Reviewer's report

Title: Inclusion of mobile telephone numbers into an ongoing population health survey in New South Wales, Australia using an overlapping dual-frame design: final weighting strategy

Version: 2
Date: 20 June 2014

Reviewer: Carol Pierannunzi

Reviewer's report:

I read with interest the manuscript entitled “Inclusion of mobile telephone numbers into an ongoing population health survey in New South Wales, Australia using an overlapping dual-frame design: final weighting strategy.” This paper is on a timely topic that is being faced by many surveys which rely on telephone interviews. Most government surveys which rely on telephone interviews are conducted on both landline and cell phones, some with dual frame samples and some with overlapping samples. Australia seems to be at a disadvantage in constructing design weights, due to the lack of information on phone use and phone ownership at sub-national geographic levels. This is less of a problem in the US, but is likely to be a situation faced by a number of researchers who have geographic areas of interest where phone use information is unavailable.

Overall the paper is a description of the issues that the researchers faced and their actions taken to deal with weighting without control totals. The methods are appropriate and draw from current established practices for dealing with probability of selection. I think that many readers would benefit from the discussions of the steps taken by these researchers. I think that the discussion would appeal to more readers if it was less specific and focused on the issue of weighting effects (cutting the imputation and data management sections)

Major revisions:

1. I think that the paper would be better cast as a discussion of establishment of design weights when control totals are not known. It reads more like a story of what happened rather than an academic paper on the appropriate use of weights. It is also cast as a description for just one case, but in fact it would be more useful to readers as a description of a method that was used in NSW but which can be used in any scenario where there are no benchmarks.

2. Parts of the discussion section should be moved to methods section of the paper. For example, the discussion on how they acknowledge that the composite factor of .5 is somewhat arbitrary includes the information that they had considered other control totals. It would help the reader understand the methods if this had been presented earlier in the paper. I was wondering about this decision until late in the paper. In addition the information that the sample was disproportionate within each of the health administration area needs to come
earlier in the paper (it is not clean until after the reader is internally asking questions about probability of selection).

3. The tables are cumbersome. I do not think that most readers will need the information in Table 2. The explanation in the text of imputation is fine. This is not the purpose of the paper and is distracting. Likewise the information in table 3 is not helpful for the reader to understand the design weighting process. A brief discussion in the text would be better. A table comparing the old weighting process (using just the landline respondents) and the new process with examples of similarities/differences in the weights is needed and essential to the conclusion that the weighting processes are “similar to, and in many cases, less than, the effects found” in earlier iterations of the survey. The authors tell us that the weights are similar, but it needs to be illustrated in tabular format with the actual weights/weighting effects.

Minor Essential Revisions:

4. I would consider taking the children out of the weighting discussion in the paper. I understand that the weights for the children were needed in the real world, but this is a discussion of a method, and the fact that there are only 138 (with only 26 in the cell phone sample), is problematic.

5. I had a couple of questions on the Figures which were not addressed in the discussion. Tell the readers why there are so many low weights for the mobile frame (even through the means of the weights were higher than for the weights of the landline). I think that I understand why this could be true, but the authors do not address it. Overall I thought the Figures were helpful and appropriate to the discussion. I would do something to label them more effectively.

6. The title of the paper correctly makes the reader assume that the paper will focus on a strategy. If the focus of the paper is changed to be more of a discussion of appropriate methods for developing design weights where there are unknown controls, I would also adjust the title. It is not just about NSW—it is a method for creation of weights in any context where there are no benchmarks.

7. There are a number of typos (p. 7 first line under Calculation of the weights “was” should be changed to “were”; p. 7 “9.8 million mobile telephone numbers” appears as “9.38” in the corresponding table—one of those is off).

Discretionary revisions

8. There needs to be some discussion on why the dual frame sample was not used or considered and the effect that a dual frame would have had on the design weighting steps. This can go in the background or methods section.

9. There needs to be some explanation of what other data might have been used to estimate phone use. There is no mention of whether there are data on the number of households and number of phones (landline and/or mobile) within any geographic area. I did not see any mention of the number of households. Perhaps the authors did look or have these data and they were not useful, but it is not clear.
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

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

Declaration of competing interests:
'I declare that I have no competing interests