Author's response to reviews

Title: Developing Methods to Assess Harmful Effects in Systematic Reviews

Authors:
Dr Heather M McIntosh (hm14@york.ac.uk)
Dr Nerys F Woolacott (nw11@york.ac.uk)
Dr Anne-Marie Bagnall (amb13@york.ac.uk)

Version: 2 Date: 9 Jun 2004

We have revised our manuscript in the light of the reviewers comments. We have changed the title to: Assessing Harmful Effects in Systematic Reviews. We have tabulated information and made the text more concise as both reviewers helpfully suggested. Below is a detailed response to each individual reviewer comment.

Reviewer 1: Yoon Kong Loke

Comment: I would suggest that the text in Results, pp 4 - 7 could be summarized in a table with the following layout...

Response: we have summarised this information as suggested in Table 1.

Comment: The rest of the results section could then concentrate on specific issues with inclusion criteria and quality assessment (or this could be moved to discussion section, with relevant subheadings).

Response: we have taken the first option suggested.

Reviewer 2: Jacqueline Dinnes

(General) Comment 3: not sure if the title conveys the content of the paper.

Response: We have changed the title to: Assessing Harmful Effects in Systematic Reviews.

Comment 4: pg 6 para 1 this paragraph was quite difficult to read - information in table form would help

Response: information now tabulated in Table 1

Comment 5: pg 7 para 2 sentence beginning 'This reflected the difference in the reviews’ objectives....' might be better to describe what they did first then comment on it.

Response: this has been resolved by describing what we did the Table to which the reader is directed at the beginning of the Results section.

Comment 6: pg 11 para 1 advice from clinicians should be sought - presumably because researchers carrying out systematic reviews often don't have sufficient topic knowledge to comment fully on the scope of a review before they've got some way into the actual work?

Response: we have removed the specific reference to clinicians to avoid obscuring the main point
we want to make i.e. the importance of a well-formulated review question.

Comment 7: pg 12 para 2, first sentence doesn't make sense - do you mean that debate continues over the usefulness of quality scales?
Response: we have rephrased the sentence which now appears in Discussion para 4 of the revised document.

Comment 8: pg 4-10. the Results section would significantly benefit from a tabulation of what was found in the reviews.
Response: we have tabulated the information as suggested, in Table 1

Comment 9: pg 4-5. Review objectives. You have included both a description of what was done and some discussion as to what perhaps should have been done for epilepsy and smoking cessation but not for schizophrenia. We don't find out until page 12 that few studies were found for schizophrenia and that an alternative approach may have been preferable. Should include a description of the success/failure of the schizophrenia objectives as you did for the other two reviews.
Response: we appreciate that we appeared to present a conflicting argument regarding the objective of the schizophrenia review. We have decided to remove the suggestion that an alternative approach may have been preferable in this review because that could not have been predicted at the outset and it confused the key point that the original review question was sensible.

Comment 10(i): pg 4-5. Review objectives smoking cessation and epilepsy. For both of these reviews you state that more specific safety issues could have been addressed if the review questions had been more focused, but do not make it clear why this could not be done within the broad scope of the reviews, presumably it was because too many studies were found to make it feasible?
Response: it is well known that addressing a broad question in a systematic review is unlikely to provide answers to specific questions. The reason is not always solely to do with the number of studies found; in the epilepsy and smoking cessation reviews the broad objectives created problems with study selection, quality assessment and synthesis because of the quantity and diversity of studies and the large number of harmful effects they reported. In the revised document we have attempted to emphasise this by stressing the importance of a well-defined review question.

Comment 10(i): You should include a summary of the number of studies that had to be screened for inclusion, and the numbers that were actually included to indicate the scale of work undertaken for each review.
Response: we have put this information in Table 1.

Comment 11: pg 7 para 2 what is the difference between uncontrolled trials and prospective case series? A box describing what you mean by these, and also PEMs and PMSs would be useful
Response: this text is now in Table 1 where it has been rephrased as 'prospective case series and other uncontrolled trials'. The distinction is that we do not know whether the 'other' uncontrolled trials were prospective or not (but we strongly suspect not). We feel that presenting definitions of study designs would detract from the focus of our paper. Definitions of primary study designs (various, as opinions vary) can easily be found elsewhere. We feel that post-marketing surveillance is a common enough concept in this area not to require definition in our paper, however we have included a reference to prescription event monitoring as this may be less well known.
Comment: 12. pg 7 para 2 by how much did the schizophrenia inclusion criteria regarding size and duration reduce the workload? Have you missed anything relevant by restricting in this way?

Response: we have no way of knowing because the review was done only with, not without, the restriction.

Comment 13: pg 8 para 1 you refer to a 'huge volume of data' in the epilepsy (sic) review - it would be useful to know how much.

Response: this information has now been tabulated in Table 1 for all three reviews.

Comment 14: you don't mention any problems in searching for this type of data - isn't this a key component of the review process?

Response: we have pointed out in the Methods section (moved from the Discussion where it was previously) that we intentionally did not address searching in this paper, together with the reason why.

Comment 15(i): pg 10 para 3 you argue that clearly defined questions are better than broad ones, but don't really say how these should be identified - are you saying that a lot more scoping work should go into identifying the right questions?

Response: Yes; we have tried to bring this out more in the Discussion paragraph 2.

Comment 15(ii): The schizophrenia review clearly benefited from review of a trial protocol, but these are known side effects and surely the point of doing a systematic reviews is both to see if adverse events occur more/less often than previously thought and to see if previously unidentified side effects are occurring?

Response: The specific objective of the schizophrenia review was to determine the incidence of these effects, not to identify them. The point of doing a systematic review is not necessarily to see if adverse events occur more/less often than previously thought or to see if previously unidentified side effects are occurring. But the assumption is not uncommon even among experienced reviewers. We have tried to bring this out better in the Background and Discussion sections.

Comment 15(iii): At the same time you say the schizophrenia review was good because it was focused, but later (pg 12) you say that perhaps it wasn't focused on the right events, so not quite sure what approach you are advocating to deciding on a review question.

Response: see Response to comment 9, above. We do not advocate any single approach to deciding on a review question: that needs to be carefully thought through on an individual review basis.

Comment 16: pg 11 para 2. is your point that, where time and resources are limited, inclusion of nonrandomised studies should be directed by review of the data available from RCTs? You're not making your point very clearly.

Response: we have attempted to clarify our point in Results, paragraph 2 under the subheading Study designs; not to be too hasty to dismiss RCTs in the simple belief that observational studies will be a more useful source of information about harmful effects in general. We removed the reference to time and resources as it obscured the key point.
Comment 17: pg 11 para 2. you also say that non-randomised studies may be of 'dubious quality' but that does not mean RCTs will answer the review question adequately as you point out later, on page 13

Response: we think we have made this clearer now in Results, paragraph 2 under the subheading Study designs and Discussion paragraph 3.

Comment 18: pg 12 para 1. you say the smoking cessation review may have provided clearer conclusions if the data from RCTs had been reviewed first, but I thought that had already been done and that was why you didn't include RCTs of effectiveness in your review? If you had looked at this data first, how would it have helped make the decision about which nonrandomised study designs to include? I'm afraid I don't understand the logic. Was it that you would have selected fewer adverse events to look at and so wouldn't have had to look at so many nonrandomised studies, rather than informing on which designs to include?

Response: See Response to Comment 16, above.

Comment 19: pg 12 para 2, third sentence. It is unquestionable that reviewers should think about what the quality assessment is for a priori, rather than carrying it out because they think they should.

Response: we agree that it is unquestionable, but in our experience it is not at all unusual.

Comment 20: pg 12 para 2, final sentence. Can you suggest how the discrimination of poor from better quality studies should be achieved? Your results section suggests that no such quality assessment tool is yet available. In the meantime, perhaps you should list the sort of questions that you found most useful?

Response: No, we are not able to do so at this time. Because of the mismatch we found between the tools and the particular features of harmful effects data, and that most study designs were only vaguely reported, we were not able to collect enough information to identify any pattern to suggest most useful questions to discriminate between studies in this area. We did however highlight this is an area for further research (Discussion, last paragraph).