MSR99

Application of Artificial Intelligence in Literature Reviews

Shiva Kumar Venkata[1], Sravani Velicheti[1], Vinayak Jamdade[1], Sreeja R[1], Monika Achra[1], Kushal Kumar Banerjee[1], Chandreyee Dutta Gupta[1], Michael Happich[2], Annabel Barrett[2]

[1]Eli Lilly Services India Private Limited, Bangalore, India, [2]Eli Lilly and Company, London, United Kingdom

BACKGROUND

- A rigorous screening phase is a vital component of a literature review, critically driving the quality of the evidence. However, screening is resource intensive.
- DistillerSR® is specialist systematic review software, designed with a user-friendly interface and artificial intelligence (AI)-based features supporting the reviewer more effectively compared to using spreadsheets or reference software.
- DistillerSR® tool manages, tracks, and streamlines the screening, data extraction, and reporting processes of systematic and targeted literature reviews (SLRs, TLRs). There are two types of artificial intelligence features in DistillerSR® which supports in semi-automation of the title-abstract screening process: 1) DistillerAI; 2) Classifiers (DAISY)
- DistillerAI applies a naïve Bayesian approach to screen title-abstracts after learning from decisions of human/manual screening.
- Classifiers, DistillerSR®'s AI system, is the "engine" behind the AI tools found throughout DistillerSR®. Classifiers is a statistical model that uses natural language processes to process information and classify it accordingly.
- DistillerAI and Classifiers can be implemented in both SLRs and TLRs. Classifiers was preferred to use in SLRs due to their advanced built-in algorithm which provides a set of evaluation metrics to report its performance. These metrics in turn provide confidence on Classifiers or helps us to identify possible gaps to consider before running AI.

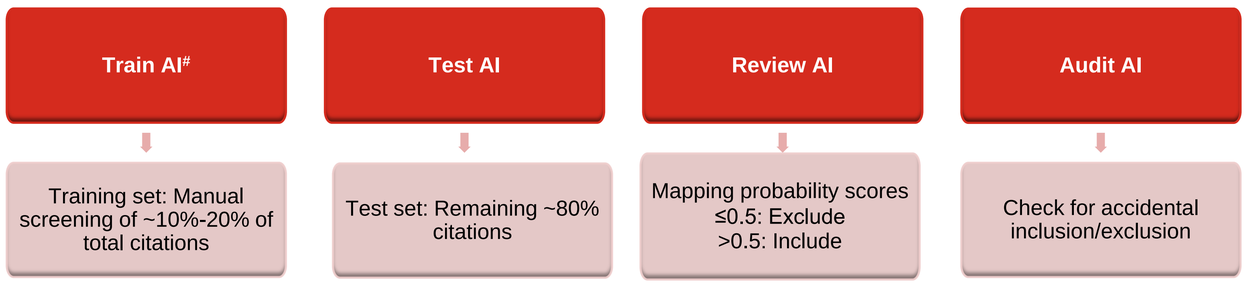
OBJECTIVE

- The purpose of this study was to evaluate the efficiency of the AI tools of DistillerSR® in conducting resource intensive title-abstract screening in TLR and SLR.

METHODOLOGY

- A total of eight TLRs and three SLRs were conducted between February 2021 and June 2023 using 'DistillerAI' and 'Classifiers', respectively on DistillerSR® platform.
- The efficiency was assessed in terms of 'screening burden' and 'accuracy' (false negatives [FN], %). At least 10% of the total citations were manually screened (one review for TLRs and two reviews with an independent conflict resolver for SLRs) from each review and used as a 'training set' for AI.
- In TLRs, DistillerAI uses responses from the training set and provides the likelihood of relevance scores that range from '0' (potential exclusions) to '1' (potential inclusions) for unscreened citations (Figure 1).
- In SLRs, Classifiers (include/exclude) uses the training set and screens all unreviewed citations in one of the two-reviewer set (Figure 2). The Classifiers are validated using a 'balance score' and 'recall score' (proportion of True positives/negatives vs False positives/negatives).

Figure 1: DistillerAI Methodology



#Manually screened citations will be used to train AI

Figure 2: Classifiers Methodology

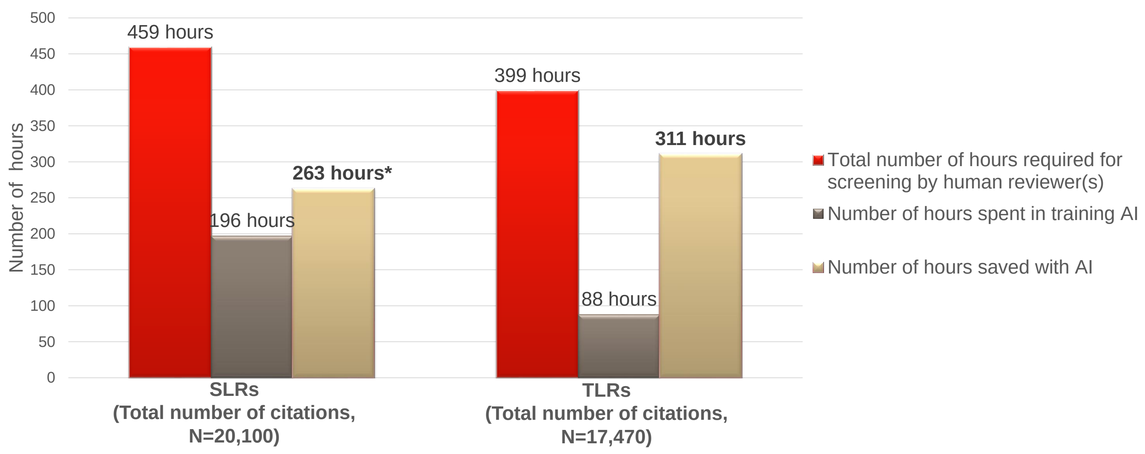
*Question: Is the citation relevant for inclusion?

RESULTS

AI results at Title-abstract screening level in literature reviews

- Number of total citations screened from February 2021 to June 2023 are presented against the time associated with the screening with and without the usage of AI in **Figure 3**.
- The median accuracy score across literature reviews was 90% (range: 85%-96%) with 1.64% of mean False-Negatives, which was comparable to manual screening (**Figure 4**).

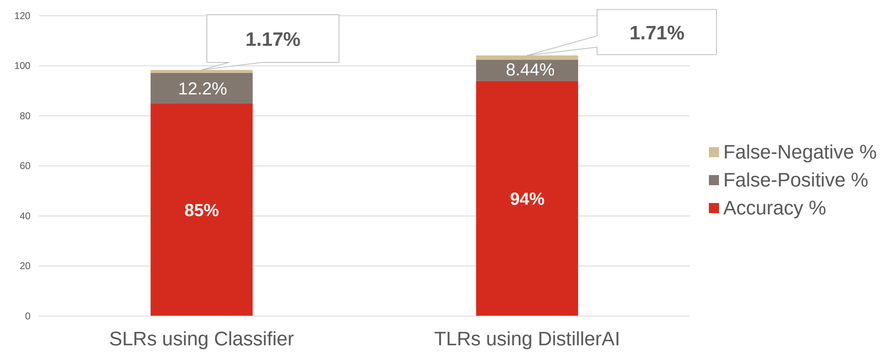
Figure 3: Time Savings generated by AI



Abbreviations: AI, artificial intelligence; SLR, systematic literature review; TLR, targeted literature review
*Spared reviewer time reference

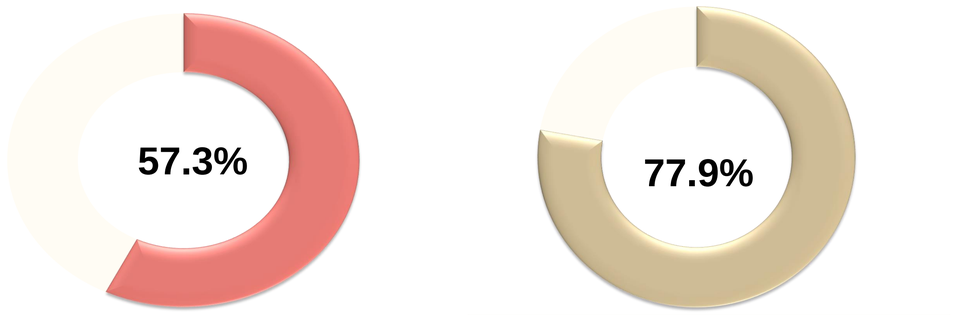
- The efficiency reported in **Figure 5** is directly proportional to the human efforts saved in the screening process for both SLRs and TLRs.
- Overall, around 67% (SLRs= 57.3%; TLRs= 77.9%) of human efforts (hours) were reduced by applying AI in the title-abstract screening irrespective of the type of literature review.

Figure 4: Accuracy attained by AI in screening evidence for different literature reviews



Abbreviations: AI, artificial intelligence; SLR, systematic literature review; TLR, targeted literature review

Figure 5: Efficiency achieved using AI features in DistillerSR® in literature reviews



Efficiency achieved using AI in SLRs

Efficiency achieved using AI in TLRs

Abbreviations: AI, artificial intelligence; SLR, systematic literature review; TLR, targeted literature review

Strengths of DistillerSR® AI tools

- It was found to be useful for literature reviews, especially with periodic updates.
- The training set can be repeatedly utilised across different literature reviews with a similar scope.
- It can partially replace the second reviewer in SLRs, by screening citations in the test set.
- The time and human savings generated are significant, and these can be further translated to cost and human effort savings.

Limitations of DistillerSR® AI tools

- DistillerSR® AI can only be run in Title-abstract screening level.
- It can provide responses only to binary questions (e.g., YES/NO), due to which exclusion reasons cannot be obtained.
- It cannot read non-English disclosures.
- Literature reviews on rare diseases can be difficult to train AI due to very low number of citations.

Study Limitations

- Unequal distribution of inclusions and exclusions in the training set can be a possible attribute for less accuracy
- Sample size for SLRs is relatively less (number of SLR projects tested [3] < TLR projects [8]) which can be a plausible explanation for lesser accuracy (85%) in SLRs when compared to TLRs (94%).

CONCLUSIONS

- AI is found to be an efficient tool for title-abstract screening, especially for large reference sets (>5,000).
- AI simulation tools are useful in prioritising likely inclusions and exclusions; however, additional quality checks are required to meet rigorous requirements of HTA.
- Further research is needed around recommendations for optimal integration of AI in literature reviews and use of AI in SLRs for HTA submissions.

Scan or click the QR code or use this URL (https://lillyscience.lilly.com/congress/isporeu2023) for a list of all Lilly content presented at the congress.

Sponsored by Eli Lilly and Company