Author’s response to reviews

Title: Determining an anthropometric surrogate measure for identifying low birth weight babies in Uganda: a hospital-based cross sectional study

Authors:

   Elizabeth L Nabiwemba (enabiwem@musph.ac.ug)
   Christopher G Orach (cgorach@musph.ac.ug)
   Patrick Kolsteren (pkolsteren@itg.be)

Version: 3 Date: 14 March 2013

Author’s response to reviews: see over
Response to reviewer’s comments (Bold and italics)

Reviewer’s report
Title: Determining an anthropometric surrogate measure for identifying low birth weight babies in Uganda: a hospital based cross sectional study

Version: 2 Date: 2 March 2013

Reviewer: Alan Horn

Reviewer’s report:
Many of the revisions recommended in my previous review have been addressed but there are a number of outstanding issues.
Major compulsory revisions

General
Many of the grammatical errors persist.
1. The word cut-off is not uniformly hyphenated. It should either be one word, or hyphenated through-out the text. – cut-off hyphenated in entire text

2. “Context-specific” and “hospital-based” should be hyphenated. - done

3. “Grammes” should be “grams” or “g” – uniformity is needed. – replaced ‘grammes’ and ‘g’ with ‘grams’

Abstract
4. 2nd line: weighing should be “weighed”. - change made

5. 4th - and 2nd - last line first page: The values for LR and sens/spec do not correspond with those in tables 3 and 4. – correction made

Introduction
6. 1st line: “.contributes 60 – 40% of infant and under 5 mortality..” – “of” should be “to”.- corrected

7. 4th line: “into” should be “in”.- corrected

8. 11th line from end of introduction: “measureis” should be “measure is”. - corrected

Results
Under descriptive characteristics:

10. The % agreement between two observers is quoted as a range for “all” the measurements: Were the two observers the supervisor and the midwife? That should be stated. It is preferable to give the agreement rate for each of the 6
different measurement types and the number of paired values compared. Kappa should be given rather than just % agreement. – **kappa statistic indicated for each measure i.e weight, HC, FL, MUAC, TC and CC**

11. 2nd last line page 8: should describe the “values” as “Pearson correlation coefficients”. - **corrected**

12. 3rd line page 9: “Foot length was the best predictor” – this statement should be in the discussion and can only be made if statistical comparison was made. You can infer significance at a basic level by referring to the degree of overlap between confidence intervals. – **the statement has been re-phrased in light of the comments in no.18 and moved to discussion.**

13. The last paragraph of the results: The line, Likelihood ratios show that all measure are good at screening for LBW..” should be in the discussion. - **moved to discussion**

14. The LR and odds ration values do not correspond with the values in table 4. – **corrected**

15. Table 1: As p-values have now been added in full, an asterix footnote is not required. – **footnote removed**

**Discussion**

16. Line 9: “…from the DOR only 58% of the variation can be explained…”. DOR should be r2. - **corrected**

17. Last line page 9: Foot length is described as the “best predictor” but no statistical comparison with the other measurements is discussed. **The term ‘best predictor’ has been dropped. Statistical comparison has been included in results and instead FL and CC are referred to as ‘most appropriate measures’**.

18. The authors state that false negatives should be minimized with this tool – in that case, a high sensitivity, a low negative LR (ie FN/TN), and a high AUC are most important.

**Comparison of means and confidence intervals suggests:**

i) The AUC of FL is significantly higher than HC and TL, but the CIs of the other measurements overlap. (Table 2)

ii) There is no significant difference between sensitivity or specificity of FL, MUAC, TC and CC, while the values for HC are significantly lower than at least some of the other measurements. (Table 3)

iii) CC has a significantly higher +LR than any other measurement. Although FL had the lowest –LR, there is no significant difference between –LR of CC and FL. (Table 4)

The above data suggests that the two most appropriate measurements are FL and CC. With the concerns regarding difficulty of measurement of CC, it is
appropriate to recommend the use of FL as the most appropriate. The above factors should all be discussed under results/discussion as appropriate. – the points above have been addressed. Statistical comparisons are stated in results and are discussed to reach a logical conclusion.

19. First line page 10: The sens/spec values do not correlate with table 4. - corrected

Limitations
20. The potential for bias by having a single investigator performing both index and gold standard tests should be stated, Further potential for bias exists when a reference measurement is derived rather than tested prospectively, this should also be mentioned. – limitations mentioned

Conclusions
21. The wording of the first line could be revised in light of above comments. The study suggests that FL is the most appropriate predictor for LBW due to a combination of high predictive values and ease of measurement. – revision done

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

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: 
I declare that I have no competing interests