Reviewer’s report

Title: Training and Evaluating Simulation Debriefers in Low-Resource Settings: Lessons Learned from Bihar, India

Version: 0 Date: 26 Oct 2018

Reviewer: Jennifer Mary Weller

Reviewer's report:

Improving opportunities for clinical training in under-resourced environments is important and it's interesting to read how the authors went about evaluating their local debriefers.

With such an extensive video resource, I did wonder if a better question could be posed, or the videos used to advantage in subsequent feedback to the debriefers. It would be interesting to know the extent to which the local debriefers met the objectives of the local training mentor training programme - the reports from the interviewees suggests that the tools used to evaluate them were not well aligned with the training objectives.

There do seem to be quite serious limitations in the quantitative arm of the study.

The title is not an accurate reflection of the study. The research questions do not evaluate if high quality debriefing can be taught, and this doesn't fit the criteria for a feasibility study. Suggest a change in the title to reflect the research questions

The manuscript is well written and organised in a way that is logical and easy to follow.

Line 137 - consider restating the following: "The aim of this study was to explore the feasibility of debrief evaluation in Bihar". The aim as stated doesn't seem well aligned with the three stated research questions.

The context is pertinent and the methods are clearly described, including the modification of the two tools. The extent to which the training of the mentors aligned with the constructs of the two tools could be clarified. Were the tools fit for purpose?

Was there ongoing training for the mentor debriefers? Was there some expectation that they would get feedback and improve?

Line 214 - 217 - this sounds more like rater training than piloting the tools. Please clarify in what way the tools were piloted, or if this was the rater training process. Please clarify if the rater training continue until a certain level of agreement was obtained? This would be best practice in rater training.

221 - please clarify who the 'separate video analysts' were. Were they the two nurses? Please clarify if the raters used both the rating tools after a single viewing of the video or if the videos watched twice. Please describe the order in which the analysts were presented with the videos to code. Was it randomised?

Please describe the sample size calculations for choosing 73 videos from a possible 4,066.

232 - It's not clear to me why unpaired data was used - wouldn't it have been a much stronger study design to use paired data from debriefing mentor pairs? Was the decision to exclude items with less than fair IRR made a priori? It would seem more robust to include all study data.

280 - Table 1. Can you include an explanation of 'Phase' in the legend?

295- Table 2 - was there a mean CAPE score? It seems a bit problematic that the CAPE tool, scored at
the individual item level, is being compared to the DASH tool, which is scored at the domain level. It's perhaps not surprising that more objective questions such as length of the analysis phase had better IRR than those that asked more about quality.

Please report on missing data - were all of the selected debrief of adequate technical quality to analyse? How was missing data on the score sheets treated?

The results relevant to improvement don't seem all that convincing of important improvements over time. There were multiple tests of significance which perhaps suggests adjustment for multiple comparisons is needed. Important things seemed to get worse. The analysis of the interviews is very honest. The description of what went on suggests that the DASH tool was not well aligned with the training programme for the debriefers. Were they familiar with the constructs within the tool? Is that what they had been taught?

Discussion
The most interesting data is coming from the interviews in my opinion. It's hard to know what to make of the ratings, and changes over time.
It seems like a real study limitation that there are no measures of mentee perceptions of the debrief, and no outcome measures of the study.
Conclusion - it does seem that the main point of this study was that it was feasible to establish some form of debriefing in Bihar. I'm not sure that the comparison of the tool tools helped that.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:
1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors’ responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal