SWAR 21: Manual versus semi-automated abstract screening in a systematic review assessing numerical formats for communicating risk

Objective of this SWAR

To compare the set of studies selected for inclusion in a systematic review by two different screening approaches: manual screening done by one reviewer (first reviewer) versus semi-automated screening done by another reviewer (second reviewer) using Research Screener.

Study area: Study Identification
Sample type: Studies
Estimated funding level needed: Unfunded

Background

The substantial growth in the number of scientific papers highlights the importance of high-quality systematic reviews (SR) for synthesizing the best available evidence to inform healthcare decisions. An essential aspect of a high-quality SR is an exhaustive and comprehensive search for potentially eligible studies to avoid missing relevant information. Therefore, search strategies should prioritize sensitivity [1], even if it means dealing with a large number of articles for screening. Additionally, it is preferable to have two independent reviewers perform title and abstract screening to ensure accuracy. However, this process is time-consuming and increases the workload for reviewers.

There is growing interest in using machine learning tools to expedite the screening process in SR while ensuring their validity [2]. Several semi-automated tools have been suggested and examined for this purpose [3,4]. One such tool is Research Screener, a web-based application and algorithm designed to partially automate the abstract screening phase in order to identify relevant articles for a specific SR. When Research Screener was used in a sample of nine systematic reviews, the percentage of articles needing to be screened to find all relevant articles ranged from 5% to 35%, with a mean of 13%, which substantially reduced the workload associated with the screening process [5].

Although Research Screener has already been used in evidence syntheses [6,7], we are not aware of any comparison to fully manual screening in a real-time assessment by researchers not involved in the development and validation of the tool. Thus, independent and prospective assessment of Research Screener is lacking.

Interventions and comparators

Intervention 1: First reviewer: manual. The first reviewer will manually screen the records identified in the searches for a SR, select the articles for full-text reading, read the full text of these articles, and select those for inclusion in the review.
Intervention 2: Second reviewer: semi-automated. The records identified in the searches for the SR will be imported into Research Screener, which will then be provided with 12 seed articles (i.e. articles that have been judged as highly relevant or representative of eligible studies based on the SR inclusion and exclusion criteria). The Research Screener algorithm will use these seed articles to rank all the titles/abstracts by relevance and will present the top 50 titles/abstracts in the first round. The second reviewer will then screen these 50 titles/abstracts, selecting those that should be read in full, and the results will be fed back into Research Screener (i.e. the reviewer will tell Research Screener which titles/abstracts were selected to be read in full). This allows Research Screener to re-rank the remaining titles/abstracts and determine the next top 50 most relevant titles/abstracts for the second round. This iterative process will continue until 10% of the titles/abstracts have been provided by Research Screener, when we will compare whether the set of selected articles for inclusion by the second reviewer is the same as the set selected by the first reviewer. The reviewers will discuss and disagreements will be resolved by consensus and/or with the help of a third researcher in order to reach the final set of included articles. If the second reviewer missed any article identified by the first reviewer, the semi-automated screening will continue. However, to make this stage more efficient, reducing the number of titles/abstracts to be screened by the second reviewer, they will only check, in each round of 50 articles, how many articles would need to be read to find all the articles in the final set of included studies. This

assessment is important to evaluate the extent to which the machine learning tool would increase the efficiency of the title/abstract selection stage in this SR.

Index Type: Searching

Method for allocating to intervention or comparator
Non-Random

Outcome measures
Primary: The final set of included articles (after disagreements between reviewers are resolved); number of articles that were selected for inclusion by the first reviewer (before disagreements are resolved); percentage of articles selected for inclusion by the first reviewer (before disagreements are resolved) among the final set of included studies; number of studies selected for inclusion by the second reviewer (before disagreements are resolved); and percentage of studies selected for inclusion by the second reviewer (before disagreements are resolved) among the final set of included studies.
Secondary: Number and percentage of titles/abstracts that will need to be read by the second reviewer to identify the final set of included studies; number and percentage of studies missed by the first reviewer, in the final set of included studies; and number and percentage of studies missed by the second reviewer, in the final set of included studies.

Analysis plans
Descriptive analysis using counts and proportions.

Possible problems in implementing this SWAR
The search for studies for the SR in which this SWAR will be embedded is difficult because there are no MeSH terms related to its subject (risk communication). Even though 22,362 articles have been retrieved in the electronic search, this is still missing some studies that are already known to be eligible for the SR.

References
1. Lefebvre C, Glanville J, Briscoe S, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, et al (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane, 2022.
2. Beller E, Clark J, Tsafnat G, et al. Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). Systematic Reviews 2018;7:77.
3. Gates A, Guitard S, Pillay J, et al. Performance and usability of machine learning for screening in systematic reviews: a comparative evaluation of three tools. Systematic Reviews 2019;8:278.
4. Valizadeh A, Moassefi M, Nakhostin-Ansari A, et al. Abstract screening using the automated tool Rayyan: results of effectiveness in three diagnostic test accuracy systematic reviews. BMC Medical Research Methodology 2022;22:160.
5. Chai KEK, Lines RLJ, Gucciardi DF, Ng L. Research Screener: a machine learning tool to semi-automate abstract screening for systematic reviews. Systematic Reviews 2021;10:93.
6. Quach S, Michaelchuk W, Benoit A, et al. Mobile heath applications for self-management in chronic lung disease: a systematic review. Network Modeling and Analysis in Health Informatics and Bioinformatics 2023;12(1):25.
7. Pung L, Moorin R, Trevithick R, et al. Determining cancer stage at diagnosis in population-based cancer registries: A rapid scoping review. Frontiers in Health Services 2023;3:1039266.

Publications or presentations of this SWAR design

Examples of the implementation of this SWAR

People to show as the source of this idea: Paulo Nadanovsky, Ana Paula Pires dos Santos, David Nunan
Contact email address: paulo.nadanovsky@gmail.com
Date of idea: 3/JUL/2023
Revisions made by:
Date of revisions: