Artificial Intelligence Integration in Health Technology Assessments: A Review of Global Policies and Practices

Eon Ting[1], Matthew Badin[1], Nishu Gaind[2], Kimberly Hofer[2], Mir-Masoud Pourrahmat[2], Luka Ivkovic[2], Thomas Haugli-Stephens[3], Johanna Jacob[4], Mir Sohail Fazeli[2]

[1] AstraZeneca, Mississauga, Ontario, Canada; [2] Evidinno Outcomes Research Inc, Vancouver, BC, Canada; [3] AstraZeneca AS, Oslo, Norway; [4] AstraZeneca AB, Stockholm, Sweden

Introduction

- Artificial intelligence (AI) is revolutionizing healthcare, including drug development, clinical decisions, and health technology assessments (HTAs)
- In HTAs, AI has potential to streamline processes, enhance evidence quality, and align strategies with the evolving expectations of HTA agencies[1]
- However, regulatory inconsistencies, methodological differences, and concerns of data quality pose integration challenges[2]

Objective

- To assess and synthesize the published literature on the use and acceptance of AI by HTA agencies, both for submissions by sponsors and for internal purposes within HTA agencies

Methods

- Comprehensive review of HTA websites for guidance documents, policy statements, and opinions on use of AI and machine learning (ML) conducted for HTA agencies in Canada, Europe, and Asia-Pacific
  - Supplementary search of Embase, bibliographies of previous reviews, and gray literature was completed on December 11, 2024
- Publications on AI/ML approaches recommended/accepted/used by HTA agencies across therapeutic area were identified using PCC framework (Population, Concept, Context)[3]

Results

Publication Selection

- Review included 18 publications (n; Figure 1): 16 from HTA agency websites and 2 from Embase
 - Embase articles included a NICE commentary and a paper from NICE on AI/ML in screening

Figure 1. PRISMA flow diagram
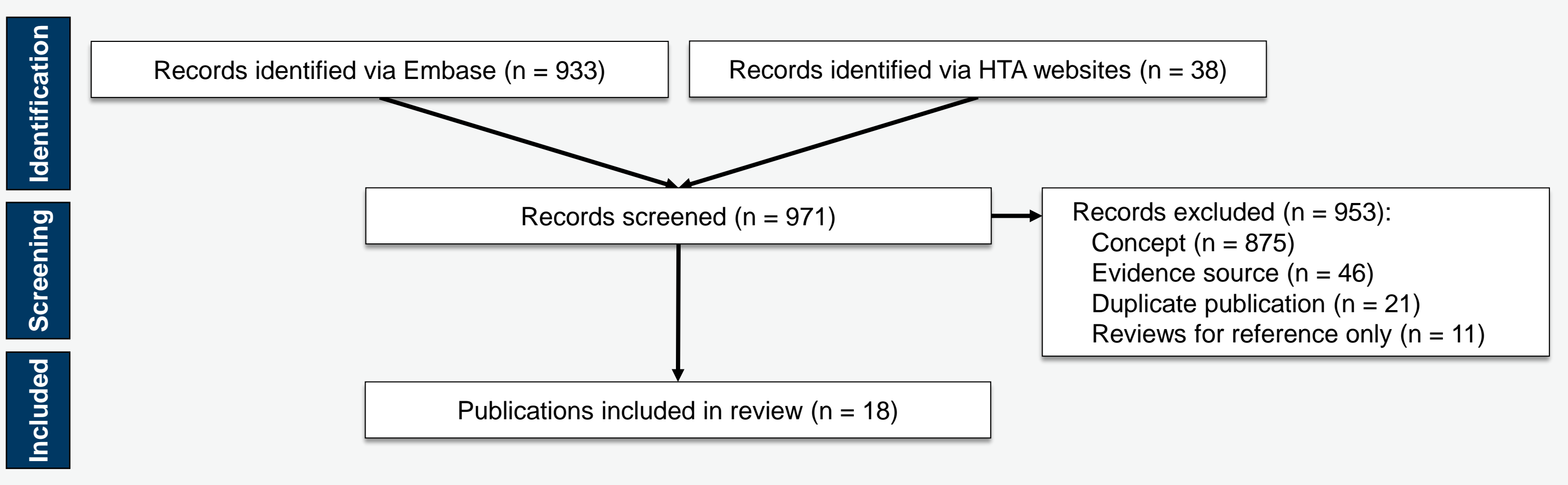
Publication Characteristics

- Few HTA agencies provided AI/ML guidance (**Figure 2**), mainly from NICE (UK)[4-7] and CDA-AMC (Canada)[8-11] (n=4 each), followed by IQWiG (Germany),[12,13] and EUnetHTA/JCA (Europe; n=2 each)[14,15]

Figure 2. Included publications on use of AI/ML by HTA agencies
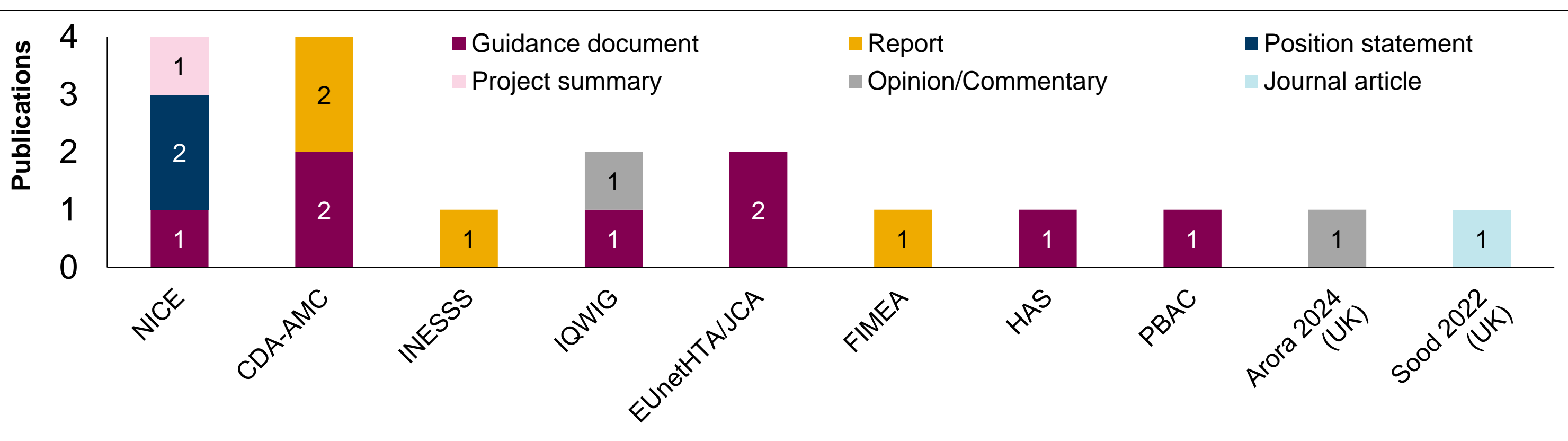
AI/ML Use and HTA Thresholds

- AI-enabled tools for publication screening from INESSS,[22] NICE,[21] and EUnetHTA[15] tended to favor high sensitivity thresholds (>95% sensitivity in correctly identifying all relevant articles), while specificity was lower (INESSS: ≥30% specificity in correctly excluding irrelevant articles)
  - INESSS tested 3 strategies by using 4 INESSS publications containing literature reviews (**Table 1**)[22]

HTA agencies are beginning to acknowledge AI/ML in submissions, particularly for literature reviews and evidence synthesis (n = 8), economic modeling (n = 3), real-world evidence generation (n = 6) and indirect treatment comparison (n = 1); Some HTA agencies are adopting AI/ML for internal use

For Internal Purposes

NICE
Position statement prioritizes staff upskilling and AI advancement in HTAs[5] NICE statement of intent outlines actions: learnings from AI projects, task automation, cybersecurity, ethical adherence, and AI literacy training[6]

CDA-AMC
Reported internal evaluation instrument on AI search tools for evidence synthesis[20]

INESSS (Quebec)
Evaluated GPT-4 for literature screening, finding a ranking strategy with 100% sensitivity and reasonable specificity[22]

FIMEA
Internal RWD/AI/ML network aims to promote the systematic sharing of topical issues and competence in these areas[18]

Sood et al. 2022
Summarized NICE's AI tools for COVID-19 surveillance: EPPI-4 for study screening and rule-based pattern matching for sub-topic categorization, enhancing efficiency and accuracy[21]

Literature Review and Evidence Generation

NICE
Stated use of ML classifier such as Cochrane randomized controlled trial (RCT) classifier priority screening tool for systematic reviews[4] Position statement highlights ML and large language models (LLMs) for evidence identification, screening, and visualization, with potential for data extraction and synthesis[5]

IQWiG
Validated study filters (e.g., for RCTs and systematic reviews) or ML classifiers such as RCT classifiers can be used if available[13]

CDA-AMC
AI may automate searches and data extraction in systematic reviews[10]

EUnetHTA
Stated using validated study filters (≥95% sensitivity) for RCT searches. RobotSearch and Cochrane's RCT classifiers are suitable tools[15]

HAS and PBAC
Referred Cochrane Handbook, which highlights AI tools like RCT Classifier and Screen4Me for streamlining study selection of RCTs[16,17]

Real World Evidence

NICE
Position statement emphasized detailed AI reporting for RWD extraction[5] Commentary outlined that in RWD analysis, NLP could be used to analyze unstructured data, or AI could assist with multimodal data integration[19]

CDA-AMC
Guidelines suggested including comments on ML methods in RWE studies, but stakeholders opted to leave the method choice to investigators[8,9]

IQWiG
Opinion in 2021 allowed the use of real-world perinatal data for AI/ML analysis[12]

FIMEA
Report indicated efforts to address RWE issues related to AI/ML both nationally and in Europe, including organizing forums and publishing updates[18]

Health Economic Modelling

NICE
Position statement notes AI's role in model development, including conceptualization, parameter estimation, and validation. LLMs can aid in replicating and cross-validating economic models[5]

NICE HTA Lab is exploring generative AI for economic modeling, from development to validation[7]

Indirect Treatment Comparison

JCA
Guidelines outlined that a wide range of statistical models, from logistic regression to ML models can be used in calculating propensity scores, used to perform indirect comparisons[14]

HTA agencies providing references to use of AI/ML in submissions

Reference to AI/ML use in HTA submissions

No reference to AI/ML use in HTA submissions

Indirect reference to AI/ML use in HTA submissions

No relevant documents found for HTA agencies in Scotland (SMC), Italy (AIFA), Spain (AETS), Sweden (SBU, TLV), Norway (NOMA), Denmark (DMC), Singapore (ACE), and Japan (C2H)

Table 1. Performance metrics of automated publication screening tool from INESSS[22]

Strategy type	Abstract Screening			Full-Text Screening		
	Basic strategy	Sensitive strategy	Ranking strategy	Basic strategy	Sensitive strategy	Ranking strategy
Sensitivity	92.3%	99.0%	100%	61.4%	92.4%	86.9%
Specificity	80.4%	55.1%	57.6%	82.6%	57.5%	60.6%

AI/ML Use and Human Involvement

- NICE advocates a "human-in-the-loop" AI model, emphasizing that AI should augment, not replace, human involvement to maintain trust in decision-making[5]

Strengths/Limitations

- **Strengths:** Recency, adherence to standard literature review methods (adapted for targeted review), and practical insights for HTA submissions, particularly in aligning with NICE guidance
- **Limitations:** Limited evidence base due to AI's emerging role in HTA and reliance on publicly available documents, highlighting the need for future studies incorporating stakeholder input

Conclusions

- AI integration in HTA submissions is evolving but remains inconsistent. NICE is the only agency with a clear policy/position statement with implementation strategies for AI. While other agencies acknowledge AI's role, formalized guidance remains limited
- Our findings highlight a growing interest in AI's potential across literature reviews, RWE, and economic modeling, yet standardization and regulatory harmonization remain as key challenges
- Increased collaboration among HTA bodies, industry, and academia can clarify acceptable HTA submission methods, enhance existing methods, and facilitate sharing of best practices

---

References

1. Fleurence R, et al. *Value Health* 2025;28(2):175-183. doi: 10.1016/j.jval.2024.10.3846
2. Zemplényi A, et al. *Front Public Health* 2023;11:1088121. doi: 10.3389/fpubh.2023.1088121
3. Joanna Briggs Institute. JBI Manual for Evidence Synthesis. 10.2.4 Inclusion criteria. 2024. Accessed 2024-12-20
4. NICE. Developing NICE guidelines: the manual, 2024; https://www.nice.org.uk/process/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869. Accessed 2024-12-20
5. NICE. Use of AI in evidence generation: NICE position statement, 2024; https://www.nice.org.uk/about/what-we-do/our-research-work/use-of-ai-in-evidence-generation--nice-position-statement. Accessed 2024-12-20
6. NICE. Statement of intent for artificial intelligence (AI). 2024; https://www.nice.org.uk/corporate/ecd12/resources/nice-statement-of-intent-for-artificial-intelligence-ai-pdf-40464270623941. Accessed 2024-12-20
7. NICE. HTA projects. 2025; https://www.nice.org.uk/about/what-we-do/our-research-work/hta-lab-projects#ai. Accessed 2024-12-20
8. CDA-AMC. Guidance for Reporting Real-World Evidence. 2023; https://www.cda-amc.ca/sites/default/files/RWE/MG0020/MG0020-RWE-Guidance-Report-Secured.pdf. Accessed 2024-12-20

9. CDA-AMC. Guidance for Reporting Real-World Evidence: Response to Stakeholder Feedback. 2023; https://www.cda-amc.ca/sites/default/files/RWE/MG0020/MG0020-Response-Document.pdf. Accessed 2024-12-20
10. CDA-AMC. An Overview of Clinical Applications of Artificial Intelligence. 2018; https://www.cda-amc.ca/sites/default/files/pdf/eh0070_overview_clinical_applications_of_AI.pdf. Evidence Synthesis. 2024; https://canjhealthtechnol.ca/index.php/cjht/article/view/AI0001/AI0001. Accessed 2024-12-20
11. CDA-AMC. Development of an Evaluation Instrument on Artificial Intelligence Search Tools for Evidence Synthesis. 2024; https://canjhealthtechnol.ca/index.php/cjht/article/view/AI0001/AI0001. Accessed 2024-12-20
12. IQWiG. das Gemeinsamen Bundesausschusses gemäß B. Kapitel 1. Abschnitt § 8 Absatz 1 Verfahrensordnung (VerfO) Entscheidung über die Gewährung der sekundären Datennutzung. 2021; https://iqtig.org/downloads/sekundaere-datennutzung/antraege/2021-06-02_Gewaehrung-sekundaere-Datennutzung-Uni-Koeln.pdf Accessed 2024-12-20
13. IQWiG. General Methods[4] Version 7.0, 2023; https://www.iqwig.de/methoden/general-methods_version-7-0.pdf. Accessed 2024-12-20
14. JCA. Methodological Guideline for Quantitative Evidence Synthesis: Direct and Indirect Comparisons 2024; https://health.ec.europa.eu/document/download/4ec8288e-6d15-49c5-a490-d8ad7748578f_en?filename=hta_methodological-guideline_direct-indirect-comparisons_en.pdf. Accessed 2024-12-20

15. EUnetHTA. Process of information retrieval for systematic reviews and health technology assessments on clinical effectiveness. 2019; https://www.eunethta.eu/wp-content/uploads/2020/01/EUnetHTA_Guideline_Information_Retrieval_v2-0.pdf. Accessed 2024-12-20
16. HAS. Choices in methods for economic evaluation. 2020; https://www.has-sante.fr/upload/docs/application/pdf/2020-11/methodological_guidance_2020_-choices_in_methods_for_economic_evaluation.pdf. Accessed 2024-12-20
17. PBAC. Guidelines for preparing a submission to the Pharmaceutical Benefits Advisory Committee. 2016; https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf. Accessed 2024-12-20
18. FIMEA. Supporting innovation and development of infrastructure. 2023; https://fimea.fi/en/supporting-innovation-and-development-of-infrastructure. Accessed 2024-12-20
19. Arora P et al. *J Comp Eff Res* 2024;14(1):e240212. doi: 10.57264/cer-2024-0212
20. CDA-AMC, Development of an Evaluation Instrument on Artificial Intelligence Search Tools for Evidence Synthesis. 2024; https://canjhealthtechnol.ca/index.php/cjht/article/view/AI0001/AI0001. Accessed 2024-12-20
21. Sood M, et al. *medRxiv* 2022; doi: 10.1101/2022.06.13.22276242
22. INESSS. Développement et évaluation d'un outil logiciel basé sur GPT-4 pour l'aide au tri de documents dans le cadre de revues de la littérature : une preuve de concept 2024; https://www.inesss.qc.ca/fileadmin/doc/INESSS/Demarche/INESSS_Outil_tri_GPT_4_VF.pdf. Accessed 2024-12-20