Reviewer's report

Title: Classifying perinatal mortality using verbal autopsy: Is there a role for non-physicians?

Version: 2 Date: 15 April 2011

Reviewer: Edward Fottrell

Reviewer's report:

This well-written paper by Engmann and colleagues details the comparability of underlying causes of death for perinatal deaths derived from review of verbal autopsy data by a duo of physicians (based on consensus diagnoses) and individual auxiliary health workers (“community coordinators”) from diverse low- and middle-income settings. Physician consensus diagnoses are designated the reference standard against which the ability of community coordinators to assign underlying causes of perinatal death from VA data are assessed. Comparability is assessed in terms of case-by-case agreement of diagnosis and overall measures of robustness, which are well described in the manuscript.

Context, rationale and background for the study are sufficiently described, with reference to relevant and up-to-date literature, and the results are discussed in relation to findings from similar studies. There are a number of details and issues that should be considered by the authors before publication of this manuscript.

Minor Essential Revisions

1. The authors define physician consensus diagnoses as the reference standard and acknowledge that this is a limitation of the current study since the “true” cause of death is in fact unknown. It would be interesting to know how often physician consensus was reached immediately and how often there were discrepancies between physicians that required discussion in order to reach a consensus. If the authors were to compare the community coordinator diagnoses separately against each of the physicians, and indeed compare the physicians against each other, one would better be able to assess whether agreement between physicians before discussion was substantially better than agreement between community coordinators, agreement between community coordinators and single physicians, and agreement between community coordinators and physician consensus.

2. In relation to point 1, the methods on Page 6 should clarify that physician CONSENSUS causes COD responses were viewed as the reference standard.

3. The Methods explain that the community coordinators administered the VA questionnaire in the community (page 5). The act of conducting the interview may have provided the community coordinators with more explicit and implicit information from the respondent, environment etc. about the circumstances, signs and symptoms of the deceased before death than was recorded on the
standardised VA questionnaire. A wide range of factors that were probably unknown to physician reviewers may have influenced the community coordinators’ ability to identify a cause. Some discussion of this would be appropriate in the Discussion section, perhaps in relation to the point about unknown validity.

4. All data are pooled, analysed and presented together in the current study. It would be interesting to know whether there was any difference in community coordinators’ performance between the different study settings. Perhaps the role of non-physicians in classifying perinatal causes of death is context specific, influenced by regional differences in literacy, recognition of symptoms etc. and thus differs between settings. In addition to the pooled results already presented, results presented by country might be interesting and some comment on this would be appropriate.

5. Robustness criteria used in the study are based on previous work by Setel et al and provide a useful measure of agreement derived from previous work on VA. However, alternative measures of agreement without the implicit designation of any one method as being “true” do exist e.g. Kappa. Some explanation as to why the authors chose the robustness measures described in the manuscript would be helpful.

6. In relation to point 5, the authors describe VA as an epidemiological tool (page 3), implying a public health utility of the results. However, assessment and discussion of relative differences in CSMFs is perhaps not the most useful measure of the epidemiological/public health utility of VA. The extent to which large relative differences matter in public health terms depends on the magnitude of the true value. For causes that account for a fairly small proportion of all causes (e.g. tetanus in Table 2, APH, maternal accident, malpresentation in Table 4), even a large relative difference can be somewhat meaningless in absolute terms. Conversely, absolute differences can be large and of public health importance if relative differences are small and the true CSMF is large. Therefore, presentation and discussion of absolute differences would be appropriate and may enable comparisons of community coordinator and PR results in terms of public health utility of the data.

7. Page 11 – the comment about non-physician coders perhaps not being a “reliable” alternative to physician coders seems inappropriate since reliability is not actually assessed in the current study.

8. Recommendations on the use or not of non-physicians to determine perinatal causes of death is difficult based on the current study since we do not know the true cause of death – discrepancies between community coordinators and physician consensus might be because the community coordinators identified the correct cause. Therefore conclusions and recommendations need to be carefully considered and limited to the evidence of the current and previous published studies, focussing on agreement rather than reliability or validity. The authors have generally done this but I am not convinced that, in the absence of “true” cause-of-death gold-standards, the call for further research is justified. Can the
authors explain the nature of this further research and how it will bring us further than where their work has taken us?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

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