Reviewer's report

Title: Adaptation of a probabilistic method (InterVA) of verbal autopsy to improve the interpretation of cause of stillbirth and neonatal death in Malawi, Nepal and Zimbabwe

Version: 1 Date: 5 February 2011

Reviewer: Abraham Flaxman

Reviewer's report:

- Major Compulsory Revisions

1. The conclusion of this work must not over-reach. While the authors state that their modified InterVA method provides reliable information, they have demonstrated only the (moderate) kappa scores comparing their technique to PCVA, which itself has limitations. In Edmond et al (reference [22] in this paper), a study comparing PCVA to hospital reference diagnostics found sensitivities of 64-74%. Thus it is possible that InterVA is doing much better than a kappa of .76 suggests, but it is also possible that it is doing much worse.

2. An important shortcoming of the data must be acknowledged. The 734 verbal autopsies used as the basis of this analysis come from previous studies in three sites: Malawi, Zimbabwe, and Nepal. While the data in Zimbabwe were gathered as part of a descriptive study conducted by the Ministry of Health, the data in Malawi and Nepal were gathered as part of randomized trials of community health interventions. In addition to the "Hawthorne Effect", where people under study act differently just because they are under study, the experiment in Malawi hypothesizes that in the experimental arms the interventions will lead to reductions in infant morbidity, increases in recognition of symptoms, and change in care-taker practices, all of which could lead to atypical responses during verbal autopsies. This being the case, it is especially concerning that the kappa score for Zimbabwe is considerably lower than for the other sites.

- Minor Essential Revisions

1. In order to replicate this technique it is necessary to know the newly refined InterVA matrix probabilities. These should be included as an appendix, and should indicate how they differ from the original matrix probabilities.

2. It is unclear to me why the 100 PCVAs used to refine the model were all taken from Malawi. Please explain the reasoning behind this, as I think choosing 100 randomly (or 20 randomly from each cause) would be much more natural.

3. The abstract incorrectly states that verbal autopsy is the only method for analyzing cause of death in the absence of vital registration systems. There is also the "hospital method" [1], which should be mentioned.
4. The end of the introduction says “The InterVA model was not designed to handle stillbirths and was limited in … (data not shown)”. I found this unclear. What data is not shown?

5. The methods section states that “Negative and missing data do not affect cause probabilities”, which makes me think that InterVA is not actually applying Bayes Theorem. Please include the mathematical formula to clarify this point.

6. The global burden of stillbirths figure quoted in the discussion needs to include units.

- Discretionary Revisions

1. The introduction states that verbal autopsies interview the “next of kin”, which does not seem like the correct term.

References


doi:10.1371/journal.pmed.0040326

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.

Declaration of competing interests:

I am currently employed by the Institute for Health Metrics and Evaluations at the University of Washington, where I am actively researching methods for computer-coded verbal autopsy for neonates and evaluating the performance of these algorithms compared to InterVA.