Reviewer's report

Title: The impact of attrition on the representativeness of cohort studies of older people

Version: 1 Date: 17 June 2010

Reviewer: Andrea Marshall

Reviewer's report:

Major comments
1. This paper provides examples of the amount of attrition and death likely in an example cohort study and the bias in terms of the prevalence of a risk factor from a simple hypothetical situation and example study. However, the conclusions drawn from this research are as expected and therefore not novel. The conclusions should be more specific to the results obtained.

2. In the background, it states that the study results may not be generalisable to the target population, however the effect that the bias in prevalence of a risk factor has on the overall results has not been explored within this paper.

3. It would be useful to have some guidance/references as to how to handle both types of attrition if it occurs. Rather than just saying it is a problem

4. In the methods section, how were the assumptions for the hypothetical situations derived? Are they based on realistic levels? How likely is the situation where bias is present at the beginning of the study? Can you give an example?

5. On page 6, it says that the length of time can be arbitrary, but what is it in the hypothetical situation considered, e.g. yearly?

6. In the ALSWH example, why was over-sampling of women from rural and remote areas intentionally.

7. Provide details of what is meant by the alcohol consumption categories.

8. More details of the multinomial logistic regression model should be provided in the methods section such as how the covariates were fitted in the model and how reference values were chosen when fitting covariates as categorical? Were all covariates fitted or a stepwise model used?

9. Why were different categories used for the country of birth in the cohort compared to the target population?

10. On page 8, what is the proportion of all women that was included in the Australian National Health Survey?

11. On page 9/10, provide the direction of the association when describing the figures. Also provide what happens if the association is in the opposite direction for Figure 4.

12. The results reported in the text from the multinomial model are selective, is there any reason for this? It should be stated that the odds are compared to
those responding. The 95% CI for physical activity categories included 1 for the women lost to follow-up and therefore should not be included in the associated list of risks factors in the text.

13. It is unclear why the low-risk drinkers in table 2 were the reference category?

Minor revisions
1. Page 4 line 5, needs and “AND” instead of a comma
2. Page 6 line 4 from bottom, need to be “through TO SURVEY 5”
3. Page 7, kilograms per metre squared is usually written as kg/m2. This needs to be amended.
4. Avoid using the abbreviation NHS.
5. On page 11/12 say that the prevalence in the cohort was lower than in the target population as this is not entirely clear at present
6. Label the figures as a) and b) for the left and right panels.

Discretionary revisions
1. The hypothetical situations considered are rather limited and the research would have been improved with a more comprehensive hypothetical study, for example, looking at different levels of prevalence, prevalence levels that change over time in the population and also the impact on the overall results of the cohort study from the differences in the prevalence.

Level of interest: An article of limited interest

Quality of written English: Acceptable

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

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