**Author's response to reviews**

**Title:**Establishing confidence in the output of qualitative research synthesis: the ConQual approach

**Authors:**

Zachary Munn (Zachary.Munn@adelaide.edu.au)
Kylie Porritt (kylie.porritt@adelaide.edu.au)
Craig Lockwood (craig.lockwood@adelaide.edu.au)
Edoardo Aromataris (edoardo.aromataris@adelaide.edu.au)
Alan Pearson (alan.pearson@adelaide.edu.au)

**Version:**4**Date:**9 July 2014

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

Response to Reviewer's Comments

Thank you for sending me the revised version of the paper. It has certainly been strengthened and some key points clarified, though a few issues still remain. I have three main points for 'major compulsory revision':

Thank you for your feedback. We are pleased that you believe the paper has been strengthened during the process of peer review. We have addressed all of your comments below.

1. The text seems to slide between assessing the dependability and credibility of individual findings, studies, and synthesised findings. I found myself getting a little confused about this, here and there. For example, you can see how the dependability of a finding from a single study can be assessed using the measurement given in Figure 1, but how does this translate to assessing dependability across studies and findings? (And on page 12 we are told that "Each finding extracted from a research report can therefore be judged with a level of credibility...") This appears to be quite different to the GRADE process, where we are interested in the level of confidence in the synthesised finding.

Thank you for highlighting this issue. We have included some additional information and highlighted when we are discussing individual findings as compared to the synthesised finding. We have now clarified the guidance for assigning the final level of dependability to the synthesised finding and not the individual findings.

Similarly, on Page 13 you say "Downgrading for credibility may occur when not all the findings included in a synthesis are considered unequivocal..."; I was under the impression that we GRADEd each synthesised finding, not the synthesis as a whole (i.e. we might have some findings with high confidence and some with low – in the same synthesis). Table 1 does not help here (but could), as the way that 'dependability' and 'credibility' are calculated is not explained. I think one of the most important amendments lies in explaining how the tool actually works and clarifying the level at which it is applied (study finding / study as a whole / synthesised finding / synthesis as a whole).

Thank you – you are correct and we have changed this from 'synthesis' to synthesised finding' in the paper. We have also included additional information in the comments section in table 1 regarding why a synthesised finding might be downgraded.

2. The meta-aggregative 'heritage' is much clearer now, and the authors use this to good effect in justifying how credibility and dependability might be determined (i.e. the assessment of dependability is rooted in the JBI-QARI tool; and credibility in the meta-aggregative definition of a finding as being a verbatim extract of the authors' interpretation + supporting evidence).

However, no evidence or argument is given in support of the tool's use outside meta-aggregative reviews, and so I do wonder whether the title / abstract needs to be more specifically focused on meta-aggregation?

The work to date has focused in meta-aggregation, and this based on previous feedback has again been emphasised in the manuscript. We have included in the abstract that this was developed for meta-aggregative reviews, but do not wish to limit the approach by creating a linear and overly reductionist title. We further note that in response to previous peer review comments that link with meta-aggregation has been clarified and that reviewer 2 is now satisfied with both the title and abstract.

If not, I think we need more detail on how the tool might be used more widely (and possibly what the "slight modifications" mentioned on page 15 might be in order to achieve this).

The slight modifications mentioned would essentially consist of using a modified system to assess credibility or dependability of findings that are different from those detailed in this article. We do not wish to detract from this article by discussing in detail how this may be done, particularly as we state the following already: 'as a ConQual ranking can be generated with any approach where credibility of findings and dependability of research are assessed.'

3a. The reference to the CerQual was certainly needed, though is slightly out of date, as I think CerQual now contains four dimensions. Maybe it would be worth checking this with Simon Lewin and colleagues, as there may be no published reference to the additional dimensions yet.

Thank you for this comment. The aim of our paper is to publish on the work of our group for wider international interest and not to provide a descriptive comparison with between ConQual and CerQual. As an author group we feel we have addressed the work of the CerQual group in enough detail considering this paper is focused on a different approach. We also do not feel we should be including any additional information regarding a methodology which has not been officially published and still appears to be in an evolving state.

3b. You mention "Glenton's protocol", but I think this is a completed and published review?

Thank you for bringing this review to our attention. We have included this as a reference. However, we have still included the protocol for the review also, as they provide a definition for 'certainty' in this protocol which we have not been able to locate elsewhere.

3c. I think you could be more even-handed in your comparison of tools, as you point out that you're not aware of any testing of CerQual in meta-aggregative reviews, but

do not say whether ConQual has been tested in anything other than meta-aggregative reviews.

We have included the following sentence: 'or of ConQual for other types of qualitative research synthesis approaches.'

3d. In your comparison of the tools, it would be useful to know how plausibility differs from credibility.

We were unable to find a clear definition of the term 'plausibility' or 'coherence' within the published CerQual materials. However, we have included how coherence is measured in the CerQual approach from the published example. We do not feel it is necessary to delve any further into a comparison between the various meanings that have been assigned to particular concepts by the authors of the CerQual approach.

4a. I think the definition of confidence still needs a little more work in the context of qualitative research, as this is still not clear. Confidence = "the belief, or trust, that a person can place in the results of research"; but what are we believing / trusting in? This is highly contested in some areas of qualitative research, so I do think we need to know a little more about the authors' take on this in order to understand the purpose and utility of the tool.

We are of the opinion that we have provided enough information here to support our definition of confidence. As mentioned in the article, we define confidence as 'the belief, or trust, that a person can place in the results of research' and have described how this can be measured (by examining dependability and credibility). This information has satisfied the requirements of the second reviewer, who brought up this issue in the first iteration of this paper.

4b. I think the definition of confidence needs to come much higher in the paper, as until we know what it is, we can't really understand why it might be useful for e.g. policymakers (on page 9).

This is a valid point and we have moved our definition of confidence to sit earlier in the paper.

5. The paper needs a careful proof read, as there are numerous typos. Here are some, but this is not an exhaustive list:
- Page 8: "Delphi-like" (missing hyphen)
- Page 12: "authors' interpretation" (missing apostrophe)
- Page 13: "downgraded based on their" (not "its")
- Page 15: "particularly suited to" (not "suitable to")

Thank you, we have made these changes and have had the paper professionally copy edited.

6. Page 10: We need to know how dependability is defined before knowing how to establish it.

We have included additional information on dependability.

7. Is this text in the wrong place? "Factors that increase confidence in the output of qualitative synthesis are dependability and credibility. The group determined that confidence was established in the output of qualitative synthesised findings when the synthesised findings arose from studies with proven dependability, and where the original findings were credible." This paragraph looked as though a cut and paste had reversed its intended ordering. More critically, we are not told *why* the group has determined that confidence relies upon these two concepts.

This paragraph has been deleted. We feel justification for these concepts as criteria for confidence have been provided in the paper. These concepts were agreed upon following extensive reading and debate.

'Discretionary revisions' are as follows:
8. Page 4: "healthcare issues that are not appropriate to be measured..." do you mean "cannot be measured"?

Text changed.

9. Page 10: If "The concepts of 'dependability' and 'credibility' are analogous with the ideas of 'reliability' and 'internal validity' in quantitative research", why do we need the new terminology?

It is beyond the scope of this paper to discuss these issues in great detail, but references are provided. It is expected that readers of this paper will have some understanding of the historical debate regarding nomenclature for establishing validity in qualitative research.

10. Page 12: I am not sure the diversion into confirmability is needed? It interrupts the flow of your account.

Thank you for highlighting this; however, we feel it is important to discuss and are satisfied with its presentation.

Reviewer 2:
At this time I believe the authors have sufficiently addressed all of the reviewers' concerns. I would suggest another review for grammatical edits

Response:
Thank you. We have done a further review of grammar and made some changes. We have also had the paper professionally copyedited.