

## OBJECTIVE

- Literature review and synthesis of clinical evidence (ES) require increasing time and human resources as the amount of available clinical literature increases.
- To limit this burden, artificial intelligence (AI) methods may be used at different stages of ES. This review aims to summarize current state of AI-based solutions used in ES.

Figure 1. Current state of use of AI methods in the ES process

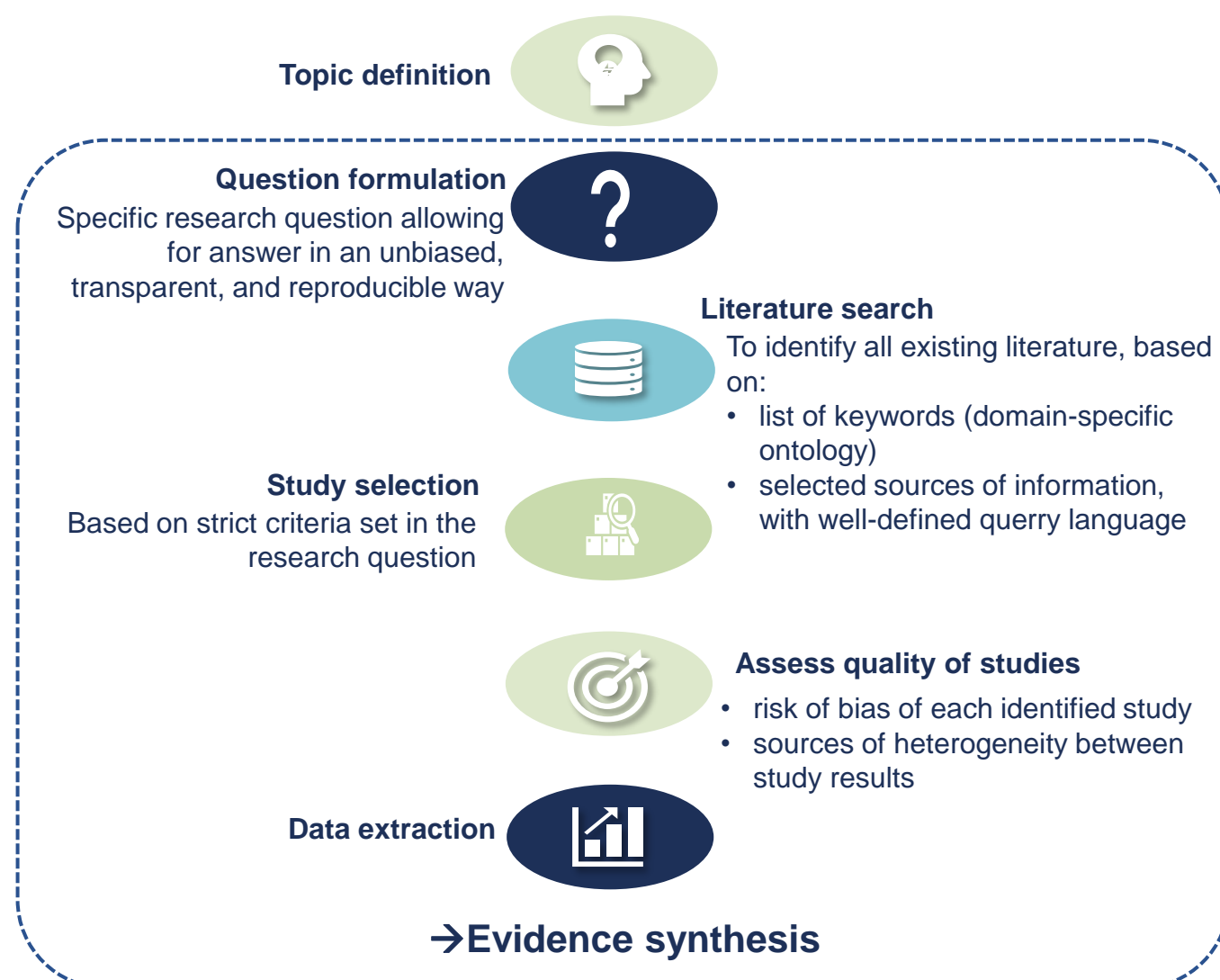


Figure 2. Combined keywords for the search

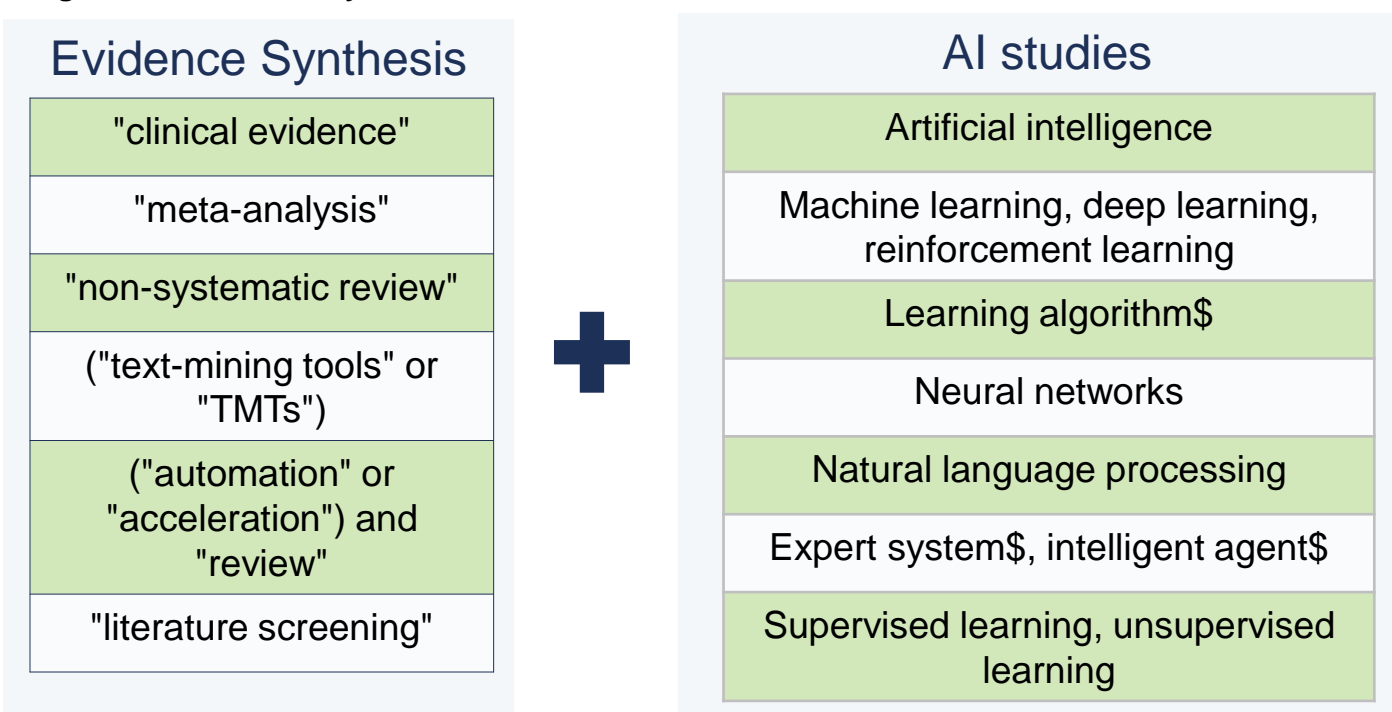


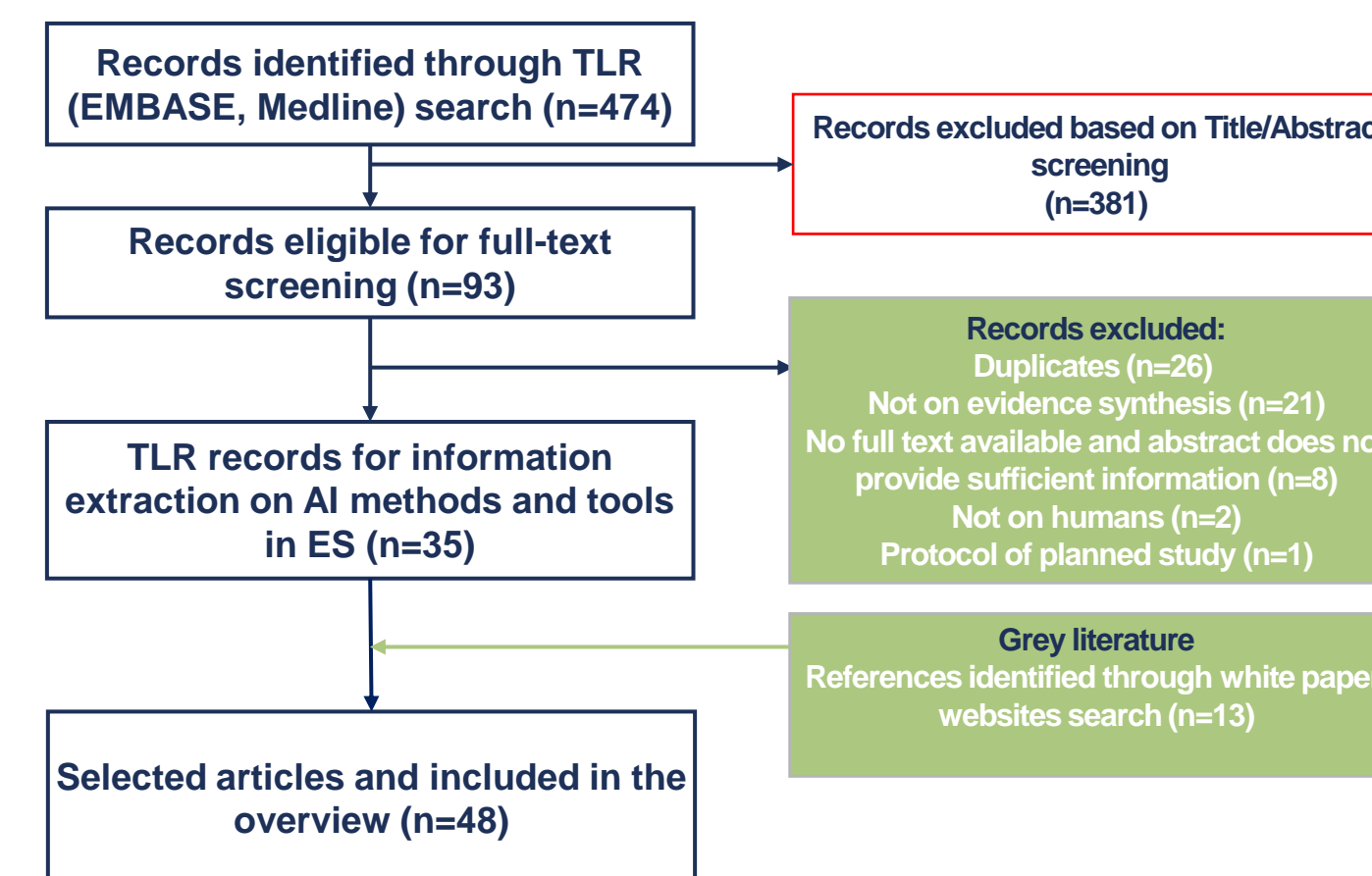
Table 1. PICOS

PICOS	Inclusion and exclusion criteria
<b>Population</b>	No restriction
<b>Interventions &amp; Comparators</b>	No restriction
<b>Outcomes</b>	No restriction
<b>Study design/type</b>	<b>Inclusions:</b> <ul style="list-style-type: none"> <li>Clinical evidence review and synthesis – any stage</li> <li>Use of AI – any method</li> </ul> <b>Exclusions:</b> <ul style="list-style-type: none"> <li>Digital application without AI support</li> </ul>
<b>Other restrictions</b>	<b>Time restriction:</b> Search was limited to studies published 1.1.2020 - 1.12.2021 to get the overview of the current status of development and use of AI methods. <b>Language:</b> English Search limited to Human studies

## RESULTS

- 474 titles and abstracts were screened, from which 35 publications were eligible for extraction of information about AI employed at any ES stage. An additional 13 publications were identified through grey literature search.
- 37 distinct AI-based solutions were identified with varying degrees of transparency and intellectual property protection, spanning from fully proprietary solutions, then free-to-use tools with no disclosed implementation, up to fully transparent implementations (n=17, none commercial) with source code available on permissive license. The most widespread way of sharing codes was GitHub software development platform. In 4 solutions, AI methods were deployed, using widely available libraries for Python or R programming language.

Figure 3. Selection



- 13 publications reported on integrated automatic ES approaches, including 5 real-time systematic literature review (SLR) or meta-analysis (MA).
- A wide range of AI methods were used at different stages of ES, with two dominant classes:
  - Machine Learning (ML) based on the ability to learn and improve automatically by finding patterns in data were employed in tasks related to grouping and classification, mainly at literature selection step.
  - Natural Language Processing (NLP) methods, aiming at making computers read, understand, interpret and manipulate human's natural language were employed, mainly in tasks of data extraction, literature selection and keywords selection.

- Reported performance of AI-based literature selection and information classification vs. human experts- assessed using a different sets of measures (precision, recall, discrimination, etc.) - was generally good (e.g. 99.9% reported by Sagheb, 2021, 79% reported by Pham, 2021), with strong emphasize on benefits from using AI:
  - Substantial decrease of human workload (e.g. 55-63%, Pham, 2021)
  - Significant reduction of analysis time
  - Including information not previously considered.
- Data extraction pose a bigger challenge than literature selection, hence fewer reported solutions and usually with worse performance (e.g. Gates, 2021 reported 48% (21–71%) median (IQR) relevance of extracted solutions).
- No tool allowed for complete elimination of human involvement, mostly in preparing training datasets or human supervision of AI-generated deliveries.
- At present, no major agency accepts submission involving AI-based review and synthesis of clinical evidence. This would require full confidence in the results is achieved, through completely transparent process and/or substantial validation.

## CONCLUSIONS

Recent advances in AI-based solutions for ES reported high concordance with human experts and a substantial reduction in time, but prior human intervention is still significant. High quality selection can be achieved, while the performance of algorithms used for data extraction remains insufficient.

Further evidence of the validity of these methods need to be established before they are considered by decision-makers at the same level as conventional methods.

## REFERENCES

- Abogunrin et al. ISPOR <https://www.ispor.org/heor-resources/presentations-database/presentation/euro2021-34077/113658>.
- Amezua-Prieto et al. *Current Opinion in Obstetrics & Gynecology* 32, no. 5 (October 2020): 335–41.
- Clark et al. *JMIR Medical Education* 7, no. 2 (31 May 2021): e24418.
- Gates et al. *BMC Medical Research Methodology* 21 (16 August 2021): 169.
- Pham et al. *Systematic Reviews* 10 (26 May 2021): 156.
- Qin et al. *Journal of Clinical Epidemiology* 133 (May 2021): 121–29.
- Riaz et al. *Journal of Clinical Oncology* 39, no. 6\_suppl (20 February 2021): 335–335.
- Sagheb et al. *The Journal of Arthroplasty* 36, no. 3 (March 2021): 922–26.
- Schoot et al. *Nature Machine Intelligence* v.3, 125–133 (2021).
- Yamada et al. *Journal of Medical Internet Research* 22, no. 12 (30 December 2020): e22422.
- Waring et al. *Artificial Intelligence in Medicine* 104 (1 April 2020): 101822.
- Wilson et al. *European Stroke Journal*, EPV0504/#786, 6, no. 1\_suppl (September 2021): 3–513.
- Zou et al. *Journal of Multidisciplinary Healthcare*. 13 (pp 671–679). 2020.

