Reviewer's report

Title: Comparing the accuracy of brief versus long depression screening instruments which have been validated in low and middle income countries; a systematic review.

Version: 3 Date: 28 September 2012

Reviewer: Charlotte Hanlon

Reviewer's report:

The authors have accommodated my comments. I have a couple of minor comments and some suggestions for the authors as they finalise the manuscript.

Minor essential revisions

Please can you include the quality grading of each included study in Table 1 or 2 to aid interpretation of how much weight to give different scales?

In Table 1, the study by Tesfaye et al. used DSM-IV diagnosis by psychiatrist as the gold standard and not the MINI

Discretionary revisions

The Kenya study by Monahan et al. also looks at the PHQ-2 as a screen. It would be useful to know whether this 2 item scale was as good as the 9 item scale. Similarly it would be worth commenting on the relative accuracy of K6 and K10.

In the discussion it would be helpful if you could comment more on the relative quality of the studies and appropriateness of the 'gold standards' for depression that were used. This will help to guide future researchers.

Similarly it would be helpful to comment on the cultural validity of scales / efforts made to maximise cultural validity.

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests