Reviewer's report

Title: Boolean versus Ranked Querying for Biomedical Systematic Reviews

Version: 1 Date: 27 June 2010

Reviewer: Julie Glanville

Reviewer's report:

This is a useful and interesting paper which was enjoyable to read. It is highly relevant to current discussions about information retrieval for systematic reviews and HTAs and well balanced in its exploration of different approaches to efficient searching within those demanding areas. It is a very useful cross-over paper which will appeal to at least two audiences. At the moment it is written for one of the audiences, and I hope the authors will be able to make some changes to bring in more of the second audience (systematic review information retrieval specialists).

The paper is not written in the same style as clinical papers, so, for me, it reads as repetitive and over detailed because of the repetition. Also the results are presented before the methods. The format and style of papers is an editorial decision but reducing the repetition in a long paper might make more readers persist – it would be a shame for readers to give up because of the length and a lot of readers who need to read this are used to papers half the length of this one. Reducing the repetition would also allow some space for additional detail on the non-Boolean approaches to searching. If space is a constraint the last set of explorations (section 5.1) could be easily lost as it does not seem to be as important as the other sections - the deconstruction (if not reversal) of the Boolean approach has some interest but felt unrewarding after the earlier parts of the paper.

Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

1. The paper provides a clear exposition of the current approaches to searching for evidence to inform systematic reviews. As this paper seems to seek to span two audiences - one familiar with Boolean searching and one more familiar with non-Boolean approaches - it would be helpful if the same detail and care used to describe the Boolean approach was also applied to non-Boolean. If this can be added then readers will be better informed. I think this can be achieved quite easily by providing more detail on how the ranked query approach works, the software used (Zettair) and how queries are entered. In particular an explanation of how the search elements in Figure 2 are operationalised to achieve a search would be helpful. I can see things are differently presented from a Boolean search but I don’t know how these search elements are put into Zettair and what Zettair does with them. The authors should tell us more about the Okapi BM25 similarity algorithm and why it was used in preference to any other algorithm.
2 P9 para 1, and elsewhere in the paper. The use of the term ‘metadata’ is not clear and the meaning of metadata with examples should be defined early in the paper. It has unfortunate confusions for me with the metadata we expect to be added to webpages, and so the term seems unhelpful at present and the authors may wish to consider a term with fewer additional connotations.

3 Page 17, para beginning ‘As a first step’. The authors describe the failure of the documented searches (in MEDLINE?) to retrieve the records of documents in the review. This is important to explore further, if I have understood correctly what the authors are reporting. Systematic reviews that result from numerous searches in a range of databases will often include records which are not indexed in MEDLINE (as MEDLINE is itself not comprehensive). It is also the case that records which are indexed and available to be found in MEDLINE are not retrieved by the MEDLINE searches used for a review. Those records are often identified at the end of the review when checking back and will have been retrieved by searches of other databases with different indexing practices. Do these two explanations help to clarify the situation described in this paragraph? If these explanations don’t help then, this makes the record set being used for the research rather doubtful, as it is not being retrieved by the searches which were recorded as retrieving some of it.

On p21 the authors again note inaccuracies in the original review search strategies. It would be useful for the authors to discuss how far they trust their own data set and to report on whether they found the strategy structure, as well as the terms used in the strategies, to be adequate for the reviews’ purposes. There is now a checklist for assessing search strategy quality (PRESS) which might be helpful to refer to even though it may be too late to assess these reviews against it (http://www.cadth.ca/media/pdf/477PRESS-Peer-Review-Electronic-Search-Strategies_tr_e.pdf). Relative recall gold standards such as the one compiled for this research are potentially useful but the authors may have encountered some of the issues which can undermine their use – a discussion of these issues would be useful.

4 P20 The discussion of explosion needs to be improved – We need to know definitely whether the Ovid implementation does or doesn’t explode all terms below the main heading. Then the discussion will have a firm basis.

• Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. p.5 There is a focus on the exhaustiveness of systematic review searches, but it would be fair to say that the ideal of exhaustiveness is giving way a little these days to a more pragmatic (resource-limited) approach – so the authors may wish to reflect this. For example, the Cochrane Handbook which guides the production of what are often referred to as the most rigorous and extensive systematic reviews acknowledge that searches can rarely be exhaustive or comprehensive (see Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 (updated September 2009). The Cochrane
2. Jesse McGowan’s surname is spelled incorrectly throughout the paper and references.

3. P5. Golder researched systematic reviews of adverse effects, so it would be helpful to add this small but important detail.

4. P6 Although many searchers do access their database via the Ovid interface, many do not – so this statement should be amended.

5. P8 para 1. It would be helpful to briefly mention other reasons why searches are recorded in detail, as these are probably more important than exact reproducibility at some future point in time. Searches are reported to allow the reader to make an assessment of the extensiveness of the search and the likelihood that it would capture relevant records for the review, and hence the likelihood of reducing bias in the review.

6. Page 9 para 1 This paragraph is difficult to follow and this may be partly because of the use of the term ‘metadata’. Some examples of new metadata fields might clarify the point. However, it would be fair to say that it is generally acknowledged in the systematic review information retrieval community that databases don’t stand still and that searches will be unlikely to be rerun exactly as they were first performed. The reason for reporting searches is not just for rerunning but to allow an assessment of likelihood of bias and recall.

7. Page 9. Para beginning ‘Within the context’; last sentence. When developing strategies for identifying studies searchers will perform various explorations using different combinations of terms and can NOT out the records they have seen previously, and can show which records are removed by an approach. So there are ways to judge whether altering a query leads to an improvement in the answer set and these are routinely used. The real issue is that these are time consuming.

8. Table 6. The changes to the search term limits and combinations have enormous impacts on the precision and sensitivity of the Boolean search: the authors describe this as simplified but it would also help to explain how ‘expanded’ it is as well. For example, replacing ‘adj n ‘with ‘OR’ replaces an implicit focused ‘AND’ with an OR, meaning that sensitivity expands enormously and precision drops. Ditto replacing Boolean AND with ‘OR’ in the same table. The Boolean approach has been deconstructed and original intentions reversed.


10. References: I think ‘Driana Yoshii’ may be ‘Adriana Yoshii’?

• Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:
I declare that I have no competing interests.