Reviewer's report

Title: Performance of verbal autopsy, physician coding and classification systems in the Indian Million Death Study

Version: 1 Date: 16 September 2013

Reviewer: Gary King

Reviewer's report:

The intended contribution of this study is the metrics of evaluation. I think the authors have this completely wrong. The key issue here is whether the CSMF's are estimated correctly. And with the MDS, from the evidence they provide, there isn't much evidence that their goal has or has not been achieved. In particular, the inter-coder reliability for this study is remarkably low -- 66%. Even if these deaths were all correctly coded, which is unlikely, adding the unobserved 33% to one or more of the CSMF rates -- i.e., where they might belong -- could completely change the results of the study. If the MDS is going to be used as the authors describe, there needs to be serious validation. (1) When physicians agree, how often do they agree on the correct COD? This is a common issue, but needs to be studied. (2) More to the point, the pattern of selection bias that would result from omitting all the missing data must be shown to be uncorrelated with the quantities of interest in the study. If not, the bias could be huge. (3) The authors explain that half of the 1/3rd of the observations where physicians disagreed where adjudicated until agreement. This is a procedure, but the authors need to provide a rigorous validation, with external information such as medical verification from a subsample, that this procedure works to reduce the bias. There is no evidence (or even clearly stated assumptions) that would prevent this procedure from increasing the bias.

Overall, to refer to these as “good quality results” is unjustified.

Re the next section: It is inappropriate to use the results of the study to evaluate the methods. If you want to do that, there’s no reason to collect the data in the first place. The methods must be justified ex ante, without looking at the results. I see now justification for them here.

Sophisticated statistical methods exist with which to accurately -- with known degrees of bias and uncertainty -- estimate the quantities of interest in this study at lower cost. I don’t see why the authors are pursing such outdated and innaccurate methods.

a few other points:

Define “SRS” before it is used.

This needs to be made precise, so that a reader could follow the same rules. as is, I can’t tell what it means. “Odds ratios to compare CSMFs in various
sub-groups were adjusted for age (linear year), sex, religion (Hindu versus other), education (illiterate versus literate), poorer or richer state, and as needed, hospital versus home and rural versus urban status"

The section on ill-defined codes is ill-defined; it doesn’t explain what they are or how to recognize them if we see them, other than that a different study defined them. I can’t even tell what the odds ratios are trying to estimate. There are different ways of estimating odds ratios; this must be clarified.

adjustments in the tables must be described in sufficient detail so that they can be replicable. I can’t figure out what they are, except in the most vague sense.

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I have a patent in a related methodological area, although we give access to the software that implements the patent ideas for noncommercial activities, including for verbal autopsy analyses; I don’t regard it as a conflict of interest, but don’t know whether this meets the formal definition in this case.