Reviewer's report

Title: Rates of coverage and determinants of complete vaccination of children in rural areas of Burkina Faso (1998 - 2003)

Version: 1 Date: 12 June 2009

Reviewer: Marek Brabec

Reviewer's report:

This is paper of substantial practical value. Technical (statistical) level is good, statistical methods that were used are appropriate. Some details (including terminology) should be improved before publication. Suggested changes and problematic spots are listed below.

It is not immediately clear whether the 44 (in 1998) and 48 health districts (in 2003 study) overlap geographically – and if they do overlap in their coverage, then to what extent. Were they essentially the same in both years (plus 3 additional new ones in the second study)? Or, were they almost completely different? Or something between? More detailed description is needed.

Description of the variables under investigation is incomplete. There is only a reference to a previous paper (which is in French, moreover). Current paper should be self-contained. Hence it should have a simple list of the variables that were considered (without complicating constructions like … those described in earlier paper with the exception of …). This pertains to the page 12 and also page 13.

Common error of using nonsensical “independent variables” terminology occurred on page 12 (and 13). Use “explanatory variables”, instead.

The analyses denoted as bivariate (page 13 and Table 1 headings) are probably univariate (according to a standard statistical methodology). One can guess that all what the authors do is that they perform a set of univariate analyses (each with one response and one explanatory variable) and not bivariate analyses (with two response variables taken simultaneously). If that is so, then the terminology should be corrected.

Similar thing occurs for multivariate analyses (Table 2 headings). What is meant there is most likely the logistic regression with several explanatory variables (analog of a multiple linear regression) and not multivariate logistic regression with several responses being modeled simultaneously.

Since the models in Table 2 for 1998 and 2003 are different (the 2003 model is larger), it is hard to compare their results. It is quite clear that one would expect that the larger one should have better fit (and might have different coefficients even on the same variable – e.g. due to collinearity effects). It is then vital to use the model of the same size (with the same explanatory variables) for both years. If there is no info about a subset of explanatory variables for 1998, then do not use it and compare the years using the smaller set of variables. Otherwise, it is
not clear whether the change in results is due to the real change in vaccination, or due to the change in the model, or due to the changes in both.