Reviewer’s report

Title: Prevalence and risk factors for stunting and severe stunting among under-fives in North Maluku province of Indonesia

Version: 1 Date: 21 April 2009

Reviewer: Thorkild Tylleskar

Reviewer’s report:

Review of Agho et al. Prevalence and risk factors for stunting in North Maluku of Indonesia

This manuscript is based on a apparently well conducted survey in one part of Indonesia documenting a high degree of stunting. The technique for such surveys has become more an more standardised and the authors mainly follows these inofficial standards. In this respect it is a paper that well merits publication.

But before publication there are a number of issues to address:

Major issues:

On page 4 it is stated that this paper uses the new growth standard from WHO and in the stat section on page 7 it is stated they are using the NCHS reference. Which one has been used in this paper? My next comment is that they should be using the new growth reference from WHO if they want to publish the paper.

The authors have constructed and household wealth index and made 3 groups. That is fine. I usually call the groups: poorest, middle and least poor; as ‘rich’ is not often a correct attribute even if they top their community. But which items did they include in the index and how was it done (Principal Component Analysis or something else?) I would also like to see a table showing the percent of people in the 3 groups owning the different items included in the index:

Poorest Middle Least poor
Proportion of households with: (%) (%) (%)
- Mobile phone 6 17 60
- Car 1 3 27
- Fridge 12 30 90
ETC.

This is a way to document that the procedure worked. Sometimes an automated procedure can do funny things and you need to verify that it did ok.

Please try to organise the factors in the same order throughout the manuscript: in the methods, results, and in the tables. And try to apply some logic in it. For instance in the tables it lists: father’s education, mother’s education, father’s
occupation, mother’s occupation, parental education, parental occupation. Why not have all the education together and all the occupation together? Why not start with Region, district and then move to background household factors like: education, occupation, socio-economic status (wealth index) and hereunder mention which factors were asked about and which ones were used, then to household size, no of meals, then child age and gender. ‘Nutrition information during pregnancy’ does it fall under education or is it inherent to the child as you imply by calling it child level factor? To me, it is rather education.

In the tables you do not need to report p-values when you have given us the confidence intervals, this is redundant information that clogs your tables: p-values should go out. Also put the 95% CI in brackets (like in table 1) and with a dash between:

OR (95% CI)
0.84 (0.59 – 1.21)
0.76 (0.55 – 1.05)

Because we know the age and the gender plays a large role for the stunting, I think you should force them into the multivariate model or possibly report the R2 for both models with and without age and gender.

In the tables, when you present education: give the number of completed school years. For someone not familiar with the local situation, it is not clear how many years ‘elementary school’ represents in Indonesia. And what is the cut-off for ‘high education’ in your analysis?

In the tables, you need to indicate how many children went into the different groups, i.e. on each line there should be a number indicating the no of children had this condition. How many children had a father that completed elementary school? Etc. This requires a new column before all the estimates. This is also to make sure that you have used the same number of children in both the univariate and the multivariate analysis. A common situation is that some information is missing in some of the questionnaires so when you are running univariate analysis, the number of children is not always the same for univariate and multivariate. In such a situation you need to drop the incomplete questionnaires and only use the complete data sets in both, otherwise the difference between the two columns may in part depend on the fact that the data sets for the two analyses differ.

Figure 2 is not informative enough to merit its existence in a publication. Instead I want a figure of the proportion of stunting and severe stunting by age. This could be a line diagram with age on the x-axis, percent on the y-axis and two lines binding the stunting and severe stunting in each year group.

I see the justification of having a model for 0-5 year-olds and for 0-2 year olds but why complicating your message with a model 24-59 months? If this paper was produced by one of my students I would have requested him/her to take out this model to give more overview of the findings.
Similarly, it becomes very congested when you have multivariate models for both stunting and severe stunting with different factors coming out in the different models. The biggest difference is in the models for 24-59 months of age. Taking this one out would make the presentation much nicer.

Minor issues:
Page 5 you indicate the population in a very exact way: 784 395 inhabitants. In fact this was probably true only for about half an hour or so, the next half an hour it had increased. So when you report this type of numbers use rounded numbers 785 000 inhabitants AND indicate the year: in 2004. This goes for all the population figures.

Reference no 16 is a thesis, by definition difficult to access. Is it available online?

**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:** Yes, but I do not feel adequately qualified to assess the statistics.

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