Reviewer’s report

Title: Reweighting National Survey Data for Small Area Behaviour Estimates: Modelling Alcohol Consumption in Local Authorities in England

Version: 0 Date: 13 Jun 2018

Reviewer: Rise Goldstein

Reviewer's report:

This paper reports a method for reweighting national survey data, which its authors then apply to the Health Survey for England. The method combines survey and publicly available data from sources external to the survey, to produce simulated locally representative survey data and provide statistics of alcohol consumption for each Local Authority in England. Based thereon, the authors identified a 2-fold difference in estimated mean alcohol consumption between the lightest and heaviest drinking Local Authorities, a 4.5-fold difference in rates of abstinence from alcohol, and a 3.5-fold difference in harmful drinking, defined as more than 35 units (approximately 12 fluid oz.) of ethanol, or about 20 U.S. standard drinks, per week for women, and 50 units (about 17.2 fluid oz.) of ethanol, or about 29 U.S. standard drinks, per week for men. The method was found to compare well to direct estimates from the data at the regional level. The authors conclude that, in addition to important policy implications in its own right, their method and the reweighted data it produces can also be used for modeling local policy effects, and for other public health small area estimation where locally representative data are not available.

Clearly there is a need to develop and implement methods that are both valid and user friendly for small-area estimation from larger regional or national survey data, since mounting small-area surveys with large enough samples for reasonable statistical power and precision quickly becomes resource prohibitive. The method presented by these investigators appears straightforward and adaptable beyond the immediate context in which it was applied in the present report. I only have a few comments:

1) It seems a bit odd to use consequences of alcohol consumption (e.g., hospital admission and mortality rates) to estimate consumption, especially since there is wide interindividual variability in the level of consumption required to trigger alcohol-related medical complications. Some further explanation and justification of the use of these measures as part of the estimation methodology would be useful.
2) More generally, the independent variables the authors use to model consumption are specific to the British context. This is of course appropriate given their specific task. However, because the readership is international, I would recommend that they take a step back, whether in Methods, in the Discussion, or both, to describe in more generic terms the kinds of measures that can be brought to bear in adapting their method to the estimation of consumption in other national contexts.

3) These investigators are correct (p. 6, lines 158-159) that multinomial logistic regression allows the most flexibility in modeling and does not impose an assumption of proportionality of odds. However, there are other ordinal models that might be appropriate here and that do not require an assumption of proportionality of odds. If there truly is ordinality in these data, then the use of the multinomial model comes at a cost of loss of statistical power. If they have not already done so, the authors should examine this issue with expert biostatisticians.

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal