**Reviewer's report**

**Title:** Assessing Nonresponse Bias at Follow-up in a Large Prospective Cohort of Relatively Young and Mobile Military Service Members

**Version:** 3  **Date:** 19 July 2010

**Reviewer:** Anne Tate

**Reviewer's report:**

Major Compulsory Revisions:
Unfortunately, the authors replies do not adequately address my main concern about this report, which is that they cannot justify the claim made in the conclusions i.e. "These findings suggest that analyses from this cohort should yield reasonably unbiased results.

In order to in any way justify these conclusions - which in any case rely on many assumptions (as pointed out by reviewer 2) - they will need to compare the characteristics of the 55k responders with the original random sample of 246k. Although, given the low response rate compared with the original sample (20%) they would still need to very cautious in making such a claim.

The reason that epidemiological cohort studies are based on random samples is that a random sample should reflect the characteristics of the population being studied.

If response is not 100% then, since the responders are not a random sample, there is likely to be bias in the results due to nonresponse.

One can compare the characteristics of responder and non-responders in order to assess this bias and, if one has enough information one can adjust for this bias.

In a follow up study, the same thing applies. In order to adjust for bias, one still has to compare with the original sample, since the sub group who replied were not randomly sampled.

There seems to be no justifiable reason for not comparing responders with the original sample, given that most of the characteristics used in the weights were measured at baseline and thus the authors will have access to the values of these characteristics. It is not sufficient to quote previous work to justify this.

If the authors are not able to do this, then they should remove this claim from the conclusions.

It is also very unclear how much the new measures that were obtained from the baseline questionnaire affected response and also the probability weights. From what I can understand they may not affect them very much. Since most of the predictors of nonresponse were based on baseline (and thus more accurate)
measures it might be better to use only the baseline measures in the first instance and then investigate how using the self-reported measures would affect these weight.

**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