SWAR 30: What are the benefits and challenges of using machine learning technologies to prioritise screening in systematic reviews?

Objective of this SWAR
The aim of this Study Within a Review (SWAR) is to capture experiences of systematic review teams using machine learning technologies for prioritised screening across a range of review projects to identify advantages and implications for the review process, including project management, and any challenges associated with the use of machine learning technologies in screening. We anticipate that the findings will inform which types of review question might benefit most from prioritised screening, and where prioritised screening might not provide tangible benefits over traditional manual screening methods.

Study area: Study Identification, Project Management, Review Processes
Sample type: Review Authors, Reviews
Estimated funding level needed: Medium

Background
Systematic reviewers are often required to assess the relevance of the titles and abstracts of thousands of records. This can be particularly challenging for reviews with complex or multifactorial research questions, and for reviews which must be completed within a short timeframe.

Prioritised screening uses text-mining and machine learning technologies so that the machine 'learns' to recognise records that are likely to be included or excluded, based on the application of the screening criteria. Machine learning technologies for prioritised screening are available in software such as EPPI-Reviewer [1,2], prediction classifier in Rayyan [3,4], 'Most Relevant' studies in Covidence [5,6] and DistillerSR's AI system [7].

When traditional methods of screening are used, 'relevant' and 'irrelevant' records will be distributed through the list of records found by the search. When machine learning assisted prioritised screening is used, this brings forward records more likely to be relevant to the start of the screening process and pushes records less likely to be relevant to the end [2]. The process is iterative, automatically clustering together records based on the text they contain, and the machine continues to learn the relevance of records, thus updating their ranking, as the number of reviewer decisions increases. The recognised benefits of using machine learning technologies have mainly been related to streamlining and speeding up the process of title and abstract screening.

The York Evidence Synthesis (YES) Group recently completed a systematic review of multi-cancer early detection (MCED) tests for general population screening [8] using prioritised screening within EPPI-Reviewer and found additional unanticipated benefits of prioritised screening across the review process as a whole. Following double screening of titles and abstracts for 10% of prioritised records, the inclusion rate of records started to slow, and the team were confident that a substantial proportion of the relevant evidence had been identified. Therefore, we decided that subsequent stages of the review process (including full-text screening and data extraction) could begin while screening of 'lower priority' titles and abstracts continued. Furthermore, team discussions of the relevance of 'prioritised' records early in the review process resulted in refinement and clarification of originally proposed eligibility criteria, and had positive impacts on the overall project output. This included better team understanding of the research question at an earlier stage, which was reflected in the quality of the final report.

We now intend to routinely use prioritised screening for systematic reviews and to document benefits and challenges associated with the use of prioritised screening, as well as associated changes to the review process, including implications for overall project management. Documentation of the use of prioritised screening processes across a wide range of projects and review teams would help to inform the types of review and review question which are likely to benefit most from the use of prioritised screening, and the types of review where prioritised screening may not provide tangible benefits over traditional manual screening methods.

Interventions and Comparators
Intervention 1: Reviewers will use a machine learning assisted technology within review software for titles and abstract screening to prioritise records identified in systematic searches.
Intervention 2: No formal comparisons will be made during the SWAR.

Index Type: Full Review

Method for Allocating to Intervention or Comparator:
Not applicable

Outcome Measures
Primary: (1) Advantages of using a prioritised screening process, (2) Challenges of using a prioritised screening process, and (3) Impact of the use of prioritised screening on subsequent stages of the systematic review process.
Secondary: Review characteristics will be recorded including review objectives, PICO criteria, search strategies and study selection methods, including the review software used for priority screening.
A PRISMA study flow diagram can also be provided to outline the study screening process.

Analysis Plans
Reviewers can provide descriptive text relating to review characteristics or provide a review protocol containing these details, as well as a PRISMA study flow diagram.

Reviewers will (individually or collectively as a review team) document the process of study screening, including any perceived advantages of the prioritised screening process, any challenges encountered and any perceived impact of the prioritised screening process on subsequent stages of the systematic review process (e.g., full-text screening, data extraction, statistical analysis, report writing). No formal comparisons will be made at a review level but review teams will be asked to consider advantages, challenges, and changes to review processes associated with prioritised screening compared to their previous experiences of conducting systematic reviews using traditional methods of title and abstract screening (without the use of machine learning technologies for prioritised screening).

Documentation of the prioritised screening process could include (but not be restricted to):
- A timeline or Gantt chart of stages of the review process (title and abstract screening, full text screening, data extraction, statistical analysis, report writing etc.)
- Numbers of records double-screened and, if appropriate, single screened, and/or the inclusion rate of titles and abstracts at various stages in the screening process
-- Graphs presenting the screening process can be produced in review software such as EPPI-Reviewer [1]
- Commentary relating to the use of prioritised screening including (but not limited to):
-- Any observed advantages for the screening process
-- Any observed advantages for subsequent stages of the review
-- Any changes to 'usual' review processes (e.g., changes in the frequency of meetings between reviewers to discuss eligibility of records, overlap of subsequent review stages, requirement of reviewers at different times in the review processes)
-- Any challenges associated with the prioritised screening process itself and for subsequent stages of the review.

Reviewer commentary of advantages, challenges and changes / implications can be solely descriptive, or supplemented by appropriate metrics or figures etc. if available.

The YES group has prepared an example of the documented prioritised screening process and impact on the subsequent review process for the systematic review of MCED tests for general population screening [8]. This example is available on request and can act as a template for future reviews, with flexibility to provide additional or alternative information as required.

No comparisons will be made between different projects or between different review teams. Where appropriate, descriptive analyses (e.g. counts and proportions) of review characteristics and elements of the review processes will be presented.

Reviewer commentary will be collated and examined for common themes, as well as experiences which may be unique to the context of the review in question. Formal thematic analysis will be considered, if appropriate.

Possible Problems in Implementing This SWAR
Review software which allows implementation of machine learning technologies for prioritised screening requires a paid subscription for this functionality. The availability of specific software packages at different institutions may vary and may need to be budgeted for. If reviewers are unfamiliar with specific software or the functions of that software, training may be required to ensure appropriate use. This may also need to be budgeted for, as well as time allocated at the start of review processes for this training.

The information collected in this SWAR will be mostly related to reviewer experiences and opinions which are by definition, anecdotal and subjective. However, such information, when collected over multiple reviews, addressing a range of research questions, has value for evaluating the practical applications of machine learning technologies in systematic reviews and for those planning future review projects.

References
[1] Thomas J, Graziosi S, Brunton J, Ghouze Z, O'Driscoll P, Bond M, et al. EPPI-Reviewer:advanced software for systematic reviews, maps and evidence synthesis. London: EPPI Centre, UCL Social Research Institute, University College London; 2022.
[2] Thomas J. Diffusion of innovation in systematic review methodology: why is study selection not yet assisted by automation? OA Evidence-Based Medicine. 2013;1(2):12
[3] Mourad Ouzzani, Hossam Hammady, Zbys Fedorowicz, and Ahmed Elmagarmid. Rayyan — a web and mobile app for systematic reviews. Systematic Reviews 2016;5:210.
[4] Rayyan Prediction Classifier. Available from: https://help.rayyan.ai/hc/en-us/articles/4406419384081-Rayyan-Prediction-Classifier (accessed 20 March 2024)
[5] Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org
[6] Machine learning – the game changer for trustworthy evidence. Available from: https://www.covidence.org/blog/machine-learning-the-game-changer-for-trustworthy-evidence (accessed 20 March 2024)
[7] DistillerSR AI. Available from: https://www.distillersr.com/products/distillersrai (accessed 20 March 2024)
[8] Wade R, Nevitt S, Liu Y, Harden M, Khouja C, Raine G, Churchill R, Dias S. Multi-cancer early detection tests for general population screening: a systematic literature review medRxiv 2024.02.14.24302576.

Publications or presentations of this SWAR design
ESG SWAR subgroup webinar, 10th April 2024
Research on Research Festival, 14th May 2024: Machine learning in systematic reviews: the hidden benefits of prioritised screening

Examples of the implementation of this SWAR
Wade R, Nevitt S, Liu Y, Harden M, Khouja C, Raine G, Churchill R, Dias S. Multi-cancer early detection tests for general population screening: a systematic literature review medRxiv 2024.02.14.24302576.

People to show as the source of this idea: Sarah Nevitt, Rachel Churchill, Sofia Dias
Contact email address: sarah.nevitt@york.ac.uk
Date of idea: 24/01/2024
Revisions made by:

Date of revisions: