Author's response to reviews

Title: Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor

Authors:

Clovis M Faggion Jr (clovisfaggion@yahoo.com)

Version: 5 Date: 27 May 2015

Author's response to reviews: see over
Reviewer 1

Minor Essential Revisions This is a cogently written discussion of the content validity of the AMSTAR instrument. The author is correct in addressing this timely and critical question because the AMSTAR is the prevalently used tool for assessing the quality of systematic reviews, which are key in our present days of the pursuit of the best evidence base presented in systematic reviews, and because any tool of measurement in research must be confirmed for its validity and reliability. Therefore, the intent of the author to establish the content validity of the AMSTAR is both timely and very important to the field. However, the author makes a very simple and dismissive statement about how the AMSTAR is scored (yes, no, cannot tell, does not apply), avoiding the very argument that this, more than anything, is the fundamental weakness of the AMSTAR: its qualitatively scoring based on the opinions of the reader rather than established criteria of excellence. This is a weakness of the AMSTAR because it precludes any and all attempts at establishing reliability of the tool and replicability of measurement (cf., "precision" and "consistency" in AHRQ’s assessment of the risk of bias). Therefore, a revision of the AMSTAR has recently been published, which does not alter the content validity of the original AMSTAR, but provides a quantitative scoring for each of the 11 questions based on the methodological criteria of systematic reviews discussed in the Cochrane, AHRQ and PCORO methodological handbooks. In brief, this paper now under review should at least mention the above discussion when discussing the scoring process of the AMSTAR. At best (for the continuing enhancement and progress of the field), the discussion of content validity ought to have been on the R-AMSTAR, rather than the AMSTAR.

Thank you for this interesting comment. I really appreciate that. I agree with the reviewer that lack of quantitative scores may limit the usefulness of the checklist. For example, a qualitative assessment does not allow the incorporation of SR quality scores into more robust statistical evaluations such as meta-analyses. Nevertheless, this may be a limitation for several tools (for example, Cochrane ROB tool), because it is difficult, in my point of view, to precisely determine the weight of each domain in different clinical issues and settings. So, the question is: providing quantitative thresholds for classifying the SR in levels of quality is reliable? I understand that some attempts for quantifying the AMSTAR scores have been provided in the literature. The idea was to correlate the number of items answered with YES with the level of quality. In other words, the higher the AMSTAR items answered with YES, the higher will be the methodological quality of the SRs. The same rationale is applied to R-AMSTAR, which allows a score of 44 points for the SR with highest methodological quality. I discussed these two sides (qualitative vs. quantitative evaluations) in the background section as the reviewer suggested.

Reviewer 2

Major Compulsory Revisions

I.96 and I.102/3 It is important to note that AMSTAR, although developed to evaluate the methodological quality of SRs, has only been developed and validated by the developers on SRs of RCTs. This should be made clear at the end of the para

Thank you for this comment. I made this information clear in the end of the paragraph.
L.22 it is mentioned in the text that the items will be presented in italics but I can’t find them in the manuscript. It is important for the readers to have the items and their guiding notes in the manuscript.

The AMSTAR items and the notes are reported in italics in the discussion section (numbers 1 to 11).

L.149 I agree that this note is misleading. However, I would appreciate it to have a guidance note whether checking someone’s work (also known as double-check) is sufficient as this is a different procedure when compared to two independent persons. I guess this was what the AMSTAR team might have had in their mind.

Thank you for your suggestion. I made the rationale clearer in lines 164-165

L.188 An increase in sensitivity is not necessarily warranted by searching two sources as it will always depend on which sources are sought (e.g. one big database might be more able to increase sensitivity than two small databases). I suggest to delete "and sensitivity" or to rephrase.

Thank you for this comment. I deleted the word "sensitivity" as the reviewer suggested

L.214-219 it is not clear to me whether the author suggest to define a minimum number of characteristics or to define characteristics that must be included.

I was meaning both: a minimum number of characteristics and the type of characteristics that would be relevant for SRs across different disciplines.

L.334 I am struggling with the term "efficiency of SR quality assessment". What is meant by that? This might be more related to validity I suggest to add a discussion para about the trade-off between more and clearer guidance (as mainly suggested by the author) at the one side and the opportunity of more subjective judgements on the other side. Both side have their pros and cons. For example, setting minimum thresholds will let many people simply count instead of think about the issue under study. Making the tool simple as much as simple makes it applicable to many researchers, including even those with very limited expertise in SE, but this might come at the cost of validity. However, this is not an only AMSTAR-related problem, but applies to all critical appraisal tools.

Thank you for this interesting comment. I added some information about the trade-off between clearer guidance and potential lack of validity of the checklist (lines 237-239).

Minor Essential Revisions

it should read "in a SR" instead of "in an SR" throughout the manuscript
Thank you for your feedback. I asked the professional editor who reviewed the manuscript for English and flow, and she answered the following: "an SR or RCT" is correct because R-C-T/S-R begins with a vowel sound (like the beginning of “an art show”). Just to confirm, I checked several style guides, and they consistently agree on “an RCT/SR”

l.319 insert "the" before "note"

This was already fixed in the sentence.

Discretionary Revisions l.99 "reporting on the use" is not the same as "using" something. Consider rephrasing

Thank you for this comment. The sentence was rephrased as the reviewer suggested.

l.167 consider to add that authors of Cochrane reviews are not routinely expected to perform handsearches. This might be a potential explanation why this was not considered in the note. However, it’s worth discussing it.

Thank you for your comment. In fact, there are some guidelines in the Cochrane Handbook for Systematic Reviews of Interventions for performing handsearching, and also in older versions of the handbook.

l.231-247 The point of the author could be presented in a clearer way. One possibility would be to state that AMSTAR whether there was any quality assessment, while the author has the opinion that it should read whether the authors assessed the quality of the primary studies with an adequate tool. In the current AMSTAR version it is not clear from the item and its guidance how to score a SR of case-control studies when the authors used the Cochrane risk of bias tool for critical appraisal, for example. The manuscript would benefit much by including examples known by the author where he experienced the problems well described in the paper. The manuscripts reads very theoretically (although personally I know that it’s not) and readers might be interested in some examples from the practice. It might also be interesting for the reader to know that the AMSTAR developers are going to develop an AMSTAR version for non-randomized studies. This information can be found on the AMSTAR website Level of interest: An article whose findings are imp

Thank you for these valuable suggestions. I added some info to explain further the difficulties on the methodological evaluation of primary studies included in the SRs (lines 257-261). I also added information on the new developments of AMSTAR as the reviewer suggested (lines 104-105).