Reviewer’s report

Title: Performance criteria for verbal autopsy-based systems to estimate national causes of death: development and application to the Indian Million Death Study

Version: 2
Date: 30 October 2013
Reviewer: Gary King

Reviewer’s report:

I have read this paper again, including the revisions and authors comments. I have no doubt that these are serious people, doing a large project, and working hard and earnestly to do as well as they can. However, the authors have made no major changes in response to my first review. Their response to the critique I provided seems to be that, among others (1) they can’t fix the problems, (2) the problems are too difficult to tackle given their research design, (3) coming up with an gold standard even on a subset of data is too difficult or impossible for them in how they set up their research, (4) there isn’t a better approach available to them, (5) other papers in "this series" (to which I do not have access by the way) do the same thing, (6) this meets the standards of the field. These justifications may each be correct, but I’m afraid none are relevant to whether the fundamental question: Does the science behind this article justify the scholarly (and policy) community relying on its results? To that question the answer seems absolutely unambiguous: The huge uncertainties -- well described in this paper by the way -- are considerably larger than almost all the "results" claimed in the paper. As such, the authors have made no detectable substantive contribution. Perhaps there is some small contribution in terms of describing how difficult it was to do this study or something, but I would very much hope that no one follows in their footsteps at least when it comes to the research design. If it is true that everyone else is doing this in the field, then maybe it will turn out I'm singling them out for inequitable, but not incorrect, criticism since I'm only reviewing this one paper; however, "everyone is doing it" is no justification for a research design that cannot support its conclusions. With regret, I recommend rejecting this paper, and encouraging the authors to develop a research design that tackles not only the very low levels of reliability but also gives direct measures of validity.

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.