Author's response to reviews

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

Authors:

Stefania Vergnano (stev@doctors.org.uk)
Edward Fottrell (Edward.Fottrell@epiph.umu.se)
David Osrin (d.osrin@ich.ucl.ac.uk)
Peter N Kazembe (pnkazembe@baylor-malawi.org)
Charles Mwansambo (cmwansambo@malawi.net)
Dharma S Manandhar (dharmasm@yahoo.com)
Stephan P Munjanja (spmunjanja@africaonline.co.zw)
Sonia Lewycka (s.lewycka@ich.ucl.ac.uk)
Peter Byass (Peter.Byass@epiph.umu.se)
Anthony Costello (a.costello@ich.ucl.ac.uk)

Version: 4 Date: 27 April 2011

Author's response to reviews: see over
Response to reviewers
To: Abraham Flaxman

Based on the response to reviewers, it is clear that I misunderstood the conclusion of the paper in my previous review. I interpreted the assertion that the algorithm is “reliable” as an informal statement that the results could be relied upon for public health decision making. However, the response makes it clear that the authors intended this to be a precise technical statement. When they say that “InterVA provides reliable information”, they mean that it will always produce the same (possibly incorrect) prediction of underlying cause of death when presented with the same VA interview results. This language is likely to be misunderstood by others as well, and should be changed. But more importantly, I find this conclusion much less interesting than the one I mistakenly criticized for over-reaching. Indeed this conclusion was known before the study began, simply based on the nature of the algorithm. I think the true focus of the paper is about the accuracy of the method, and the conclusion should be about accuracy (in balanced terms). Accuracy, not reliability, is what is important when considering using this method to gather information on neonatal mortality patterns for decision making.

We believe that it is not possible at present to be certain of the accuracy (or approximation to the truth) of any VA method to diagnose causes of death in populations where the majority of deaths occur at home. The technical concept of “accuracy” or “validating” VA is therefore complex as it is impossible to eliminate uncertainties and errors in both the model’s output and in any reference standard – this is discussed in our paper. Our study presents a thorough evaluation of the modified InterVA method and assesses performance in terms of similarity to independent physician assessments of the same data, acknowledging the limitations of this reference standard. The comparison presented in the manuscript tells us that, notwithstanding the discrepancies that are discussed in the manuscript, InterVA produces plausible cause of death profiles that are generally comparable to physician review of the same data and are likely to be useful in informing policy makers and for monitoring and evaluation activities. The “known” fact that the method is completely standardised (“reliable” in a technical sense) is also highlighted in the manuscript and we think it is worthwhile restating this fact as it is a major advantage of all automated methods compared to physician review.

To avoid ambiguity, we have modified the text of the conclusion to state that InterVA produces plausible results for stillbirths and newborn deaths, broadly comparable to physician review but with the advantage of complete reliability, allowing standardised cross-country comparisons and eliminating the inconsistencies of physician review in such comparisons.

- Minor Essential Revisions
Bayes formula in first paragraph of methods section is incorrect (the denominator should be a sum of products).

Amended

- Discretionary Revisions
None.

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 and evaluating the performance of these algorithms compared to InterVA.