Reviewer's report

Title: Understanding the complex determinants of linear growth and adiposity in disadvantaged day-care preschoolers in Salvador, NE Brazil through structural equation modelling

Version: Date: 16 September 2014

Reviewer: aryeh Mr. stein

Reviewer's report:

This paper presents the results of an attempt to 'understand' the 'complex determinants' of growth and adiposity of preschool children in a community in Brazil. The authors recruited children from several day-care facilities and obtained a series of measures on them and their mothers. They then used SEM (implemented as a path analysis) as a modeling tool.

The population is a strange one to try and understand issues of 'double burden'. The sample does not appear to have a 'problem' with either stunting or overweight. In a well-nourished population, one would expect 2.5% of the sample to be outside the threshold of 2 SD (upper threshold for obesity, lower threshold for stunting, and 16% to be outside for overweight and for moderate stunting. The observed proportions are within those values, suggesting an adequately-nourished sample. This can also be seen from the mean HAZ (which is positive)) and mean BMIZ (which is negative). Thus, it is not clear what there might be to explain that would have any public health relevance.

Other than attending the daycare centers, were there any other inclusion or exclusion criteria? Among those attending the daycare centers, how many parents agreed to have their child participate and how many declined? In this community, how do children in daycare centers differ from those not in daycare centers, or from those in non-philanthropic daycare centers?

When were the data collected?

Why were the 100 excluded? Did the authors consider using multiple imputation strategies to guard against any biases that might have resulted from excluding this many observations?

The title speaks about 'linear growth' – we do not know anything about 'growth' (which is a change in size over time) but just 'stature' or height. Understanding 'growth' requires longitudinal data. Similar considerations apply throughout the manuscript.

The authors make the point that there are children (and mothers) who are both short and overweight. True, there are, but the proportion of short children and mothers who are overweight does not differ from the proportion of non-short children and mothers who are overweight. Again, what is there to 'explain'?

The sample was 3-6 years of age when studied. As the authors are no doubt aware, most growth failure, if it is going to occur, happens in the first two years.
Hence, metabolic and other parameters that are measured after stunting has occurred cannot be predictors of stunting, and may in fact be consequences. If the model is mis-specified, even if the statistical relationships are present, it cannot provide useful tools to guide policy.

It is not clear that there were enough ‘white’ children (n=22) to provide useful inference. It might be advisable to exclude these children and re-run the analysis. Furthermore, given that SES and skin color showed ‘no differences’ (whatever that means – if indeed that is what the author meant by their sentence on page 7 para 1 -, it is not possible from the data in Table 1 to determine, but one might infer from knowledge of Brazil that high SES children were white), the colinearity is likely to be a problem in a small study.

The prevalence of helminth infection was 17.8%. Even if the estimate are unbiased and all helminth infection were eliminated, given the weak coefficient (0.08) how much improvement would there be in BMIZ and HAZ?

Which variables in the model highlight the importance of clean drinking water, regular physical activity, and maternal nutrition education? Where is the evidence that any of these are deficient in this community or that they are associated with nutritional status? Certainly not in the tables provided.

All the biomarker data were measured simultaneously in blood. What is the logic behind the causal direction of the arrows? Are these motivated by an understanding of their uptake and circulation?

Was information on child diet collected? If so, why was that not considered?

Why do the authors think there is a strong causal relationship between gender and helminth infection? Given their coding, is it boys or girls who have higher prevalence? Similarly for selenium – was this arrow pre-specified or added based on the correlations in the data? If the latter, what other arrows were post-hoc? Finally, given that both HAZ and BMIZ are generated from sex-specific distributions of height for age and BMI for age, why did the authors think that a main sex effect on HAZ and BMIZ was likely?

What is the evidence that speaks to the hypothesis of this paper, that ‘use of SEM would broaden our understanding…’? Were alternative methods used and shown to have narrower understandings? How could one test this question? Perhaps the study question might be rephrased. Given the centrality f the UNICEF framework to our understanding of child nutrition and growth, what additional understanding has been provided by this analysis? It has, at most, confirmed the relevance of the UNICEF framework once again.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**
I declare that I have no competing interests