164. doi:10.1136/bmi.m3164 Rivera SC. et al. BMJ 2020:370:m3210. doi:10.1136/bmi.m321 Hawksworth C, et al. Health Economics; 2023 Vasey B., et al. Nat Med 2022; 28:924-33 Collins, G. S., et al. (2024). bmj, 385 Sounderaiah V et al BMI Open 2021-11:e047709 SMC. Guide to Submitting Companies 2022 [Available from: https: HIQA. National Guidelines for the Economics Evaluation of Health Technologies in Ireland 2025 HIQA. Guidelines for Evaluating the Clinical Effectiveness of Health Technologies in Ireland 2019 [Available from: https://www.hiqa.ie/reports and-publications/health-technology assessment/guidelines-evaluating-clinical-effectiveness] HAS Choix methodologiques pou l'evaluation economique a la HAS 202 Availablefrom:https://www.hassante.fr/upload/docs/application/pdf/202007/guide_methodologique_evaluation_economique_has_2020_vf.pd HAS. Choix methodologiques pour le developpement clinique des pansements 2013 [Available from: https://www.has-sante.fr/upload/docs/application/pdf/20 12/guide_methodologigue_17122013.pd IQWiG. General Methods 2023 [Available from: https://www.iqwig.de/en/about-us/methods/methods-paper SBU. The process of producing a systematic review 2023 [Available from: https://www.sbu.se/en/method2/the-process-for-conducting-a-systematic-revi CADTH. Finding the evidence: Literature searching tools in support of systematic reviews 2024. [Available from: https://www.cadth.ca/finding-evidence-lit

PBAC. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee 2016[Available from: https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-versi PBAC. Appendix 2 Literature search methods 2016 [Available from: https://pbac.pbs.gov.au/appendixes/appendix-2-literature-search-methods.htm Canada's Drug Agency. Position Statement on the Use of Artificial Intelligence in the Generation and Reporting of Evidence. Ottawa: Canada's Drug Agency; 2025

Financial Disclosure: Authors are employees of ConnectHEOR Limited. No external funding received to conduct this research and no conflict of interest to declare.

