

## **Author's response to reviews**

**Title:** Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.

### **Authors:**

Beverley J Shea ([bevshea@uottawa.ca](mailto:bevshea@uottawa.ca))  
Jeremy M Grimshaw ([jgrimshaw@ohri.ca](mailto:jgrimshaw@ohri.ca))  
George A Wells ([gawells@ottawaheart.ca](mailto:gawells@ottawaheart.ca))  
Maarten Boers ([mboers@vumc.nl](mailto:mboers@vumc.nl))  
Candyce Hamel ([chamel@uottawa.ca](mailto:chamel@uottawa.ca))  
Lex Bouter ([lbouter@vumc.nl](mailto:lbouter@vumc.nl))  
Neil Andersson ([CIEtinter@compuserve.com](mailto:CIEtinter@compuserve.com))  
Ashley Porter ([ashleyclaireporter@hotmail.com](mailto:ashleyclaireporter@hotmail.com))  
David Moher ([dmoher@uottawa.ca](mailto:dmoher@uottawa.ca))

**Version: 2 Date:** 20 November 2006

**Author's response to reviews:** see over

November 20, 2006

Dear BioMed Central Editorial Team and Referees,

Thank you for the opportunity to respond to suggestions to improve our manuscript entitled '**Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.**'

We would also like to thank our reviewers for their constructive comments. We believe there is a need for a new methodological quality assessment instrument for several reasons. The Overview Quality Assessment Questionnaire (OQAQ) in our view is one of the best instruments (out of more than 2 dozen) to assess the quality of methods. An enhanced version was published 5 years ago in a BMJ book, but is not widely known or used.

We believe it's not good enough and decided to base a new tool on the OQAQ (enhanced) and Sacks (which came out in a previous review as another more or less acceptable instrument). A 37-item assessment tool was formed by combining 1) the OQAQ scale, 2) a checklist created by Sacks, and 3) three additional items recently judged to be of methodological importance. This tool was applied to 99 paper-based and 52 electronic systematic reviews. Exploratory factor analysis was used to identify underlying components. The results were considered by international methodological experts using a nominal group technique.

There is a major push amongst methodologists to improve the methodological quality of systematic reviews, while simultaneously improving reporting bias. There are a number of instruments available to do this, but not all have been empirically developed and are therefore evidence-based measurement tools. Our instrument addresses the key issues of efforts made by authors to minimize bias. We hope our arguments will help resolve this, by providing systematic reviewers access to this valid and reliable methodological quality assessment instrument. We admit that we contribute an incremental improvement over existing methods, but we believe it's relevant and should be published.

Based on the excellent comments received from our referees, you will find a revised version of the manuscript. We thank the editors and reviewers for their help. We very much hope that the paper will be accepted in the current version.

Best wishes,  
Beverley Shea

## Appendix 1

### Reviewer's report

**Title: Development of AMSTAR: a measurement tool to assess reviews.**

**Version: 1 Date:**

**11 September 2006**

**Reviewer #1: Graham Mowatt**

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

**Comment:** P5 line 6 'assessing' should be 'assessed'.

**Response:** Thank you for pointing this out—the word now reads assessed.

**Comment:** p5 line 13 there is a reference to Appendix A but there was no Appendix A in my version of the paper.

**Response:** I have removed the word appendix A.

**Comment:** p6 line 14 'processed' should be 'process'.

**Response:** The word now reads process.

-----  
Discretionary Revisions (which the author can choose to ignore)

**Comment:** p3 3rd para 3rd sentence 'the available instruments...are not particularly user friendly'. In terms of? Suggest adding a few words at the end of the sentence providing more information on why the available instruments are not considered to be user friendly.

**Response:** The sentence now reads: 'Nevertheless, a majority of the available instruments are not widely used. Several are lengthy and include complicated instructions for their use.'

**Comment:** p4 line 11 suggest giving more information on what is meant by 'good' face and content validity or just removing the word 'good'.

**Response:** The word good is removed.

**Comment:** p16 Table 2 question 4. In all of the other questions the optimal answer is 'Yes'. Suggest rewording question 4 likewise to make it consistent with the others, e.g. were reports included irrespective of their status of publication (i.e. grey literature was not used as an exclusion criterion)?

**Response:** This is a very helpful suggestion, the question is reworded as follows: '**Were reports included irrespective of their status of publication (i.e. grey literature was not used as exclusion criterion)?**' The authors should state that they searched for reports

regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.'

**Comment:** p17 question 11. Conflict of interest may extend beyond sources of support. Suggest amending the sentence along the lines of 'Sources of support and other potential conflicts of interest should be clearly acknowledged both in terms of the systematic review and the included studies.'

**Response:** This question now reads as follows: '**Were potential conflicts of interest stated?** Sources of support and other potential conflicts should be clearly acknowledged both in terms of the systematic review and the included studies.'(Page 16)

We thank our reviewer for his positive review. These suggestions have certainly helped improve our paper.

## Appendix 2

### Reviewer's report

**Title: Development of AMSTAR: a measurement tool to assess reviews.**

**Version: 1 Date: 29 August 2006**

**Reviewer 2: Joseph Lau**

**1. Comment:** The manuscript describes the development of a new tool to assess the methodological quality of systematic reviews. This tool builds upon items from a previous scale and another checklist. Factor analysis was conducted on these items using a database of 99 paper-based reviews and 52 Cochrane systematic reviews and identified 11 components. One item from each component was selected by an expert panel, which then forms the new tool. This work represents minor tweaking of prior work by others and provides no new insight into assessing the quality of systematic reviews. Even though the objective of the research was to develop a user-friendly tool, the final list of items has not really been thought through as to the user friendliness or practicality.

**1. Response:** We appreciate Dr Lau's views and understand the criticisms that he is making. We feel that he is applying a test of originality that is beyond the general aims of this project. In our view his recommendation that the paper is "too small an advance to publish" would be justified if we were claiming a major original breakthrough, but that was not our intention. Our aims were to achieve to consensus and a degree of standardisation and validation in a field that is currently characterized by lack of agreement. Despite the advances that have been made in systematic review methodology, there is no general agreement on a quality assessment instrument for systematic reviews.

Existing tools did not reflect current evidence on systematic reviews and were generally not validated (with exception to the OQAQ). Frequently users of Oxman and Guyatt had to develop their own rules for operationalising the instrument that would likely lead to variable use of the instrument and poor reliability. We believe the AMSTAR instrument, which indeed relies heavily on work done by others, will be widely used and will achieve a degree of standardisation.

**2. Comment:** Additional items added to the prior checklist or scale were rather arbitrary (there is no specific criteria to define "advances in the field") and redundant with items already on the list (e.g., publication bias). While publication bias is a major threat, there is no proof that methods of assessing publication bias are valid or that assessing publication bias leads to a less biased review.

**2. Response:** We appreciate the reviewer's comments, but respectfully disagree for a number of reasons. Several important items are not thoroughly assessed by the enhanced OQAQ. New methodological research appeared regularly during the development of our project. We were concerned there were still items missing from the evidence-based pool on which the factor analysis would be conducted. With the consensus and agreement of the expert panel we added the following empirical items with clear descriptions for

evaluation. 1) *Language restriction*: Language restriction in systematic reviews remains controversial. Some studies have suggested that systematic reviews that include only English language publications tend to overestimate effect sizes<sup>1</sup>, whereas other studies suggest that such language restrictions may not do so<sup>2</sup>. An item was added to determine whether a language restriction was applied in selecting studies for the systematic review.

2) *Publication bias*: Publication bias refers to the tendency for research with negative findings to get published less frequently, less prominently, or more slowly, and the tendency for research with positive findings to get published more often. Publication bias has been identified as a major threat to the validity of systematic reviews. Empirical research suggests that publication bias is widespread, and that a variety of methods are now available to assess publication bias<sup>3 4</sup>. An item was added to determine whether the authors assessed the likelihood of publication bias.

Although there remains controversy about the optimal methods of assessing publication, we would argue that it is important that reviews explore the likelihood of publication bias and its implications when interpreting the results of a review. This is an area that has generated substantial methodological interest in recent years and we anticipate that better approaches will emerge. 3) *Publication status*: The importance of including grey literature in systematic reviews has been and continues to be explored by methodologists<sup>5 6</sup>. (Page 4)

We tried to a) be more complete, and b) to have a clear operationalisation. We used an expert panel to design the new instrument (AMSTAR) on the basis of the two methodological instruments, plus additional empirical items, to ensure that our new instrument would cover all important aspects (i.e. items that should be conducted in systematic reviews. We believe that should offer at least some protection against making arbitrary choices. We agree that not yet for all AMSTAR items we have compelling evidence that they bear on review validity and can be assessed adequately. This is now mentioned in the discussion section. (Page 9)

**3. Comment:** While “the tendency for research with positive findings to get published more than once” does lead to bias, this redundant publication is not “publication bias” *per se*.

**3. Response:** We partially agree with our reviewers comment. If indeed there is redundant publication and it is discovered, then there is no bias in the review. If however it is not discovered, then there is double counting within the review which also leads to bias. As indicated above, we reworded the relevant sentence on page 5 as follows: ‘Publication bias refers to the tendency for research with negative findings to get published less frequently, less prominently, or more slowly, and the tendency for research with positive findings to get published more often.’

**4. Comment:** Even though the objective of the research was to develop a user-friendly tool, the final list of items has not really been thought through as to the user friendliness or practicality.

**4. Response:** Thank you for this comment. However, we disagree with the reviewer. Initial feedback has been positive about the instrument. The Canadian Agency for Drugs and Technologies in Health (CADATH)<sup>7</sup> undertook an independent assessment of available quality assessment criteria for systematic reviews. A team of international experts assessed all available instruments and recommended the use of AMSTAR by CADATH. Feedback from CADATH reviewers has been very positive. Further preliminary experience suggests that AMSTAR has good reliability and convergent validity suggesting that appraisers can apply it in a consistent way.

**5. Comment:** Since this tool is specific for systematic reviews, the title should include the term “systematic” before “reviews”.

**5. Response:** We agree with your suggestion to include the word ‘systematic’. The title reads ‘Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. AMSTAR will now stand for ‘Assessment of Systematic Reviews’

**6. Comment:** The selection, composition and size of the expert panel were not given.

**6. Response:** Thank you for pointing this out. Under objective 2 on page 6, we added the following sentence to the text to better describe the selection, composition and size of expert panel: ‘We convened an international panel of eleven experts in the fields of methodological quality assessment and systematic reviews. The group was selected from three organizations involved both in the conduct of systematic reviews and in the assessment of methodological quality. The group was made up of clinicians, methodologists and epidemiologists, and new reviewers to the field. Some individuals were previously involved in the Cochrane Collaboration, while a number were not. By examining the results of the factor analysis, they reflected critically on the components identified and decided on the items to be included in the new instrument. The nominal group process took place in San Francisco during a one day session’.

**7. Comment:** The way that some of the items in Table 2 are framed is trivial, quite subjective or open to various interpretations. For example, #5 “Was a list of studies (included and excluded) provided?” 99%+ of the published systematic reviews I came across listed included studies. For the excluded studies, at what point are we talking about exclusion?

**7. Response:** There are still a number of published systematic reviews that do not provide a list of all included studies. This item is directly related to replication of the review and avoiding bias. We believe it is extremely important to provide a list of all included studies. We agree that the issue of rejected studies is a valid point for discussion. Mostly the flow diagram deals with the number of citations that were retrieved and reviewed. Specifically, for excluded studies we are referring to those studies seriously considered, meaning selected on the basis of title or abstract, but after studying the full text rejected for the review. Journals (e.g. BMJ) often put that list with study details on their web site.

We have revised this question and it now reads: ‘**Was a list of studies (included and excluded) provided?** A list of included and excluded studies should be provided either within the article or on a designated website. Included studies are only those incorporated in the systematic review and/or the meta-analysis. Specifically, for excluded studies the authors should list those studies seriously considered, meaning selected on the basis of title or abstract, but after studying the full text rejected for the review. (Page 15, question 5)

**8. Comment:** Journals typically do not want long list of references of rejected studies.

**8. Response:** Also see our previous response. Journals (e.g. BMJ) often put that list with details of excluded studies on their web site. Alternatively, reviewers can put them onto websites.

**9. Comment:** Statement of conflict of interest is required by all peer reviewed journals I came across, the usefulness of item #11 is doubtful.

**9. Response:** We tend to disagree somewhat with Dr. Lau’s comment on conflict of interest. Recent studies conducted by David Moher<sup>8</sup> suggest that the majority of systematic reviews are published in specialty journals that may not require declaration of conflict of interests. In our preliminary validation work we also found a number of systematic reviews that did not include the sources of support. Similar studies have been published by Lisa Bero<sup>9</sup>. We are even more convinced that funding sources are associated with bias in systematic reviews. What is most important about question #11 is that it asks not just about the sources of support for the systematic review itself, but for all primary studies included in the systematic review.

As noted, conflict of interest may extend beyond sources of support. This question relating to this issue now reads as follows: ‘**Were potential conflicts of interest stated?** Sources of support and other potential conflicts should be clearly acknowledged both in terms of the systematic review and the included studies.’(Page 16)

We would like to thank Dr. Lau for his constructive comments. We respectfully disagree with his suggestion that there is not a need for a new instrument. We agree that we contribute an incremental improvement over existing methods, but we believe it is relevant and should be published.

## References

---

- <sup>1</sup> Egger M, Zellweger-Zahner T, Schneider M, Junker C, Lengeler C, Antes G: **Language bias in randomised controlled trials published in English and German.** *Lancet* 1997, **350**:326-329.
- <sup>2</sup> Moher D, Pham B, Klassen T, Schulz K, Berlin J, Jadad A, Liberati A: **What contributions do languages other than English make to the results of meta-analyses?** *Journal of Clinical Epidemiology* 2000, **53**:964-972.
- <sup>3</sup> Sterne J.A.C., Gavaghan D, Egger M: **Publication and related bias in meta-analysis- power of statistical tests and prevalence in the literature.** *Journal of Clinical Epidemiology* 2000, **53**:1119-1129.
- <sup>4</sup> Sutton A, Duval S, Tweedie R, Abrams K, Jones D: **Empirical assessment of effect of publication bias on meta-analyses.** *BMJ* 2000, **320**:1574-1577.
- <sup>5</sup> Hopewell S, McDonald S, Clarke M, Egger M: **Grey literature in meta-analyses of randomized trials of health care interventions.** In *The Cochrane Library*, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd.
- <sup>6</sup> McAuley L, Pham B, Tugwell P, Moher D: **Does the inclusion of grey literature influence the estimates of intervention effectiveness reported in meta-analyses?** *Lancet* 2000, **356**:1228-1231.
- <sup>7</sup> Canadian Agency for drugs and technologies in health [www.cadth.ca](http://www.cadth.ca)
- <sup>8</sup> Moher D, Tetzlaff J, Tricco A, Sampson M, Altman DG; Epidemiology and Reporting Characteristics of Systematic Reviews. (*under-review PLoS Medicine* 2006).
- <sup>9</sup> Bero LA; Title Managing financial conflicts of interest in research. *Source Journal of the American College of Dentists.* 2005; 72(2);4-9.