Reviewer’s report

**Title:** Population Health Metrics Research Consortium Gold Standard Verbal Autopsy Validation Study: Design, Implementation and Development of Analysis Datasets

**Version:** 1 **Date:** 25 April 2011

**Reviewer:** Carla AbouZahr

**Reviewer’s report:**

This is an excellent article that addresses a topic of major public health importance. The authors have taken great care to carefully define the research questions and to develop appropriate methods for answering them. The methods are well described and the discussion and conclusions well supported by the data. The article is very well written. The findings represent a huge step forward in understanding the potential and limitations of BA for ascertaining causes of death. Given the extreme care that the authors have taken in developing the methodology and describing their findings, it seems almost churlish to express any criticisms. Nonetheless, there are some issues that will need to be considered prior to publication.

**Major essential revisions**

The most important issue concerns the selection of the health care facilities that were used to develop the gold standard cause of death criteria which provide the foundation for the validation database. As the authors note, the key challenge was to find deaths with sufficient laboratory, medical imaging and pathological details so that a gold standard cause of death assignment could be made. While the eventual cause of death lists were substantially more comprehensive than other used in verbal autopsy studies, they nonetheless reflect the rather special circumstances pertaining to the particular medical facilities in which the studies were undertaken. As the authors note, it was difficult to identify any deaths from some causes that would meet the strict gold standard criteria. None of the health facilities was “typical” of the developing countries in the study, several were private facilities; others were specialist centres that recruited patients with well defined diseases (cancers, HIV/AIDS) from a wide catchment area. The deaths identified in such settings are hardly likely to reflect overall community mortality patterns. Deaths among poor, disadvantaged and remote communities are less likely to reach these kinds of medical facilities that those among more fortunate population groups. It is possible that the facilities did indeed recruit from a broader population base but the authors provide no evidence of this. Some information on typical population intake (in terms of income, educational level, occupation etc. would be helpful for the interpretation of the results. Even if one accepts that it would not be possible to develop gold standard cause of death criteria in settings without such high end medical facilities, the authors should discuss the potential implications for the VA tool. Given, for example, that there were no deaths due to neonatal neonatal tetanus or measles, what are the
implications for the design the VA instrument? How should the finding that only obstructed labour emerged as a frequent cause of maternal death be interpreted? Community studies have found that deaths due to obstetric haemorrhage are less likely to occur in hospital settings because haemorrhage is so rapidly fatal. Does the absence of this cause of death in the health facility settings imply it cannot be validated for use on VA studies?

Minor essential revisions

In the Discussion section, the authors note the poor quality of hospital records and implications for reliable cause of death ascertainment even in health care facilities. It would be useful to provide some quantitative assessment of the reliability/accuracy of cause of death ascertainment based on routine medical certification compared with VA. If the high level laboratory, medical imaging and pathological details generate cause of death data that are considered 100% accurate, where do routine medical certification and VA fall on the scale (say 80% and 60%)?

Discretionary revisions

In the methods section, the description of the study sites is somewhat repetitive and reader understanding would be greatly facilitated by a table showing the distribution of deaths by age and sex across the various sites. Perhaps the table could also include data on average life expectancy in each study site (or failing that nationally) so as to enable readers to judge the extent to which the age and sex distribution of death is plausible in each setting. Not sure that religious affiliations add much to the argument, data on average household income might have been more useful for the analysis.

Section headed Data Collection, Identification of gold standard deaths para 6, starting “In Pemba…..” Line 6, inser “s” after “sign”

Line 8, lower case “Computer”

Section headed VA interview 2nd para, 1st line, suggest rewording “blinded interviewers” or repositioning “blinded”

Section Shortened Cause Lists, line 6, replace “AIDS and tuberculosis” with “AIDS with tuberculosis”

Level of interest: An article of outstanding merit and interest in its field

Quality of written English: Acceptable

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

Declaration of competing interests:

I have in the past been in receipt of short-term contracts from one of the authors. However, I do no belive that this impedes my review of this paper and therefore
declare have no competing interests.