Author’s response to reviews

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

Authors:

Cyril Engmann (cengmann@med.unc.edu)
John Ditekemena (josol70@yahoo.fr)
Imtiaz Jehan (imtiaz.jehan@aku.edu)
Ana Garces (anagarces@imsalud.org)
Mutinta Phiri (drtinta@yahoo.com)
Vanessa Thorsten (vthorsten@rti.org)
Manolo Mazariegos (mmazariegos@incap.int)
Elwyn Chomba (echomba@zamnet.zm)
Omran Pasha (omranapasha@aku.edu)
Antoinette Teshefu (antotshe@yahoo.com)
Elizabeth McClure (mcclure@rti.org)
Dennis Wallace (dwallace@rti.org)
Robert Goldenberg (rgoldenb@drexelmed.edu)
Waldemar Carlo (Wcarlo@PEDS.UAB.EDU)
Linda Wright (wrightl@exchange.nih.gov)
Carl Bose (cbose@med.unc.edu)

Version: 4 Date: 3 June 2011

Author’s response to reviews:

Dear Editors,

Please find enclosed a review of our manuscript with text that readdresses 3 of the reviewers comments highlighted in green. Please also find a point-by-point response in italics.

Reviewer1, comment 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.

We have included in the manuscript a paragraph which reads” Since the most commonly used method for coding COD uses physician panels and these panels represent the reference standard, other measures such as kappa, which are used to compare two alternatives considered of equal weight, and not to compare an alternative approach against a reference/gold standard, would be
less appropriate in this case.
Regarding robustness measures, we have included a sentence that states “We defined robustness using criteria previously described, utilized and published by Setel et al.”

Reviewer 1, comment 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.

The reviewer identifies an important consideration in the interpretation of relative differences. We did not include absolute differences in the tables because the current tables are relatively dense and we felt that the volume of data was becoming somewhat overwhelming. Furthermore, because the actual CSMF(PC) and CSMF(CC) are located in close proximity within the tables, the absolute differences are readily available to the reader.
We have added a statement which reads “It is worth pointing out that where relative difference values are small, it may be because there are very small CSMF’s”.

Reviewer 2: comment 16,
16. RD calculations: Is it possible for you to calculate 95%CI for the RD? As a conservative estimate you could simply base it on the RD between the upper 95% CI for CSMF(PC) and the lower 95% CI of CSMF(CC) as the lower limit, and the lower 95% CI for CSMF(PC) and the upper 95% CI of CSMF(CC) as the upper limit.

We note that this comment by the reviewer was a discretionary revision. First, because these relative differences are a measure of discrepancy for our particular study, not the estimate of some larger population parameter, providing confidence intervals does not change the central message of the manuscript and may in fact be misleading in interpreting these relative differences. Furthermore,
the mathematical nature of the absolute value function that is embedded in the
calculation of these relative differences substantially complicates the computation
of the confidence intervals for the RD, as standard large-sample methods for
approximating confidence intervals are mathematically inappropriate for RD
confidence intervals. As such we would need to develop programs for computing
these confidence intervals using bootstrap simulation techniques, substantially
delaying our response to this request. Given the questionable value of these
interval estimates, we don’t believe that such a delay is warranted.