Reviewer's report

Title: Overview of the effectiveness of reminders in improving professional behaviour

Version: 1 Date: 2 February 2012

Reviewer: Duncan Chambers

Reviewer's report:

Please find below some comments and suggestions for the authors:

Major compulsory revisions

1. The authors should explain and justify why they considered reviews with an AMSTAR score >5 to be high quality. As far as I can tell the AMSTAR papers themselves don’t give any guidance on this but Rx for Change classifies scores of 4 to 7 as medium quality and only those with 8 or more as high quality.

2. Linked to this, the summary score by itself isn’t very informative. It would be helpful to have the answers to the individual questions for each review. This could be added to the additional table; a shorthand version would be fine.

3. Vote-counting as a method of synthesis clearly has major limitations, i.e. it takes no account of sample size, study quality or the magnitude of effect. I think the authors need to acknowledge these limitations and justify the approach they have taken. The Cochrane handbook chapter cited by the authors recommends attempting to assess the strength of the evidence for each major outcome. Would it not have been possible to do this for the primary outcome, perhaps using a GRADE-type approach?

4. Conclusions, first sentence: ‘reminder systems are more likely to be successful if they are patient specific and meet the specific needs of the clinical setting...’. This rather seems to come out of nowhere given the limitations of the reviews focused on settings and patient groups. Can the authors justify the statement more specifically using the evidence presented in the overview?

Minor essential revisions

1. The authors should clarify their reporting of the search. The abstract states that MEDLINE, EMBASE, DARE and the Cochrane Library were searched but from the Methods it appears that these databases were searched indirectly via Rx for Change. In practice it probably wouldn’t make much difference but the abstract could give the impression that MEDLINE etc. were searched specifically for this overview. The statement in the second paragraph of the discussion that the overview was informed by ‘a comprehensive search strategy...’ rather adds to the confusion.
2. Report whether a protocol was drawn up in advance for this overview.

3. Figure 2, labelling of years needs improving.

4. Figure 3, it would be more informative to present actual numbers (as in Fig. 2) rather than percentages.

Discretionary revisions

1. Although the overview focused on process outcomes, it would be interesting to know which of the reviews reported health/patient outcomes and how strong the evidence is that reminders can improve these outcomes.

2. Last sentence under ‘reviews of specific settings’: perhaps remove or rephrase the reference to number of studies as this is not necessarily related to quality of the review?

3. Under ‘reviews of specific behaviours’, it would be helpful to have some interpretation of the numerical results from Durieux et al. Was the effect clinically significant?

4. Under ‘reviews of specific patient populations, the authors could comment on the apparent discrepancy between the results of the overall vote count and the findings of the higher quality reviews.

5. The discussion (last paragraph) could reflect on why more narrowly focused reviews are done, i.e. they are more likely to be helpful to decision-makers.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests