Reviewer's report

Title: The development of socioeconomic health differences in childhood: results of the Dutch longitudinal PIAMA birth cohort.

Version: 1 Date: 26 August 2010

Reviewer: Henrik Ohlsson

Reviewer's report:

This paper focuses on the development of socioeconomic health differences in childhood health. Moreover, the authors investigate possible explanations for these socioeconomic differences in childhood health. The authors employ an interesting dataset from the Netherlands with 3,963 children with individual measurements during each year from 3 months up to 8 years of age. The research question is important and well posed. However, there are some problems with the statistical and epidemiological analysis performed, that needs to be solved before any decision can be made.

Major compulsory revisions

In their second aim (possible explanations for these socioeconomic differences in childhood health) the authors use an effect decomposition approach to explore their research question. While this approach may (?) be valid for continuous outcomes, it often requires unjustifiable assumptions (see A further critique of the analytic strategy of adjusting for covariates to identify biologic mediation – Kaufman, J, MacLehose, R, Kaufman, S. Epidemiologic perspectives and innovations Oct 2004). At the least, a discussion of this issue is warranted.

The use of this effect decomposition approach may be even more problematic when several determinants are included in the model. The authors conclude, for example, that the relation between educational level of the mother and children’s poor general health is almost completely (75 %) explained by the biological and lifestyle factors. I have some hesitations regarding this approach – when several mediation hypotheses are each tested with a simple mediator model, these separate models may suffer from the omitted variable problem, which can lead to biased parameter estimates. Moreover, this approach can also lead to the conclusion, if you use enough mediators, that you explain more than 100 % of the studied relation. For example on page 9 it is stated that “about one third of
the higher prevalence of weight problems among children from lower educated mothers can be related to the biological and lifestyle determinants studied. How about the correlation of the biological and lifestyle determinants?

When numerous mediators are used in a multiple mediation model, the likelihood of parameter bias due to omitted variables is reduced. Furthermore, it is possible to determine to what extent specific variables mediate the relation between maternal education and the outcome, conditional on the presence of other variables in the model. It is important to remember that a specific indirect effect through a mediator (say, breastfeeding) in the multiple mediation context is not the same as the indirect effect through breastfeeding alone, except in the unlikely circumstance that all other mediators are uncorrelated with breastfeeding.

Moreover, table 4 includes very little information regarding the different variables included in the model. The authors must address these issues in a revised manuscript.

Additionally, there is an ongoing discussion in the literature on how to calculate the amount of the mediation effects and uncertainty measures (e.g. use of bootstrapping methods). At the least, a discussion of this issue is necessary. The authors have used 15% as cut-off point to determine if the determinants are explanatory or not. Why did the authors use 15%? Why did not the authors include an uncertainty measure of the mediation effect?

Page 8, line 18. Lower educated mothers are less likely to breastfeed their children for more than 16 weeks… I can not find the results from this analysis in the tables. In the results part there are several results, regarding associations between the biological/lifestyle determinants and education and or the health outcome I can not find in the tables. It is difficult for the reader to judge the importance of the results if they are not shown.

Maybe the authors should consider the use of directed acyclic graphs in order to clarify for the reader the different pathways in their data material.

Page 8, line 2 – regarding the phrasing statistically significant: First of all the authors have not defined what they mean with statistically significant. Secondly, as pointed out in several articles the phrase is obsolete and is often misinterpreted. Please rephrase and focus on the implications of the range of values in the 95% confidence interval. This problem is also apparent at page 9, line 10 – However the increased risk for frequent respiratory infections among children with intermediately educated mothers is no longer statistically significant raised. The point estimate changed from 1.17 to 1.12 – the question is not if it is statistically significant but rather if the reduction of the point estimate (and the interval) is relevant or not.

The tables are hastily done. The headings in the table section need much more information. In table 3 a and b are not explained. In table 4 a and * are not explained. No confidence intervals are included in table 4. Moreover, I would like information of the share of the different health outcomes for the different educational levels.
The figure legend and figure information are also hastily done. What does OR mean (I know it is odds ratio – however I believe that this information must be included). The error bars round the point estimate - is it a 95 % confidence interval or?

The abstract does not really represent the manuscript – there are no results presented regarding the possible explanatory determinants.

Please include and discuss in the discussion part that almost 35 % of the information regarding obesity and overweight was missing.

I also have considerations regarding the policy implications the authors propose. I think it is too simplistic to state that for example – “our findings emphasize that health professionals should promote breastfeeding, especially among parents with low socioeconomic backgrounds”. First of all the authors have not reported the association between breastfeeding and their health outcomes. Secondly, the authors have definitely not discussed the variation within the group of mothers with low socioeconomic background. Therefore, in order to avoid trivializing a complex process and to increase the likelihood that policy discussions are treated with the seriousness that they deserve I believe that this policy implication might be excluded from the manuscript.

Page 5, line 29 – The biological and lifestyle determinants used in the explanatory analyses were selected on theoretical grounds. Please provide references.

Minor Essential Revisions

There are some small language mistakes in the manuscript that disturbs the reader – maybe it is a good idea to have the manuscript proofread.

For example:

The second heading on page 8 – Please delete Text for this subsection

Page 8, line 15 – change are almost completely explained… to is almost……

Page 5, line 8, …weight en height – change en to and

Page 5, line 20……next tot status – change tot to to

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, and I have assessed the statistics in my report.
Declaration of competing interests:
I have no competing interests.