Reviewer's report

Title: A Survey of Systematic-Review Automation Technologies

Version: 2
Date: 17 April 2014
Reviewer: Clive E Adams

Reviewer's report:

Minor comments
1. It should be mentioned both in abstract and introduction that this paper has been focused on systematic review of randomized/controlled clinical trials and some parts of the content might not be applicable to systematic reviews in other systematic reviews.

2. An operational definition of ‘automation’ might be required as the inclusion criteria of the studies. It will be helpful to understand if the paper considers automation and semi-automation as synonyms.

3. Development of the new registry system (Cochrane Register of Studies = CRS) in Cochrane Collaboration is a missing part of this paper. It worth noting that CRS is related to some of the tasks and should be mentioned in the paper as the process of de-duplication (references and studies) and studification (assigning related references of one study to the same study or as you defined “Study De-Duplication” or merging) could be done semi-automatically by this system based on either field-based setting by user or fuzzy searching (user defines the percentage of similarity). Also, there are some registers in Cochrane groups with valuable coded studies. It means trial search coordinators of some of Cochrane groups have already extracted PICO information (so-called coding) of many studies manually which could be a valuable collection instead of medical dictionaries for getting the potential titles for systematic review, research gaps, number of available trials for a specific title (helpful in prioritization of potential review questions), automatic coding, developing automatic search strings, etc.


   B. crso.cochrane.org (online CRS which is helpful for getting drug names from DrugInfo and design the search strategies related to drugs automatically)

4. Having a register of trials in each major topic is helpful and makes the process of automation so much more specific and efficient. Mentioning the available trials registry systems listed in WHO, CRS, EU-Psi, and the systems by which they are held – Procite, EndNote, MeerKat etc., would, perhaps, add value.

5. Perhaps a missed piece of the puzzle is dealing with language bias. The paper majorly is focused on English trials. Automation of translation and developments in this field could be mentioned. i.e. DOI: 10.1186/2046-4053-2-97

Also, almost all of OCR systems have problems with Arabic, Chinese, and Hebrew characters.
6. In the task related to getting the full text of trials ‘Ferret’ has been missed as a
software program which is able to connect to online library catalogs and locate
the closest physical location of hard copy and probably an automatic ILL request
for the copy in the future:


I know there are other systems as well but they might be worth a mention.

7. Search strategy converter/runner could be another suggestion. If might be
possible to convert EMBASE search strategy for other resources and search all
databases by pressing one button when the program is running in the university
campus.

This might be worth a consideration


10. De-duplication in EndNote is almost a semi-automated process and there are
some related papers about it worth citing: DOI: 10.13105/wjma.v1.i3.97; DOI:
10.1371/journal.pone.0071838

Also for rule-based de-duplication see DOI: 10.1093/database/bat086

For future developments in automatic screening, there is an approach we are
currently working on it which is rule-based and weight-based screening of search
results so that having or missing defined keywords will be used as the rules and
existence of keywords in specific field of record, frequency of keywords in the
records and maybe collection, and location of keywords in the field will change
the weight of the records to enter in one of the output collections (reviews,
relevant studies, irrelevant studies, and needs checking).

Implications for practice/research/policy

1. ‘Implications for Journals’ could be a heading of the paper. I think there is a
need for collaboration between major systematic review producers, PROSPERO,
Online Trial Registers, ICMJE, WAME, and EASE to contribute in registration of
trials and systematic reviews. Making it mandatory to provide the registration
number in submission step will be helpful for future of automation systems
(studification), prevention of duplicate works, and standardization of writing trial
reports and systematic reviews. The same applies for following CONSORT and
PRISMA by the authors.

2. Reporting the abstract/text in PICO format or PICO as part of the paper could
be another suggestions.

3. The next implication might be for software developers to develop a software
program for reporting standard format of trials – and, perhaps at least some of
the dataset. So, for example, using an standard XML tagged text export for
production of PDF, an XML export from trials written by the software program
could be parsed by systematic review programs.
4. Some other less related suggestion are as follow:
A register of grey literature
Automation of Summary of Finding for guidelines
Automation of PRISMA flow diagram
Added value for the paper
A comprehensive table of all of current tools for automation to summarize all of available automation/semi-automation tools for systematic reviews will add a great value to this paper. During reading the paper I had to frequently leave the text and look at the reference then search for the tool to see what it is. This table could be designed in a way that makes referring easier.

Very minor revisions
There is a reference error in Page 9
Page 12, the sentence at the end of second paragraph needs revision as it is long. It is better to be “is that they did not achieve” and also “about new trial/study”.
Page 18 first paragraph a reference error

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

None.