The Potential Use of Artificial Intelligence in Streamlining the Literature Review Process to Support Timely Evidence Generation for JCA Submissions

O'Donovan P[1], Metcalf T[1], Dodd O[1], Shukla S[1], Yakob L[1], Heron, L[1]

[1]Adelphi Values PROVE™, Bollington, Cheshire, Great Britain

HTA136

Introduction

- Joint clinical assessment (JCA) in Europe is quickly approaching with the first medicinal products set for assessment in 2025 (Figure 1).
- In preparation for this, manufacturers continue to examine what population, intervention, comparator and outcomes (PICO) criteria will be needed as they look to create a roadmap for reimbursement.
- However, given the short time frame from PICO confirmation to submission date (100 days), this creates challenges even with the most comprehensive planning.
- Given that SLRs are a vital and time-consuming part of this process, we anticipate that artificial intelligence (AI) will play a key role in successful implementation of the guidance.
- We propose the use of AI classifiers to semi-automate the abstract screening process to help streamline this process (Figure 2).

Methods

- To test this hypothesis, we utilised a previously conducted SLR and trained an AI classifier based on human decisions as per the agreed PICO criteria.
- The specialised AI classifier was developed based on the 3,084 dual-screened abstract screening results of a previously conducted SLR (Table 1). The average F1 score of the classifier was 0.54 (+/-0.03).
  - The classifier was trained by several experienced reviewers and 30% QC'd by senior reviewers to ensure that decisions were accurate.

Table 1 AI Classifier SLR update summary findings

Decision in training set	Number of references
Excluded	2,831 (91.8%)
Included	253 (8.2%)
Total	3,084

CI: confidence interval; RCT: randomised controlled trial.

- The same strategies from the original SLR databases (e.g., Embase, Medline etc.) were run on OVID SP ~3 months after the initial search and de-duplicated versus the studies initially identified. This resulted in 113 references being included in the SLR update.
- Studies were uploaded to an online platform. The specialised AI classifier was used to tag each article with a binary screening decision of include or exclude only. This tagging process was performed without the awareness of the human reviewer.
- We then conducted an SLR update, where the human reviewer was asked to screen each of the 113 references based on the same screening criteria as the initial SLR. The human reviewer remained blinded to the decision of the specialised AI classifier.
- Results were exported to Microsoft Excel and the decisions were compared, with a focus on the match rate between human reviewer and specialised AI classifier.

Figure 1. JCA Process

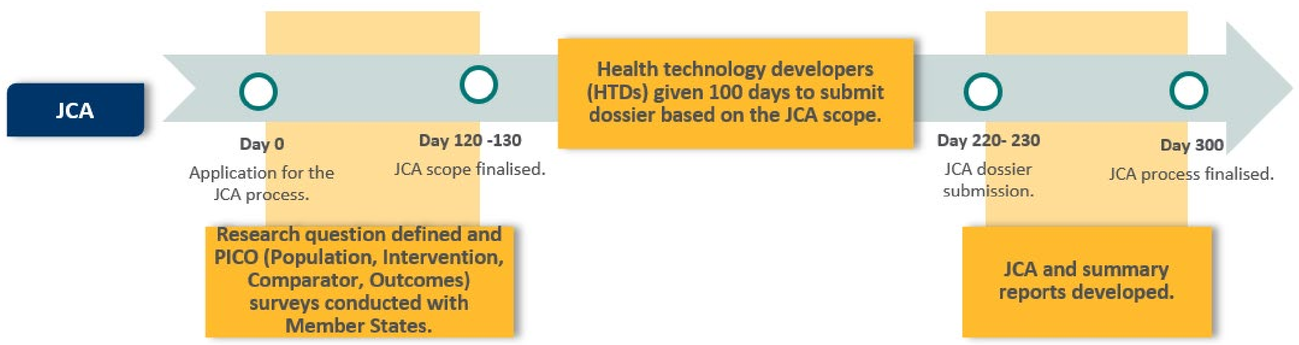
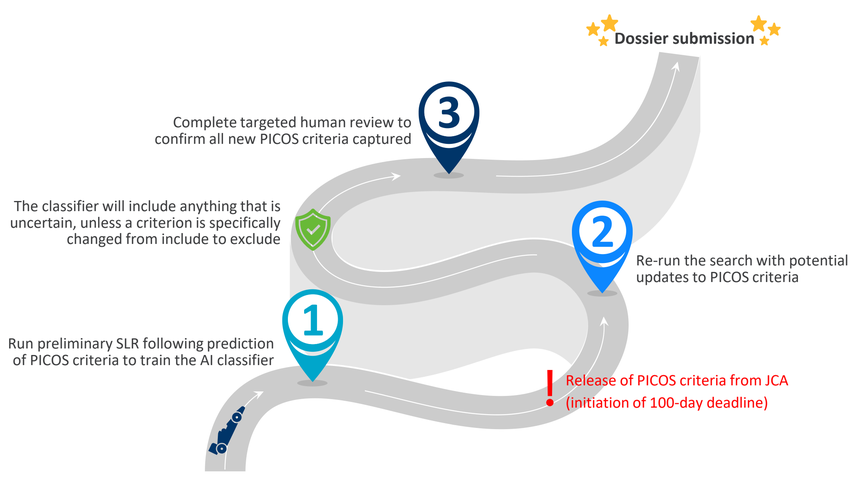


Figure 2. Methodological approach to using AI classifiers for JCA dossier development



Results

- Overall, when results between the human and AI screener were compared, we found that 91.2% of decisions were matched showing a high level of agreement.
- Of the 113 abstracts screened, three articles were deemed relevant by the human screener, all of which were also identified as includes by the specialised AI Classifier.
- Regarding excluded studies, all 100 studies excluded by the specialised AI classifier were also excluded by the human reviewer.
- The specialised AI classifier demonstrated very high sensitivity, with limited specificity.
  - 10 articles were not matched between human reviewer and specialised AI classifier. These were all excluded by the human reviewer and included by the specialised AI classifier.
- These results highlight that the use of AI classifiers in the role of the second reviewer on an SLR would result in a substantial reduction in time screening and streamline the SLR process, and we propose a framework by which we can implement this as part of the SLR methodology.

Discussion and conclusions

Discussion

- Based on this study, the unmatched decisions suggested an over-inclusive approach by the AI classifier, adding confidence that all relevant articles would be identified.
- AI classifiers present a unique opportunity to streamline the abstract screening process in SLR updates, automating the selection of relevant studies and reducing the reliance on manual screening when identifying studies from the vast amount of research can be incredibly time consuming.
- AI classifiers are comparable to human screening ensuring minimal error or bias is introduced during this important stage of the SLR development.

Conclusions

- By employing AI in this way, researchers can drastically shorten the timeframe needed for abstract screening at the SLR update stage, allowing them to identify and confirm relevant abstracts much faster compared to traditional methods.
- This acceleration in the process will be important, when working within tight deadlines, such as the 100-day window for JCA dossier development and submission.
- Appropriately leveraging AI in the literature review component will be key to ensuring an efficient JCA submission process.