Author's response to reviews

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

Authors:

Edward Fottrell (edwardfottrell@yahoo.co.uk)
Peter Byass (peter.byass@epiph.umu.se)

Version: 2 Date: 4 April 2008

Author's response to reviews:

4th April 2008

Dear Editors

We are grateful for the detailed comments provided by the reviewers and we welcome the opportunity to respond and modify our manuscript accordingly.

We feel that many of the points raised by the three reviewers can be addressed by providing more background to Demographic Surveillance Systems (DSSs) in general, and specifically to their initial design and the sampling methods that they commonly employ. We now do this in the Introduction (page 5) and, as recommended by Samuel Clark, our revised manuscript now gives specific examples of DSSs in Ethiopia and South Africa. We feel that this more detailed description of DSSs also addresses Shea Rutstein’s comment that “usually all households within a site are interviewed to gather information on demographic events”, which is a common misconception.

We have added another table to the manuscript to address Samuel Clark’s desire for more detailed descriptions of the sampling methods and to summarise how the methods used relate to real field survey and DSS situations (Table 1). By clarifying the sampling methods used and their relevance to population surveys in rural African settings, we feel that this table also addresses Shea Rutstein’s confusion over what the aim of the paper is. As the title of the paper suggests, our work is applicable to survey sampling methods in a rural African setting in general. Although much emphasis is placed on DSSs, the approaches we use are relevant to surveys nested within DSSs as well as stand-alone surveys. We do not feel that sampling within a site is conceptually very different to sampling methods employed in establishing a representative DSS – the aim of getting a representative sample is the same and the biases caused by parameter distributions within the population are the same. In our study, the broad Immpact census data that we have used can be considered, for modelling purposes, as complete national, regional or DSS data. We are unsure what Shea Rutstein means by a “proper” national sample survey within the context of limited sampling frames.
Our use of the word “better” (page 6) refers to a subjective assessment of the quality and value of different sampling methods. In other words, we state that more complex methods may be perceived as being superior in any of a variety of ways than more simple methods. In response to the concern raised by Shea Rutstein, we have altered this sentence so that it is clearer that we are describing a possible subjective perception rather than a statement about statistical fact.

Shea Rutstein’s opinion that PPS is often used to control the size of the sample is true but it is incorrect to state that it is not used to control the probability of selection. Our definition of PPS, i.e. that it assures that those in larger units have the same probability of getting into the sample as those in smaller units, and vice versa, is widely accepted and is a justifiable reason for its use. We have cited additional work to support our definition of PPS methods (page 8).

We have modified the sentence highlighted by Samuel Clark beginning “The selected parameters…” on page 9 to clarify our meaning.

Somnath Chatterji feels that we should clarify that the strategies that we present apply to the level above households. We feel that this is made clear in the Methods section when the definition of concession is given and we have now added a paragraph to the Discussion (page 11) highlighting this but also emphasising sampling at the concession level in our study is conceptually very similar to sampling at the household level. In the same paragraph, and in response to comments by Somnath Chatterji, we also emphasis that our study reports on household-level parameter estimates and that the sampling methods employed might produce different results for individual-level parameters.

This study does not attempt to address the issue of sample size and we do not agree with Shea Rutstein’s comment that the paper “seems to regard the 1% sample as ideal”. The 1% sample size used in this study relates to the premise that active follow-up in DSSs can only be justified if it can meaningfully be extrapolated into the surrounding 100-fold population. We agree that this sample size may not be appropriate for measuring all indicators and have emphasised these points in the new paragraph added to the Discussion on page 11.

In response to comments by Samuel Clark, we have modified the final paragraph in the Discussion (page 16), suggesting ways in which sampling strategies may be better informed by empirical studies and population modelling.

Samuel Clarke describes confusion over the choice of parameters that we used to assess the variability in sample statistic values across the different sampling methods (point 3). While it is true that we have made somewhat arbitrary choices, this reflects what is often done in practice.

In response to Samuel Clarke’s other comments (point 3), the figure presentation of our results as they currently are facilitates comparisons with an earlier paper on the same topic that uses English census data (reference 17 in the manuscript). Taking 20 samples for each parameter is intuitively consistent with the arbitrary p<0.05 criterion often used in the analysis of sampled data.
We look forward to your response.

Yours faithfully,

Edward Fottrell and Peter Byass