Reviewer’s report

Title: Population survey sampling methods in a rural African setting: measuring mortality

Version: 1 Date: 27 February 2008

Reviewer: Samuel Joseph Clark

Reviewer’s report:

I enjoyed reading this manuscript and applaud the authors for investigating such an important set of questions.

Major Compulsory Revisions:

0. NONE

Minor Essential Revisions:

1. In general I would like to see more specific description of the methods and exactly what was done. For example on page eight in the first paragraph that discusses the methods of reproducing DSS sampling methods, I think there needs to 1) be more discussion of what a DSS typically is, and 2) a more detailed description of how the methods employed relate to real DSS designs. Perhaps a real example of a specific DSS would help. On the same page paragraph three, I don’t know what The selected parameters were selected from the Burkino Faso database means? And on the following page (9) I would like to perhaps see some pseudo code and/or more quantitative description of the sampling methods.

2. One of the main themes of the manuscript is that DSS data are and/or could be used more effectively to substitute for more traditional types of data in developing world settings, but that they have a potential problem with representativity. The manuscript then proposes to evaluate the extent of that problem. In order to set up this problem the average reader will need more information on what a DSS is and exactly what sampling methods are employed at real DSS sites; perhaps with one or two specific examples. My experience with DSS leads me to believe that very few were designed using a rational sampling strategy apart from put everyone in this geographic area under surveillance or conduct a clinical trial enrolling the population of this geographical area. I would appreciate a more detailed and nuanced discussion of this set of issues. I would also like to have the authors thoughts on how to do a better job at designing a representative DSS, taking into account that a DSS should also provide sufficient prospective, individual-level data to study risk factors and complex causal associations of other types.

3. I am confused by the choice of parameters that were used to assess the variability in sample statistic values across the different sampling methods.
For example, why not use full distributions for age and socio-economic status and compare the sampled distributions with the “real” distribution from the full census. I am of course curious about how well the different sampling methods reflected the top quintile of wealth or an age group other than 0-5, and I believe comparing the full distributions would be more informative. Along these lines, I wonder about the choice of mean values instead of median values for the type of comparison that is conducted? Likewise, I wonder if 20 samples for each sampling strategy is enough; I would have thought that on order 100 would be better?

Discretionary Revisions

4. A small point is that figures 1-7 may be more readable using box plots rather than the scatter charts (?) currently used.

5. Finally, I wonder if the authors could provide more detailed suggestions for how to improve both sampling design and choice of appropriate sampling designs in the developing world health and population contexts that interest them. Specifically, would population modeling of some sort be an aid in choosing a good sampling strategy that optimized on accuracy, logistical practicality, cost and other important dimensions? Would it be possible to simulate populations and develop a more sophisticated understanding of regular biases associated with different sampling methods, and subsequently defensible adjustments for those?

What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.