Reviewer's report

Title: Changes in the Determinants of Exclusive Breastfeeding in Nepal: Comparison of National Surveys 2006 and 2011

Version: 2 Date: 2 July 2013

Reviewer: James Stanley

Reviewer's report:

This is a statistical review for Khanal, Sauer, and Zhao, “Changes in the Determinants of Exclusive Breastfeeding in Nepal: Comparison of National Surveys 2006 and 2011” as submitted to BMC Public Health.

Dr James Stanley

This is a topic of interest: the study question is well defined, and important to address, and the use of a strongly developed national survey to approach the question is a great use of resources.

Unfortunately, I have very strong concerns about some statistical issues – most importantly the fact that data analysis methods have not taken account of the complex sample design of the original surveys. I feel this would need to be addressed before the paper could be accepted: otherwise the results will be misrepresentative of the population sampled (i.e. inferences drawn from the results will be incorrect.)

I also had a few concerns with some other elements of the statistical design, which are outlined below. There are a few other minor points that could potentially be addressed as well (but which aren’t statistical issues per se.)

MAJOR COMPULSORY REVISIONS:

1) There is no information presented here about whether the data analysis took account of the complex sample structure (clustering by PSU, weight of selection.) If such methods are not used, then statistical inferences drawn from regular data analysis techniques will be incorrect (and can be quite substantially incorrect, particularly in respect to underestimation of standard error.)

I think that SPSS (as per page 6) requires an add-on package to handle such considerations: if this add-on was used, then this needs to be mentioned in the text. If the sample structure was not considered in data analysis, then the analysis needs to be repeated using such methods.

I’m afraid that it looks to me like this is the case – for instance, I could replicate the confidence intervals for reported proportions, e.g. top of page 10, and the p-values for chi-squared tests in Table 1 (for sex differences in EBF), with no information about the survey design.
(n.b. SAS, Stata, and R all handle complex survey methods quite well – R being the only one requiring an add-on package to do so, with both core R and the Survey package being free to use.)

The survey report online for 2011 at http://www.measuredhs.com/pubs/pdf/FR257/FR257%5B13April2012%5D.pdf makes mention of the complex survey analysis issues: the average design effect across measures was 1.86: the report authors state p. 269 “This means that, due to multi-stage clustering of the sample, the average standard error is increased by a factor of 1.86 over that in an equivalent simple random sample.” which means that there is considerable over-optimism in the statistical results presented in the paper.

On p. 11/12 for examination of changes between surveys: the comparisons here should also take into account the complex survey structure. Combining results across two complex surveys is a hard enterprise: it might be sufficient to compare the point estimate prevalences and confidence intervals (appropriately calculated, as per previous notes, to account for the survey design) to draw inferences here rather than relying on hypothesis tests.

2) Bivariate screening of variables to include in regression model

p. 10 states “The variables that were significant in Chi-square test (Table 1 and 2) were then further analysed using multiple logistic regression with a backward elimination procedure (Table 3)” Screening of candidate variables based on bivariate associations is not considered a suitable predictor selection tool: the most important issue is that if a particular variable (e.g. ethnicity) is confounded by another variable (e.g. father’s occupation) then the first variable can very well be not significant, even if there is an association with the outcome (that is confounded by the second variable.)

For more notes on model selection methods, see e.g.
Harrell “Regression Modelling strategies”, p. 56-60 (should be available on Google books preview?)
Vittinghoff et al., “Regression Methods in Biostatistics: Linear, Logistic, Survival, and Repeated Measures Models” -- Predictor selection chapter, particularly p. 150-151

See also e.g. http://www.ncbi.nlm.nih.gov/pubmed/8699212

3) Use of Stepwise backward elimination I’m not a particular fan of stepwise model selection methods, as they likely lead to inflated Type I error rates and overoptimistic standard error estimates, but stepwise backward elimination is generally considered the most robust method. I should also note that the process of using such methods is complicated further as per point (2) regarding the bivariate result screening process, and point (1) regarding the computational issues to take into account the survey design in calculation of statistical results.

One starts to see issues with the backwards selection method appear in the first
paragraph of p 11 and corresponding Table 5 – with the exception of infant age, factors that were significantly associated with EBF in 2006 are no longer “significant” in 2011 – comparing the odds ratios and 95% CI between the two surveys reinforces this point in that there is a quite considerable degree to random error in the estimated associations. This is probably due to the problems with predictor selection methods for the most part, but also small sample size for 2011 (see point 4.)

4) consideration of sample size

p. 10 says that “Based on the result of the Chi-square test, only the age of the infant was a significant determinant”

The reduced sample size in 2011 (which is quite small from a cross-sectional survey perspective) means that relying on significance per se to decide on the relative importance of variables is haphazard.

MINOR ESSENTIAL REVISIONS

5) on p. 5: The power calculation seems retrospectively added in (as per a reviewer’s request): why is it important to be able to detect a 32% difference in EBF rate, and is the 4 months vs. 5/6 months comparison the main comparison of interest? There could be better justification here.

Also, simple power calculations are likely to be over-optimistic in the case of complex survey data: usually one might multiply the sample size required by a design effect derived from the study sampling parameters.

6) p. 7/8: Father’s occupation is not mentioned here, but in Tables 1/5 it is described as “Agricultural”, “Non-agricultural” and “Other” – on first glance the latter two categories are somewhat confusing as they suggest the same thing?

7) p. 8 “Haemoglobin was measured in every alternate household” Was this variable used in the data analysis? If so, what are the implications of this for the model selection process, given that this would limit the pool of survey members who had “complete” data? Anaemia is mentioned on p. 9, but appears to be missing from subsequent tables.

8) p. 38 Table 7 has a highlighted heading “contraceptive” which seems incorrect.

DISCRETIONARY REVISIONS

p. 6: “The rate of EBF was reported at less than one month, one, two, three, four, five and six months to identify changes in the rate of EBF according to the age of the infant.” It isn’t clear whether this means that historical EBF information is being elicited for each child at each of these ages, or is being reported based on the current age of the child at the time of the survey (later it seems clearer that it is the latter.)

p. 7 First mention of “dalit” should be capitalised? (second paragraph) And perhaps the “relatively advantaged” category membership (“Brahmin, Chhetri…”)

could be shifted to the start of this paragraph since the other caste/ethnicity groupings are considered there.

p. 25 Tables 1 to 3: I found it somewhat confusing to read these tables as the percentages in the two columns mean different things: the left hand column ("Total N [%]") describes the relative breakdown of the survey sample by sociodemographic characteristic (i.e. a column percentage) while the right hand column ("EBF n [%]") gives the breakdown of EBF at each level of the sociodemographic characteristic (i.e. a row percentage.) This could be signalled more clearly as the visual temptation is to try and compare percentages in the columns (which is exacerbated the further one gets from the top of the table.)

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I have no competing interests.