

## **Reviewer's report**

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

**Version:** 3 **Date:** 28 June 2014

**Reviewer:** James Thomas

### **Reviewer's report:**

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':

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. 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).

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? 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).

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.

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

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.

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

‘Minor essential revisions’ are as follows:

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.

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).

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”)

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

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.

‘Discretionary revisions’ are as follows:

8. Page 4: “healthcare issues that are not appropriate to be measured...” do you mean “cannot be measured”?

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?

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

**Level of interest:**An article of importance in its field

**Quality of written English:**Needs some language corrections before being published

**Statistical review:**No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

As previously mentioned, I have been involved in a couple of meetings relating to the similar 'CerQual' tool; but do not feel this has affected the review I have given here.