

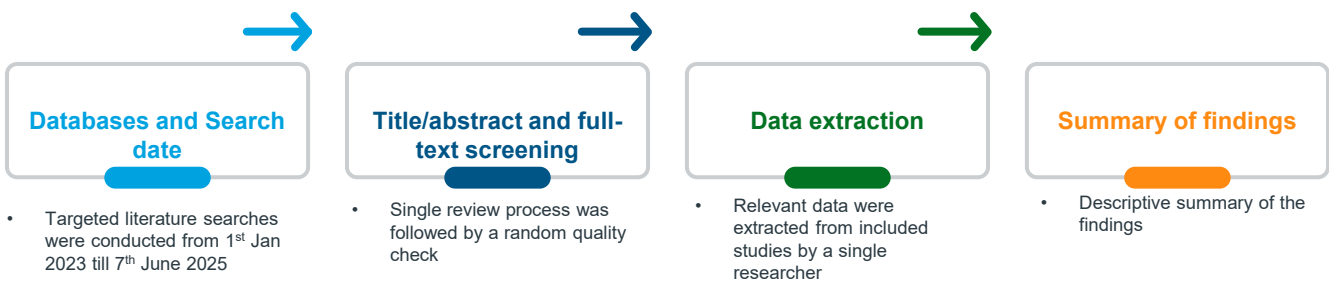
Background and Objectives

- SLRs are essential to evidence-based decisions and often require screening thousands of records. As volume of available data increases, more resources are needed for review and analysis¹
- This increased burden, combined with the emergence of AI/ML tools, presents an opportunity to automate certain SLR tasks, potentially reducing manual effort and expediting the review process²
- While many researchers have assessed the technical performance of AI/ML tools in SLRs, there is uncertainty on whether these tools truly reduce manual burden, increase efficiency, or improve accuracy³
- Extending on our previous work on the use of AI in SLRs, this review aimed to evaluate the functionalities of AI/ML-enabled web-based and software tools across the SLR workflow, with a particular focus on their application in data extraction, quality assessment (QA), and reporting

Methods

- An update of previous review published in 2023 was conducted to identify AI/ML-based SLRs (Figure 1). Search terms included “AI,” “ML,” “deep learning,” “SLR,” “meta-analysis,” and specific AI/ML-enabled platforms. No restrictions were applied on indication, treatment or geography

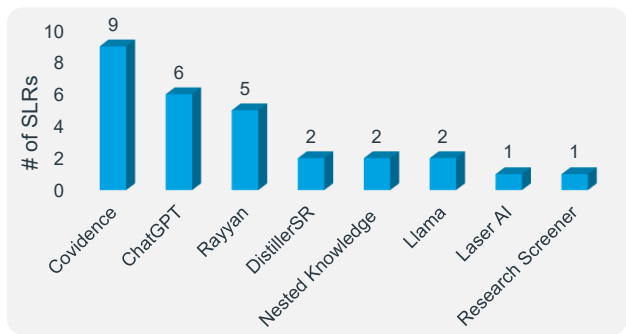
Figure 1: SLR methodology



Results

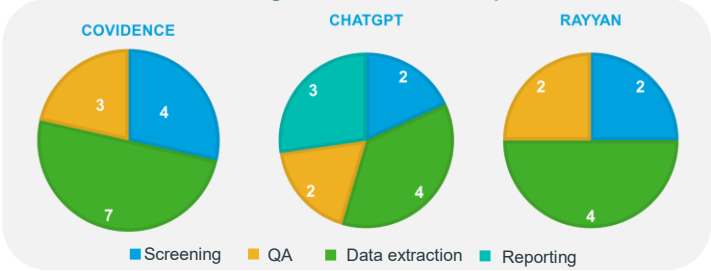
- Twenty-five SLRs using AI/ML-assisted tools were identified, with a substantial increase in AI/ML tool usage observed over the past two years (from 3 in 2023 to 11 in 2025)
- The most frequently utilised tools were Covidence followed by ChatGPT, Rayyan, DistillerSR, Nested Knowledge and Llama (Llama2-13b, and Llama3-8b) (Figure 2)
- Among these identified SLRs, AI/ML was predominantly used in data extraction (n=20), followed by screening (n=10) and QA (n=7), with Covidence, ChatGPT, and Rayyan being the most utilised tools for these steps (Figure 3). Table 1 shows capability of the most frequently used tools
- A limited number of SLRs reported the use of AI/ML for report drafting (n=3) and search strategy development (n=2), all of which utilised ChatGPT

Figure 2: Most frequently utilised AI/ML tools



Abbreviations: AI, artificial intelligence; ML, machine learning

Figure 3: Distribution of selected (most frequently used) AI/ML tools usage across the SLR steps



Abbreviations: AI, artificial intelligence; ML, machine learning; QA, quality assessment; SLR, systematic literature review

Table 1: Capabilities of selected (most frequently used) AI/ML enabled and web-based tools in the SLRs

	Covidence	ChatGPT	Rayyan
AI model	 - ML model: Cochrane RCT Classifier - Active-learning relevance model (EPPI-Centre)	Web-based GPT-5 model	- SVM classifier - AI agents
Data extraction	 - Side-by-side view of PDF and form - Export to Excel/RevMan	- Use in extractions with tailored prompts for predefined, pilot extraction grid - Perform independent duplicate extraction on critical/subjective items	- Customise tables and forms - Side-by-side view of PDF and form - Capture structured answers from PDF
QA	 - Built-in templates - Templates can be customised for QA	Use only as supportive input by giving specific prompts	- No dedicated RoB module - QA via custom fields or external tools
Reporting	 - Automated PRISMA 2020 flow diagram - Generates tables for QA and extractions, which can be exported and included in the report	Draft methods, PRISMA section, tables, lay summaries by giving specific prompts	- Automatic PRISMA 2020 flow - Exports available
Summary of AI integration	 AI-integration is only available for extraction, not for QA or reporting	AI-integrated prompts can be developed and tailored to support data extractions, QA and reporting	AI-integration is only available for extraction specifically for institutional customers, not for QA or reporting

Abbreviations: AI, artificial intelligence; ML, machine learning; PRISMA, preferred reporting items for systematic reviews and meta-analyses; QA, quality assessment; RCT, randomised controlled trial; RoB, risk of bias; SVM, support vector machine

Conclusions

- Use of AI/ML tools in the SLRs has evolved rapidly over the past two years, with a growing emphasis on data extraction. This represents a notable shift from our 2023 review, where these tools were primarily used for screening. However, its application remains limited in tasks such as QA, report drafting, and search strategy development
- These findings suggested that human intervention is necessary to ensure methodological rigour and transparency in evidence synthesis, as AI/ML tools cannot reliably be used independently to conduct end-to-end SLRs

References: 1. Kamra S, et al. Value in Health. 2023 Dec 1;26(12):S6.Dos Reis AH et al. Systematic Reviews. 2023 Dec;12(1):1-4; 2. Gates A, Guitard S et al. Systematic reviews. 2019 Dec;8(1):1-1