Reviewer's report

Title: Selection bias: neighbourhood controls and controls selected from those presenting to a health unit in a case control study of efficacy of BCG revaccination

Version: 1 Date: 28 November 2006

Reviewer: Philip Hill

Reviewer's report:

General
This study is potentially a useful addition to the literature on TB case control studies in developing countries. The group has a good track record in TB research, providing key articles.

There are certain problems with selecting controls when the population is not clearly defined and the drivers for health seeking behaviour are complex and poorly understood. What is perhaps not emphasized enough in this paper is that the selection of neighbourhood controls and health facility controls are both vulnerable to bias. Neighbourhood controls may include individuals who would not have been treated at the same facility if they got the disease and being treated at the facility can be related to an exposure under study eg when a health facility focuses only on particular types of people. It could be argued that using neighbourhood controls will relatively underestimate BCG vaccination as it may select individuals who are less likely to turn up for BCG vaccination and less likely to turn up to a health facility when coughing. Hospital based controls are a select group of individuals and bias is introduced in particular when the disorder for which they present to the facility has a link with an exposure under study.

With respect to BCG vaccination. Since it is done at health facilities, one would be concerned that individuals that are more likely to attend the facility to receive their vaccination may also be more likely to attend if they have symptoms of any disease. So facility controls may be over-represented with respect to BCG scar and this does seem the case in this study. However TB cases may also be over-represented with respect to BCG scar.

The published case control study that was referred to enrolled neighbourhood controls and not health unit controls. Therefore the authors at least appear to have got this mixed up in the introduction/methods. This does beg the question as to what the authors were trying to do with this study. Playing devil’s advocate for a second: it appears that what actually may have happened was that they initially did a case control study using health unit controls and didn’t publish those results as they were so different from the RCT results. Perhaps they then proceeded to do a study with neighbourhood controls to see if they would be more like the RCT results and published that study instead. It would be important to hear from the investigators what actually happened. Why did they not present their initial case control results with a good discussion of why they might be biased, or at least wait till they had results from both sets of controls before publishing them?

The published case control study found a trend towards lower than expected efficacy in the youngest age group, while the overall efficacy was similar to the RCT. It is not surprising that neighbourhood controls are also subject to some form of bias. Considering the issues with selection, mentioned above, it is not surprising that there might be under-estimation of the efficacy of BCG, in certain individuals.

--------------------------------------------------------------------------------------------------

Major and minor Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Intro:

I think the intro is too long and much of this could wait until the discussion for comparison with the present results.

It is of note that the confidence intervals around the efficacy estimates using the two groups of controls actually overlap with each other-this is not pointed out by the authors and should be mentioned I would have thought. One cannot say for certain that the efficacy estimates are actually different.

Methods:
It would be helpful to have more detail on the local health system—what are health units in relation to other providers in the community and the hospital—who provides TB treatment—ie. what are the providers of treatment?

Why were the controls an average of 1 year older than the cases when the control selection was actually 2 years later?

Results:

In this study the neighbourhood controls were more commonly males—is this related to refusals? It is not clear in this study how many potential controls of either type refused participation and whether their characteristics were different to those that accepted. It is clear that some controls were excluded after agreeing to participate, because their cases were excluded, but not how many potential controls refused and what were their basic characteristics?

Rather than using the neighbourhood controls as the reference comparison, there should probably be more detailed comparison of both control groups with the RCT findings as they are likely to be subject to the least bias.

Tables and figures:

P values should be either the actual p value or whatever the journal requires when <. Eg. <0.0001

Table 2 is really too small an amount of information to warrant being a table.

Table 3. When a variable is binary, just one component could be mentioned in the table eg. Proportion male n(%), rather than give both results, as one can be easily deduced from the other.

Discussion

One would expect a paper such as this to have a more full discussion of the design issues if it is really to help us understand the issues around case control studies in TB in developing countries better.

There should probably be some discussion about the usefulness of the socio-economic surrogates mentioned—ie. there is no guarantee that anyone is getting this right in developing country research.

Discretionary Revisions (which the author can choose to ignore)

One does also wonder whether this study exposes any issues over the RCT design—are the results of the RCT, having strict entry criteria etc, actually representative of the real situation in the setting? Should there be some discussion about this issue?

There is no mention of the issue of scar formation as many individuals were excluded with no scar while perhaps having a record of having been given vaccine.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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
Declaration of competing interests:

I declare that I have no competing interests