Reviewer's report

Title: Optimizing Search Strategies to Identify Randomized Controlled Trials in Medline

Version: 3 Date: 5 January 2006

Reviewer: K Robinson

Reviewers's report:

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

1. Previous Comment 1. I don't feel that the added text fully addresses the issue raised in my previous review (#5) nor is it, I think, entirely correct. It is not clear to me on what basis other studies concerning the development of optimal search strategies are dismissed as being focused on comprehensiveness alone. For instance, precision was considered in my paper referenced, including in the testing of the addition of CROSS-OVER-STUDIES. The authors provide the Lefebvre and Clarke chapter as a rationale for testing the four terms but do not address how this study fits with other work. I think there needs to be a stronger context set for this study in the introduction -- Why these four terms? Have they been tested previously (yes)? Why were other terms that have been tested not included?

Authors' response: We do not set out to dismiss other studies, rather they are out of scope for this paper. We will try to express that more clearly. We set out merely to validate these variants proposed without supporting evidence in Lefebvre and Clarke's chapter. Robinson implies in this comment that the variants have been previously tested.

KR: My comment (copied above) may have been more clearly stated as separate points.

(i) the previous version of the paper did sound dismissive of previous research suggesting that what research had been conducted had focused on comprehensiveness alone. This was an incorrect statement, as pointed out by another reviewer as well. The revised text in this version of the paper is better.

(ii) The second component of my comment was about setting this study within a context of existing research. Why did the authors choose the terms that they did - for instance, 'latin square' was also mentioned in the Lefebvre and Clarke chapter. I was also puzzled as to why the authors chose to go back and test the addition of terms to the original Cochrane strategy rather than modify/test the revised strategy. I was not suggesting that a comparison of results is necessary but rather asking for some further explanation as to the rationale for the design of this study and some interpretation of how this study and its results fit with the other studies. As an example, the revised Cochrane strategy (in OVID) tested the modification or addition of three terms to the two phases of the original Cochrane strategy. An additional term was tested and deemed too imprecise for the top two phases - CROSS-OVER STUDIES - and this term was tested again by the authors. There may be a very good reason to test this term again, and to not test the other terms, but I don’t see or understand that reason from the text.

(Please also see comment 4ii under Discretionary Revisions.)

2. I appreciate in this version of the paper some metric for the cost-benefit tradeoff in the use of the
different terms. I had previously suggested Number Needed to Read and had not thought of adding the time needed for review. This is a nice addition. However, I am still concerned about the bottom line message for reviewers. I believe that the authors' would like to say that in those situations where reviewers decide to use the two phases of the original HSSS, rather than including the more resource intensive phase 3, they should consider adding the term 'versus'. I am not convinced that the results allows any stronger statement and I think that the statement in the conclusions that the addition of 'versus' balances sensitivity and precision is an overstatement.

-------------------------
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. There remain some issues with the text. Some examples:

-- the last line of background in Abstract is an awkward sentence

-- page 3, last line of conclusions in Abstract, I think they mean "weigh"

-- awkward sentence, first line last paragraph, page 8

-- different spellings of author Moeher in text versus in reference list

-- page 14, last sentence in Conclusions, drop 's' on evidence

-------------------------
Discretionary Revisions (which the author can choose to ignore)

1. As point of clarification, I did not ask the authors' to run the strategy with ti,ab. rather than tw. In my first review I asked for clarification as to why the authors used tw, and in the most recent review noted that there may be some confusion between tw and ti,ab. The term tw means results in different searches in the different interfaces for MEDLINE and there was some discrepancy in the terms used in the strategies presented in the Appendix. Thank you to the authors for now making this consistent.

2. page 4 (also in abstract) there is a sentence that the UKCC pilot study determined that the terms in the third phase was too broad to “warrant their inclusion”. The UKCC concluded that we should not include phase 3 in the MEDLINE Retagging Project. In other words, phase 3 is too broad when searching the entire MEDLINE database (see page 80 in L&C Chapter). I think it is misleading to say that this work suggested phase 3 generally not be included as phase 3 may be useful, and even essential (see Brand) for subject specific searches.

3. Related to above, there seems to be a discrepancy between the statements made on page 4 and those on page 11. On page 4, the authors suggest that phase 3 is not worthwhile and on page 11 recommend use of phase 3.

4. In response to a previous comment, the authors clarified how they identified and evaluated the subject specific search strategies. There is also a nice discussion of this issue on page 12 (though I would suggest adding "subject specific" to sentences, as appropriate). I am wondering, however, if it would be useful to provide further details on how they identified and confirmed those reviews that used the original HSSS. Specifically:

(i) how was the search strategy developed for the HSSS? Were other terms like 'Dickersin' and 'optimal' tested?
(ii) How did the authors check whether the reviewers used the original HSSS? Several reviews, as noted on page 12, refer readers to the CRG strategy. A quick look suggests that some CRGs use the revised Cochrane strategy (see point 1 under Major Revisions), such as the Musculoskeletal CRG (see ref 21 of reviews in paper). Other reviews provide details of the strategy - were these checked to confirm that it was the original HSSS? For instance, Cody (ref 5) seems to also use the revised Cochrane strategy (ie, latin square and CROSS-OVER STUDIES are included). Would the use of revised versus original Cochrane search impact the results of this study?

5. page 8, second paragraph - I think the authors mean "detailed subject specific search strategies" in the line comparing 61 and 33 reviews.

6. page 13 - I am not clear how the authors' study provides a basis for them to make statements about an improvement in the conduct and reporting of searching. Likewise the last sentence about use of information specialists. Depending on the definition of information specialists this statement may be true but the pervious paragraph is about the reporting of search strategies.

7. Page 5, first line - While we revised all Phases of the Cochrane strategy (first in OVID, then translated into PubMed), we did not make any recommendations as to what phase(s) reviewers should use. The inclusion of the specific versions of the strategies in the paper reflect space considerations, an issue pointed out by the authors.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: No

Declaration of competing interests:

I am an author of one of the papers referenced in this paper.