Reviewer's report

Title: Differential misclassification of confounders in comparative evaluation of hospital care quality: caesarean sections in Italy.

Version: 3
Date: 2 June 2014
Reviewer: Timothy L Lash

Reviewer's report:

The authors demonstrate the potential importance of differential misclassification of a confounder (high fetal head) in Italian surveillance data on the proportion of births by caesarean section within different hospitals and regions. Overall, the study is well conducted and well written.

Major compulsory revisions

1. The number of births at different hospitals is likely to be substantially different, with some hospitals having relatively few births. Typically, the lowest and highest proportions (of both HFH and CS) will be recorded at low volume hospitals. Semi-empirical Bayes shrinkage methods should be considered to avoid the potential for extreme proportions, measured with poor variance, to unduly influence the results.

2. At least in the discussion section, if not implemented in the analysis, the authors should consider the potential for bias analysis to analytically address the problem of differential misclassification of HFH. The only apparently proposed solution is to avoid adjustment for HFH, but would it be possible to reduce the bias by using bias analysis to adjust for the misclassification?

3. It would also be useful to suggest, if not implement, a validation substudy to examine the actual rates of misclassification of HFH. The widely variable proportion, in correlation with CS proportion and the historical information on demands to justify CS, make a compelling case. And the authors allude to the potential for HFH to be difficult to assess, so maybe validation is not possible. Nonetheless, it is generally good practice in a situation like this one to use validation substudies to document the classification errors, and that good practice should be mentioned even if not implemented in this study.

Minor essential revisions

4. In general and throughout, "misclassifications" (plural) should be "misclassification" (singular).

5. In general and throughout, data are presented to too many digits beyond the decimal point. The sample size is not likely to be sufficient to support the reported level of accuracy.

Level of interest: An article of importance in its field
Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I am co-author of two textbooks that advocate for the use of quantitative bias analysis in analyses of epidemiologic data and for which I receive royalties.