Reviewer’s report


Version: 0 Date: 30 Aug 2019

Reviewer: Sanja Stanojevic

Reviewer's report:

Madanhire and colleagues summarize a large study of healthy children from Zimbabwe and highlight how the GLI reference equations can be appropriate used in this population. The overall study design and methods are strong. The overall presentation and interpretation of results could be improved to highlight how these results impact the interpretation of pulmonary function in practice. In addition there are several statements that need to be corrected.

Specific comments.

1. The GLI are not ERS equations per se, it would be more appropriate to just use GLI2012 when describing

2. In some cases the text says urban in other peri-urban please specify and use consistently

3. In the abstract - please specify that z-scores using the African-American GLI equations are summarized in the results

4. Table 1 - it would be better to have the total column first. Which of the GLI equations were used to calculate the z-scores. Please specify
   Table 2 - instead of 'standard normal GLI' it would be more appropriate to say 'Caucasian GLI'

5. Table 3 - I am not sure this table is very useful without accompanying confidence intervals (adding these correlation coefficients to the figures might be better.

6. Figure 2/3 - adding the +1.645 line would be helpful to understand the upper and lower bounds

7. In the methods, the statement that z-scores are calculated as observed-expected/standard deviation is not incorrect, but with the GLI equation the calculation of z-scores is a bit more complicated. It might be simpler to state that GLI z-scores and LLN were calculated using the available calculators, which provide a age, height, sex and ethnic specific value.
8. In the methods it states that the ECCS/Polgar equations were used, but the ECCS are only available for those >20 years, so I assume that only the Polgar equations are compared. Please include a citation for these equations. The statement in the discussion regarding the Polgar reference is not correct. The ECCS were developed in Coal workers, but the paediatric equations by Polgar were not, there equations combined multiple published equations and have their own set of limitations.

Results:

9. The gender differences in z-scores require further investigation. The calculation of z-scores should adjust for the sex related differences in lung development, therefore the observed differences after adjustment suggest that the two groups are different due to other factors.

10. Since z-scores were calculated for all of the GLI equations, it is important to highlight which results are described in the text. I assume after the comparisons are done in Table 2, that the rest of the results are for African-American equations but this is not clear.

11. Reporting the % of subjects with values below the LLN would be more intuitive, as we would expect 5% to be below this value. It does look like for FEV, a greater proportion than 5% have values below the LLN which will influence interpretation.

12. The comparison of GLI to Polgar should include a Bland-Altman plot.

13. A multivariable regression to adjust for ses, school etc. might be more informative than the supplemental material. In the current version it is not clear whether univariate or multivariable regression results are presented. Is there evidence that some of the offset observed is explained by SES? A multivariable model might help explain this.

14. In supplemental file 2 there seems to be a bias with z-scores and age in the high income group. Additional data, including the average z-score(SD) and % below the LLN in each of these sub-groups (in a Table) would be informative. For direct comparison, the scales (y-axis in particular) should be the same for each figure. In the high income group, there is also a bias with zheight and FEVFVC.

Discussion

15. The dysnaptic growth differences during puberty are unlikely to have occurred in children between the ages of 7 and 13.. possible if puberty was shifted to younger ages, but then there would have been a clear age bias observed (which does not seem to be the case - Figure 2).

16. Avoid terms like diagnosis, since a diagnosis requires more than PFT results. Misclassificaiton would be more appropriate.
17. Discussion, spirometry on its own cannot provide evidence of restrictive lung patterns, please consider re-wording

18. Discussion of the lower on average z-scores and what this means for use in clinical practice is important and currently not addressed in the manuscript. While overall the differences observed may simply be due to sampling bias, there is a chance that the LLN may over-classify children with low lung function. A comment about particular attention to interpretation of results around the LLN is needed to avoid over-interpretation of results in practice.

19. The conclusions generalizes the results to all children in Zimbabwe. The authors may want to add a disclaimer that this assumes that the schools in this study are representative of all children.

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

Yes

**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 am able to assess the statistics

**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 am the co-chair of the global lung function initiative network. I have had previous discussions with the authors about this work but have not seen the manuscript.

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