Title: Improving the Use of Research Evidence in Guideline Development: 8. Synthesis and presentation of evidence

Authors:
 Andrew D Oxman (oxman@online.no)
 Holger J Schunemann (hjs@buffalo.edu)
 Atle Fretheim (atle.fretheim@nokc.no)

Version: 2 Date: 25 September 2006

Author’s response to reviews: see over
Improving the Use of Research Evidence in Guideline Development: 8. Synthesis and presentation of evidence

Responses to comments

John Lavis
Reviewer's report: General observations
The paper addresses an important issue and is well reasoned and balanced. The paper is reasonably well written but requires some additional work to ensure that the abstract and text are more closely linked and that the key messages are well articulated.

Major compulsory revisions
None

Minor essential revisions
1. The authors should ensure they are consistent in the questions that the paper addresses and in the language they use to address the questions. By my count they invoke three variations on the questions:
   a. the title suggests that the questions are about how evidence should be synthesized and presented;
   b. the abstract and results section suggest that the questions are about how existing reviews should be critically appraised (and used - although it's not clear in this context what is meant by use), when and how WHO should undertake or commission new reviews, how the findings of systematic reviews should be summarized and presented to committees responsible for making recommendations, and what additional information is needed to inform recommendations and how this information should by synthesized with information about effects and presented to committees; and
   c. the text on page 7 suggests that the questions are about how existing reviews should be critically appraised, how priority setting should be undertaken, how reviews should be commissioned, and how the findings of systematic reviews should be summarized and presented to committees responsible for making recommendations.

We believe this is consistent. The title does not contain a question and states the broad topic that the paper addresses. We have added the following to the abstract to clarify that we have focused on the four key questions that are stated there, which are, as stated, key questions relevant to the topic of the paper:

Objectives
We reviewed the literature on the synthesis and presentation of research evidence, focusing on four key questions.

The rationale for focusing on these key questions is provided in the background.

2. The authors provide a list of the other types of information needed to inform recommendations, however, their list pertains primarily to clinical practice guidelines. They don't comment on the other types of information, such as the views and experiences of stakeholders, which can be important to developing policy guidance.

The types of information that are needed to inform decisions about clinical, public health and health systems interventions or policies are the same, although the sources and characteristics of these different types of information differ. The views and experiences of stakeholders are important for all three types of recommendations, but this is a source of information and not a type of information and may be relevant to any of the types of information that we have
identified, including information about effects, need, modifying factors, the availability of resources and values. We believe we have captured the key types of information that are needed for clinical, public health and health systems recommendations in the categories we have used, but there may be other types of information that are sometimes important and there are, of course, other ways of categorising information. We have therefore prefaced the categories we use with: “Additional information that needs to be considered in a recommendation includes,” which does not exclude other types of information.

3. The authors don't comment on whether there is an ethical obligation on the part of WHO to take reasonable steps to ensure that a full systematic review is undertaken when a rapid appraisal is chosen as an alternative because of time pressures and to ensure that any deviation between the findings from the full systematic review and the rapid appraisal is assessed in terms of whether the recommendations warrant changing.

This is a good point, although we prefer not to address it in terms of WHO’s ethical obligations. We have made the changes shown in italics to the discussion:

The available evidence suggests that, generally, in situations where time or resources are limited, thorough quality assessments should likely take precedence over extensive literature searches. When a full systematic review is not undertaken, for example because of the need for a rapid response, explicit consideration should be given to the need and urgency of undertaking a full systematic review and putting in place appropriate mechanisms for timely updating of the recommendations.

We have also added this to the key messages in the abstract:

- When time or resources are limited it may be necessary to undertake rapid assessments. The methods that are used to do these assessments should be reported, including important limitations and uncertainties and explicit consideration of the need and urgency of undertaking a full systematic review.

Discretionary revisions
4. The authors should address a number of minor wording and formatting issues:
   a. bullet points 1 and 2 in the abstract should likely be reversed;
   b. bullet point 1 in the abstract talks about the relevance of the review to the questions being asked, however, this point is not mentioned in the text;
   c. in the abstract the authors provide a list of other types of information needed to make recommendations and it would help the reader if they always maintain the same ordering as in other papers -- "factors that may modify effectiveness in specific settings, need (prevalence and baseline risk or status), availability of resources, costs, and values" -- and as elsewhere in the text (e.g., pages 5 and 7);
   d. in the first paragraph of the text (on page 4) the authors use the words "summary," "synthesis," and "reviews" in ways that suggest they are synonyms but in other places in the text the words seem to be used to mean different things;
   e. the second-from-last and last sentences in the background section (on page 4) seem to be making similar points;
   f. on page 5 "be WHO committees" should be "by WHO committees;"
   g. on page 6 "produce guidelines" should likely be "produce clinical practice guidelines to make clear what type of guidance is being discussed;
   h. on page 6 the authors should use the new name of NICE;
   i. on page 6 the authors may want to re-word the awkward sentence about
NICE reviews being available on its website;  

j. on page 6 the authors should remove the implication that AHRQ produces guidelines, which as far as I know it no longer does;  
k. on page 7 the phrase "and used" could likely be dropped from the first subheading in the Finding section given this domain is not touched on;  
l. on page 8 the authors should clarify the difference between study quality and study validity in this context;  
m. on page 9 the authors distinguish between guidance about clinical and population interventions whereas in other papers they distinguish among clinical, public health and policy recommendations;  
n. on page 9 the authors discuss priority setting but don't make clear how this relates to the overall focus for their paper and they don't include seminal work about priority setting for research relevant to managers and policymakers (such as Lomas et al. in The Milbank Quarterly);  
o. on page 9 the authors introduce a series of three questions but the questions don't appear to map onto the paragraphs that follow and the paragraphs that follow often contain information (such as how extensively one searches for studies) that don't appear related to the focus of the paper;  
p. on pages 9-11 the authors have written an extremely long paragraph that could be broken into smaller paragraphs;  
q. on page 10 the authors argue that "the interdisciplinary nature of the research [on non-clinical interventions], use of research designs other than randomized trials, and limitations of what and how the research is indexed" whereas they may want to clarify that the challenge lies in where and how the research is indexed and use disciplinary mix as a reason why many databases can be important and the use of designs that are not consistently indexed (unlike trials) as examples to support their point;  
r. on page 10 the authors say "public health and areas other than clinical medicine" but they may want to say "public health and other non-clinical areas;"  
s. on page 11 the authors say that "half the guidelines were outdated in 5.8 years" and "no more than 90% were still valid after 3.6 years" but the wording "no more" is very awkward because on first reading the two pieces of information seem logically impossible;  
t. on page 13 the authors make normative statements but they should clarify that the normative statements are being made by others, not themselves (which is implied by the citation but could be made clearer); and  
u. on page 13 the authors may want to clarify that "these should adhere to standards" means "these full reviews should adhere to standards."

We have taken into considerations each of the comments above and made most of the suggestions changes. Thanks very much for the careful reading and thoughtful comments.

Reviewer: David Tovey

General

Thanks for sending this to me for review. It is an impressive piece of work, and my comments need to taken in the context of only having read one other of the sections. Some of my queries may therefore have been addressed elsewhere. I have tried to keep
my comments to content, but I have also struggled with some of the language, so will also mention these as they arise, since they probably reflect my ignorance - which might be shared by other target readers.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)
None

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)
Page 6: NICE is the National Centre for Health and Clinical Excellence

We have corrected this.

Page 13: The statement describing standards for systematic reviews is called the Quorom statement.

We have corrected this as well, thank you.

Discretionary Revisions (which the author can choose to ignore)
Abstract: No major comments, but I am surprised that the context of present WHO guideline practice, which reflects the importance of this work is not mentioned within the abstract.

We have added a new background section to the abstract of all of the articles in the series:

Background
The World Health Organization (WHO), like many other organisations around the world, has recognised the need to use more rigorous processes to ensure that health care recommendations are informed by the best available research evidence. This is the eighth of a series of 16 reviews that have been prepared as background for advice from the WHO Advisory Committee on Health Research to WHO on how to achieve this.

Background / What WHO is doing now: No major comments, but a very clear description of the need to improve the use of evidence based resources in setting guidelines.
What other organisations are doing / methods:
no major comments
Findings:
I found this section difficult to understand even on repeated re-reading. In particular the "five systems" that "represent best practice" do not seem to be described clearly enough - I presume they are the original seven domains minus data abstraction and search strategy, which if true seems a surprising conclusion.

We have clarified that the ‘five systems’ refer to the four checklists and one scale that were considered the best available instruments by the authors:
The authors concluded that based on coverage of the seven domains that they considered key, these five systems (four checklists and one scale) represented “best practice” (i.e. were the best available instruments) for appraising systematic reviews.

In the succeeding paragraph I would have liked more information on why the AMSTAR was chosen as the preferred instrument. As it presently appears, it reads almost as though the authors are inviting the reader to trust a process that is insufficiently described. Some description of the rationale would have been welcome, albeit I understand that it would cause the document to be longer.

We agree, but a report with that detail has not been published. We have clarified this by adding a sentence at the end of this paragraph:

Based on the second expert consultation, the AMSTAR 2005 was selected as the best instrument for appraising systematic reviews (Box 1). A description of the rationale for selecting that instrument is not available.

When and how should WHO undertake or commission new reviews?:
The description of research findings and conclusion that more limited database searching could be recommended was well argued, as were the arguments for different skills and exploration for subjects closer to Public Health and non clinical areas. Notwithstanding my obvious conflict of interest I would have been interested for the authors to consider whether alternative secondary source synopses (including Clinical Evidence) could be utilised where adequate systematic reviews either do not exist or are obviously out of date.

Neither the papers we reviewed or we explicitly considered this question. Checking secondary source synopses might sometimes be a useful component of a search strategy, but do not provide sufficient information for critical appraisal using an instrument such as AMSTAR.

I was interested in the evidence on frequency of updating, since I believe that this is an essential area of concern with respect to credibility and accuracy. The reporting of the Cochrane study does not provide any information as to whether the updated reviews can be considered representative of the whole. One could hypothesise either that those with either less material to update, or more, might be more likely to be updated. [40] The AHRQ review that reported that only 3 of 17 guidelines remained valid after an unreported duration presumably reflects the fact that multiple systematic reviews are required for any given guideline, and this therefore seems to argue for more frequent updating. In addition, there will be substantial heterogeneity in the growth of literature between topics, as the authors acknowledge in the discussion, and this perhaps could also be mentioned in this section.

We agree and have taken ‘under three years’ out of the relevant bullet point in the abstract and stated:

Consideration should be given to undertaking or commissioning a new review whenever a relevant, up-to-date review of good quality is not available.

In the discussion we have added more emphasis to the brief statement we had that “The frequency with which reviews or guidelines need to be updated is likely to vary,” by making
the rest of that sentence more consistent with this, without going into detail about reasons why it is likely to vary:

but as a rough rule of thumb, based in part on a study of clinical practice guidelines, the need for updating should be considered routinely after three years and more often for areas that are developing rapidly.

How should the findings of systematic reviews be summarised...:
I can understand that the authors might have described the GRADE approach within another document, but I think the inexperienced author might have valued some additional explanation of the guiding principles and values within the GRADE process.

We have added a reference to the article in this series where this is discussed.

What additional information....:
No major concerns, although I believe that within the GRADE group there have been discussions about methods to explore local patient preferences and some work relating to this, that could have been reflected in this section, if it is not covered elsewhere.

We have added a reference to the article in this series where this is discussed.