

## BACKGROUND AND OBJECTIVES

The volume of digital healthcare data being generated is increasing rapidly, particularly as devices are becoming more complex and mechanisms to gather data become more accessible and sophisticated

With the help of AI, it is now possible to leverage RWD and begin to understand patient populations in holistic ways

**Objective:** To provide an overview about trends in current research and status of development and use of AI methods in RWE studies conducted using electronic health records (EHR) or claims databases.

## METHODS

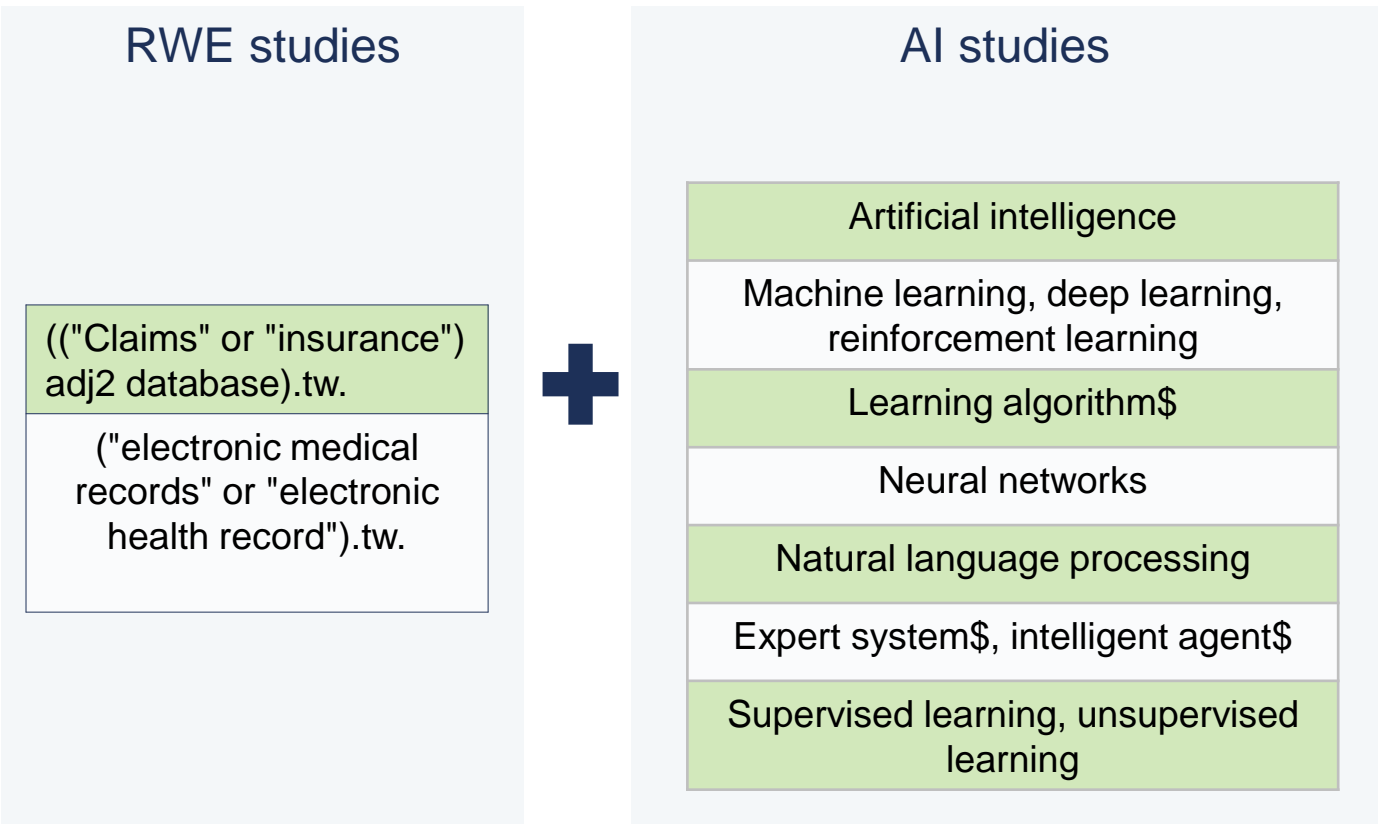
A search strategy was implemented in MEDLINE and EMBASE via OVID and covering the period from 1st January 2020 to 1st December 2021.

Both titles and abstracts were screened. It was completed by a grey literature search on specialized media, relevant conferences websites, providers websites and research reports.

Table 1. PICOS

PICOS	Inclusion and exclusion criteria
Population	No restriction
Interventions & Comparators	No restriction
Outcomes	No restriction
Study design/type	<b>Inclusions:</b> <ul style="list-style-type: none"><li>RWE study</li><li>Data source: medical records or claims database</li><li>Use of AI in the study methods (explicitly reported in the study abstract)</li></ul> <b>Exclusions:</b> <ul style="list-style-type: none"><li>Narrative reviews, RCTs, case reports, conference abstracts</li></ul>
Other restrictions	<u>Time restriction:</u> Search was limited to studies published 1.1.2020 - 1.12.2021 <u>Language:</u> English Search limited to Human studies

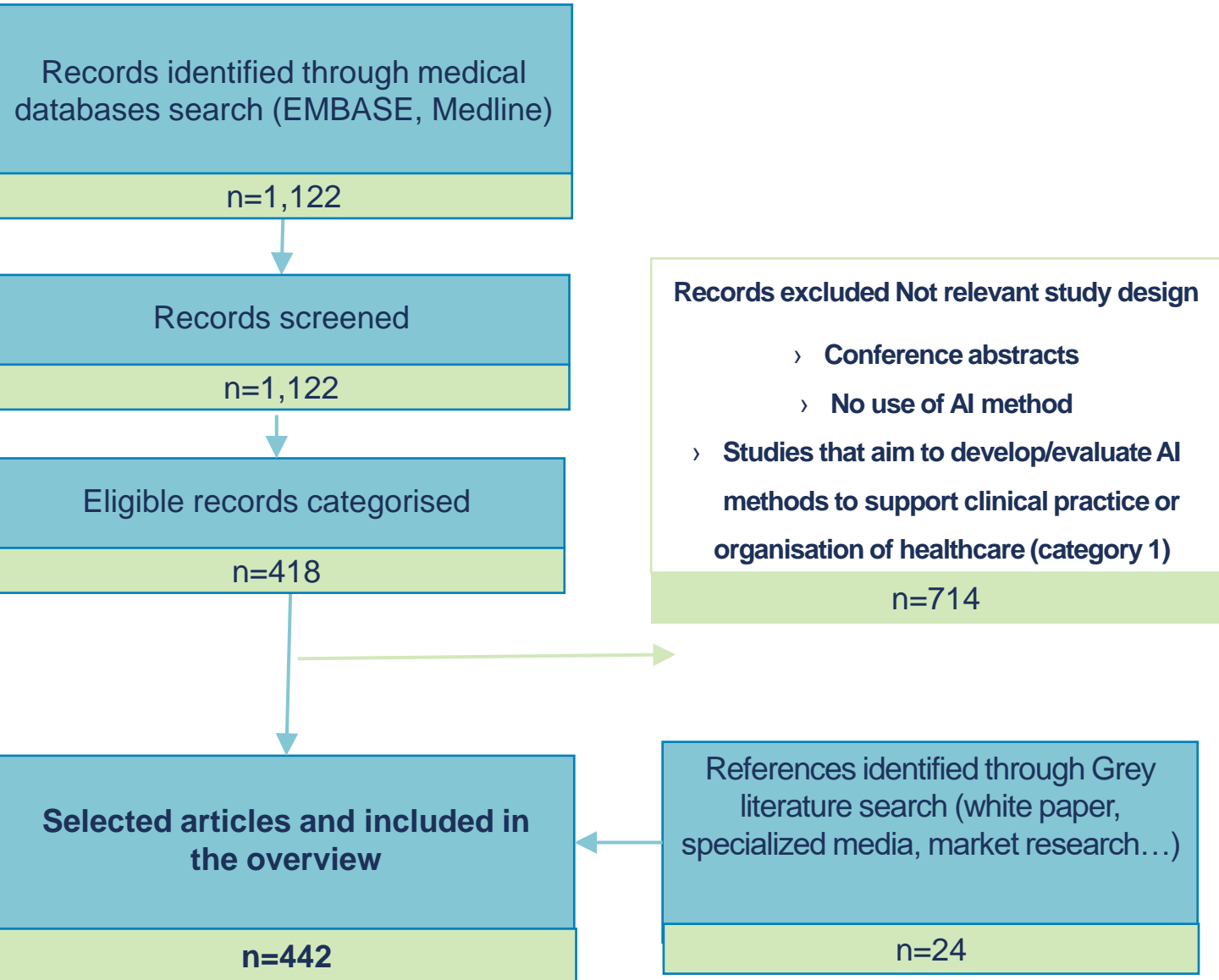
Figure 1. Combined keywords for the search



All RWE studies using electronic medical records or claims database, where use of AI in the study was explicitly reported in the study abstract were eligible.

Different types of applications of AI were categorized based on titles and abstracts and selected examples were reviewed within each category.

Figure 2. Selection



## RESULTS

The search yields 1,122 hits and 418 articles were selected after screening. Additionally, 24 publications from grey literature were reviewed and included.

Studies were divided into following categories

- Category 1 (n=74/18%): AI is used in study methods to identify population, prevalence of a condition, patients' characteristics – risk factors, treatment patterns, health outcomes and adverse reactions.
- Category 2 (n=314/77%): Research that aim to develop and validate AI tools to assess prognosis, predict survival, identify drug interactions, or identify health outcomes
- Category 3 (n=20/5%): Research that aim to develop and validate AI tools that address challenges of RWE studies, such as imputing missing data, measuring patients' similarity for heterogeneous EMR data, predicting clinical variables not completely reported in EHR and leveraging temporal sequential patterns from EHRs

Figure 3. Category 2 is defined in 5 subcategories

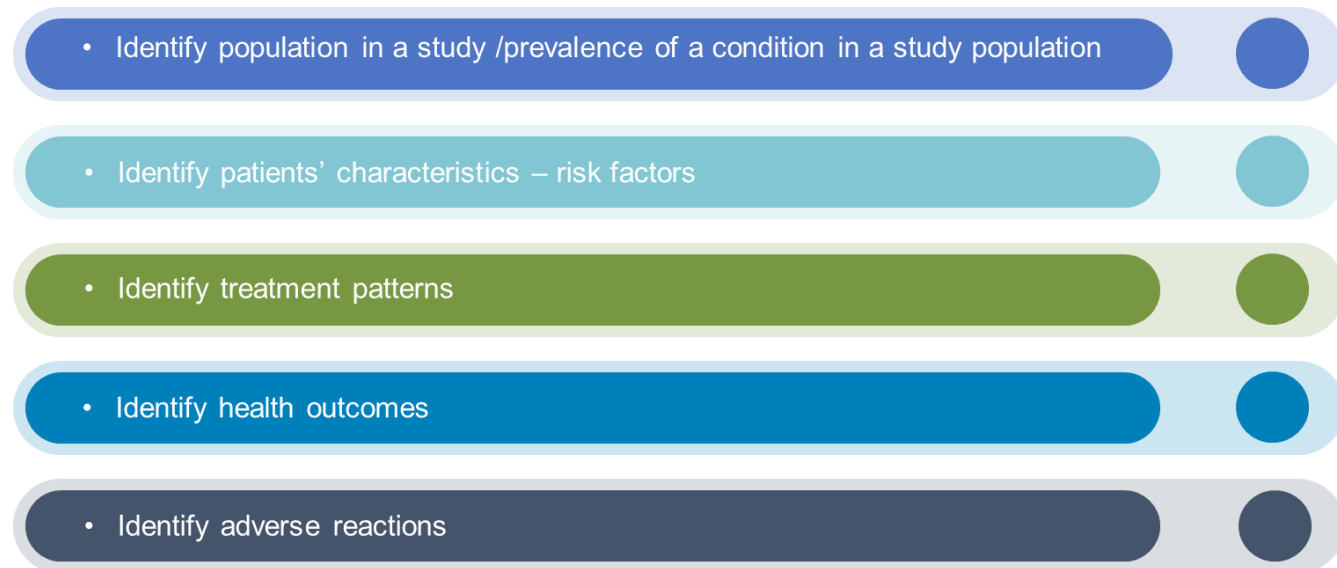


Figure 4. Category 3 is defined in 5 subcategories

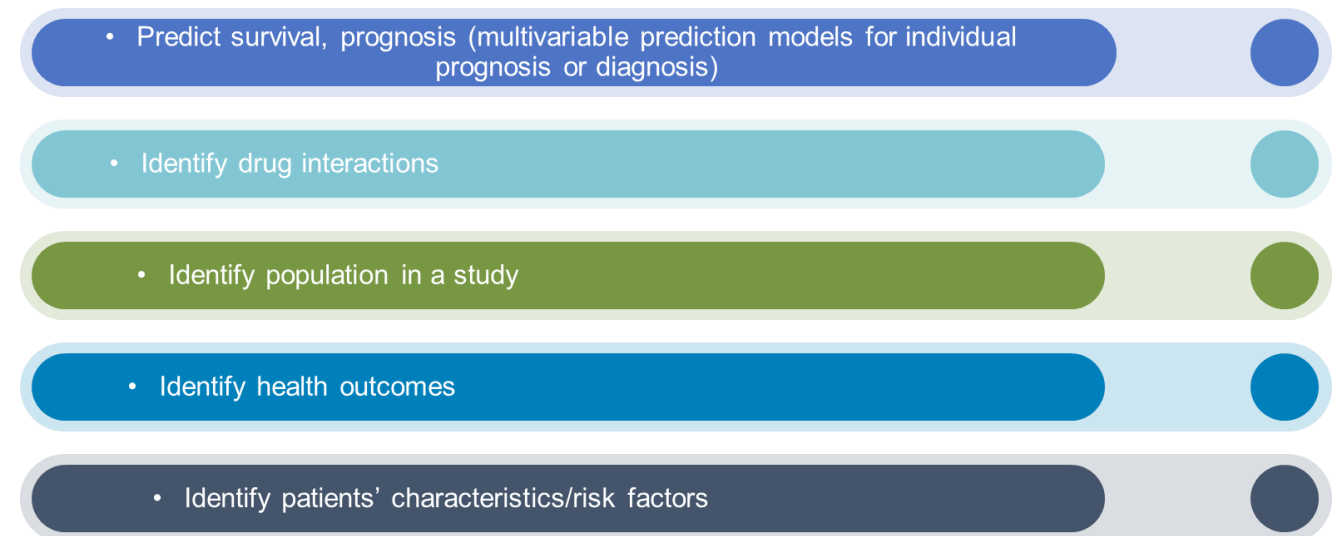
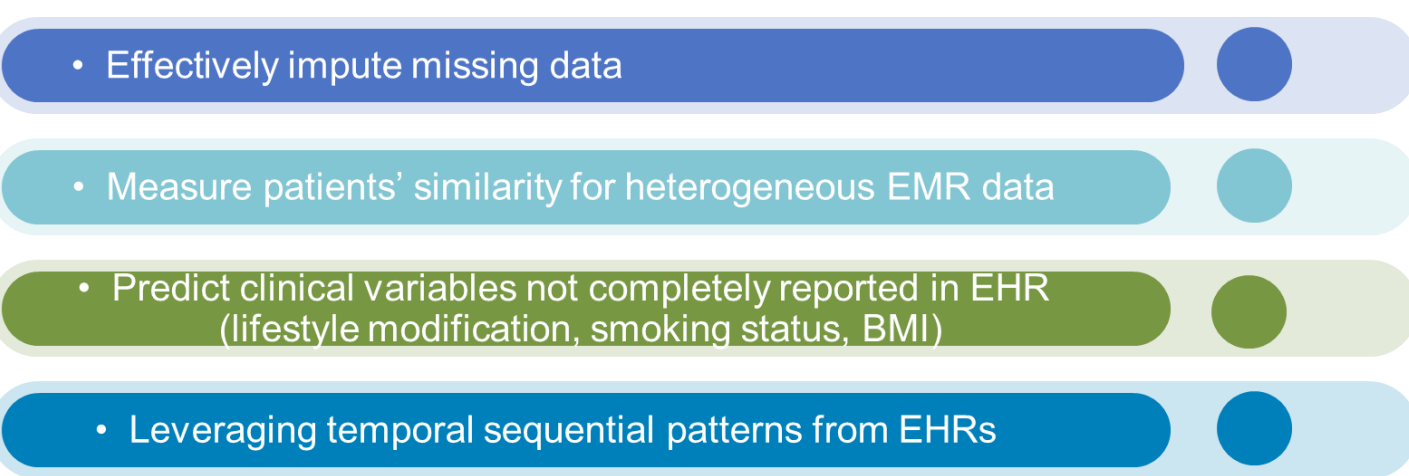


Figure 5. Category 4 is defined in 4 subcategories



The main 4 types of AI methods used were machine learning, deep learning, NLP and process mining.

Key challenges encountered are related to data source, data bias, data aggregation, lack of receptivity and black box effect.

- Data are critical for delivering evidence-based health care and developing any AI algorithm. Refinement, or training, of AI algorithms is essential to improving the accuracy of RWD analysis. However, researchers must first ensure that the data obtained are complete and relevant to the condition, patient population, and treatment analysis
- Low data recall (incorrect inclusion or exclusion of data from an analysis) also appears to pose a great challenge to the use of AI in RWD analysis.
- Lack of receptivity by external stakeholders is a major barrier.
- The black box problem is serious enough that specific regulations are beginning to emerge. The EU's GDPR now includes a right to explanation to deal with algorithmic opacity.

Cross-disciplinary expertise will also be necessary to ensure that the software and subsequent refinements are tailored to each analytical approach.

## CONCLUSIONS

The pace of adoption of AI is accelerating. AI methods provide potential for significant advances in the field of RWE, but key challenges including lack of receptivity and black box effect remain to be addressed.