Reviewer’s report

Title: An alternative to the hand searching gold standard: Validating methodological search filters using relative recall

Version: 2 Date: 28 April 2006

Reviewer: Carolyn J Green

Reviewer’s report:

Discretionary

My comments are philosophical rather than technical as I found the paper well thought out, executed and written. Well done.

To provide an alternative to the de facto gold standard I expect relative recall to be directly compared. As this wasn’t done I suggest it be mentioned as a strategy for further studies. The paper would be improved by outlining how the conclusion could go from ‘promising’ to ‘acceptable’ for practice. The definition of â€œgold standard™ as the â€œbest™ available test is practically useful and to go off the gold standard should be done after careful consideration of the consequences. It does appear through this analysis that the consequences could be minimal but I appreciate that this is analysis is useful as a proof of concept and the conclusions appropriately couched.

The method that the authors have devised is indirect. The dataset provided by Cochrane reviews was developed with hand searching as part of the search strategy often and for validating the search strategies used. This makes the systematic review an appropriate denominator for exploring the usefulness of relative recall and therefore useful and innovative.

A more efficient standard for developing/validating hedges/filters/search strategies would provide benefits by saving resources and making rigor more widely available to under resourced group. Therefore I think the paper would be more grounded by alluding to these benefits and limitations as part of the discussion.

Without a direct comparison I have remaining questions about how much more efficient this is comparatively. E journals make hand searching much more efficient than in the past. In an electronic age perhaps â€œhand searching™ should be renamed â€œjournal issue by issue searching. Still all these issues do not have to have to be resolved by this one study. Consider extending discussion.

There is a good argument for using the true best gold standard to develop a filter that will be widely used. If clinically relevant trials are missed then the consequences of â€œgoing off the gold standard™ could be consequential. The importance of relevance is therefore worth more explicit discussion. Searches that are informing public health policy and clinical practice warrant rigor and resources and so the emphasis of groups like Cochrane on exhaustive strategies to identify all available primary data is worth reinforcing. Hedges and the importance of validating are worth placing in the big picture of what systematic reviews can contribute to the optimization of human health. In my opinion it would ground the discussion. The work of Haynes and colleagues at McMaster on hedges is worth referencing.

Recall and precision though widely used have been critiqued. The work of Kagolovsky and Moehr are worthy contributions to this discussion.

Generalizability for diagnostic studies is not good given that only 1 systematic review was available. Perhaps a limitations section would be useful in this paper.

The European literature is so rich that I think it is worth mentioning the importance of Embase in avoiding a North American centric perspective.

The category of 30 studies as an â€œother™ category in table 2 looks like insufficient categorization. Consider labeling as a single groups as described in the text.

What next?: Accept after discretionary revisions
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I declare I have no competing interests.