Author's response to reviews

Title: Systematic reviews of adverse effects: framework for a structured approach

Authors:

Yoon K Loke (Y.Loke@uea.ac.uk)
Deirdre Price (deirdre.price@clinpharm.ox.ac.uk)
Andrew Herxheimer (a.herxheimer@ntlworld.com)

Version: 3 Date: 11 January 2007

Author's response to reviews:


Response to Reviewers’ Comments

General Point: Thank you for the very helpful suggestions. As suggested by both reviewers, we have completely changed the flow of the paper to give a clearer idea of how to conduct a review of adverse effects in a structured manner.

Specific points:

Reviewer: Barney Reeves

Reviewer's Comments:

1. I should state at the outset that I applaud the authors for taking on this subject. It is a very important area.

2. However, I don't think it provides guidance but rather structures the problem. I give more detailed comments for this opinion below. I think that badging the paper as guidance may frustrate and mislead readers.

3. Conceptually, I think that the structure of the paper has not been thought through as clearly as it might have been which makes the task of the reader much more difficult and limits the usefulness of the paper because the existing structure is difficult to assimilate with other guidance on systematic reviews. With a bit more work, I feel the paper could be much better.

4. I think it would help if the structure/headers more clearly followed the steps that are already familiar to reviewers: e.g. setting the question, searching for potentially eligible studies, defining inclusion/exclusion criteria, assessing susceptibility to bias (new recommended Cochrane term, rather than quality assessment), synthesising the data, interpreting the findings, etc.

5. It might also be helpful to set the problem of reviewing adverse effects in the context of other frameworks. e.g. for applying evidence. For example, it is widely assumed that the relative effects of treatment generalise to populations at varying risk of the (primary) outcome. Is the problem with studying adverse effects in RCTs primarily because: (a) this assumption does not hold true; (b) it is difficult to quantify the relative risk of relevant outcomes precisely in RCTs?

6. For me, the central difficulty is the trade-off problem. I feel that this is given relatively emphasis. I can see the value of doing the leg work without reaching a clear conclusion but am not sure that reviewers will.

7. It is brave of the authors to set out such a wide range of examples in Table 1. I agree with the three main areas identified in which systematic reviews of adverse effects are justified but am concerned that some
of the examples aren't ideal (see below).

8. There is general, not serious problem with inconsistent use of terms / language which is likely to make the paper more difficult to understand for the reader than it need be.

Authors' Reply:

Point 1: Yes, revised as suggested.

Point 2: We have changed the title of the paper to reflect that this is a framework.

Point 3 & 4: We have completely restructured the paper to follow the steps that are familiar to reviewers.

Point 5 & 6: We have completely revamped the 'Scope', 'Types of studies' and 'Interpreting Results' sections to concentrate on how the evidence could be applied, and trade-offs considered.

Point 7: We have amended Table 1 to make the examples much clearer.

Point 8: Terms and language have been revised throughout.

Reviewer: Lisa Berro

Reviewer's Comments:

A much more detailed description of the consensus process used to develop this document should be presented.

Table 1 is important to set the context for the paper. However, I found it confusing. For example, in the first section "the margin between benefit and harm is narrow," the 3rd example mentions only benefit and not harm.

p 5-6. In the section "What types of adverse outcomes" why were outcomes categorized by how the information was collected (e.g., diagnosed by physician, self-reported, etc.) I'm not clear how these would help determine the most "relevant" outcomes, although they would help determine which ones are measured most rigorously. Relevance, however, would also be influenced by the severity or reversibility of the outcome.

Page 7 - 9. This section is a bit repetitive as it lays out the same problem -- benefits and harms may be difficult to compare directly -- but doesn't offer any solutions. I'm left with thinking that this comparison should not be made.

Page 12 - 15. This section has a number of problems. All the quality instruments that are referenced are for RCTs whereas the possibility of including studies of other design is acknowledged earlier in the document. There are quality assessment instruments for observational epidemiological studies. If the authors choose to ignore these, they should explain why. I agree that these instruments (as well as many items on RCT instruments) lack an empirical basis, but I was not clear why they were not being discussed.

Page 14: It seems like a criteria (similar to one used for efficacy data) is that that data must be extractable from the source document.

Page 14. Under general principals, study design itself could be added as a question to establish a hierarchy of design (even without specific quality assessment).
page 16-17. Using withdrawal / dropout rates as an outcome measure is an important issue. I'd like to see more discussion of this .. what's the evidence that low rates are associated with underreporting of adverse events?

page 18. What's the evidence for the "rule of 3" - more discussion of reference #21 is required.

Discretionary Revisions (which the author can choose to ignore)

The middle paragraph on page 7 re different study designs that can be used to assess adverse effects seems to fit better with the section on quality assessment (page 12)

page 11. Why was limiting by study design not included in the search strategy? By not doing so, the number of identified studies could be overwhelming. Although, the entire search seems geared towards high sensitivity and low specificity, so this rationale should be explained.

Authors' Reply:

Point 1: We have clarified that this framework was developed at the meetings of the Adverse Effects subgroup

Point 2: We have amended Table 1 to make the examples much clearer.

Point 3 & 4: We have rewritten the sections on outcomes and study selection to guide reviewers on the why the use of certain categories might help in sorting out the data, so that important problems might not be inadvertently missed out. We agree that the section on pages 7-9 are repetitive, and have removed it completely.

Point 5 & 8: We have highlighted the use of the Newcastle Ottawa Scale, and the use of the study design hierarchy in the 'Assessing susceptibility to bias' section.

Point 6: We have added a much more detailed discussion about the potential problems in using withdrawals/ dropouts as an outcome measure.

Point 7: Evidence for the rule of three is now discussed more fully.

Point 9: We have clarified the sensitive nature of the search strategy, and the limitations of using study design as a search criterion.