Reviewer's report

Title: Sleep disturbances in an arctic population: The Tromso study

Version: 1 Date: 22 January 2008

Reviewer: Anthony Staines

Reviewer's report:

I was tasked as follows :-
"This study protocol obtained ethical approval but not external funding. In these cases we ask advice regarding the following questions:-
1. Will the study design adequately test the hypothesis?
2. Are sufficient details provided to allow replication of the work or comparison with related analyses: if not, what is missing?
3. Is the planned statistical analysis appropriate?
4. Is the writing acceptable?"

The aim of this study is :-
"Using validated sleep assessment tools, our aim is to describe and analyze sleep disturbances, with their correlating and predisposing factors."

1. Will the study design adequately test the hypothesis?

There are two proposed studies - a basic prevalence study, and a nested case-control study within the cohort, where cases are identified by a 2-stage screening process. Controls are matched to the cases (how?), and both receive a more detailed sleep questionnaire (Q2). You might have a look at some other nested case-control studies.

Basic prevalence study

Issues that emerge are representativeness presumably the study collects enough data to allow you to say if participants are reasonably representative of the general population.

Screening question it's obviously a sensible strategy to get a subset of the study group at higher risk of sleep disorder but is it known how reliable this question and this cut-off are?

Sample size you expect to get 10% to 15% of about 12,000 people screening positive quite a big sample, should be big enough for most purposes, but a formal sample size calculation would reassure. There must be one for the main study. The overall sample size must be big enough to measure the prevalence quite reliably.
Analysis of risk factors

One wrinkle not addressed is what happens if a 'control' reaches the 'case' levels on the sleepiness questionnaire - this should allow you to measure the reliability of the screening questionnaire.

How will you match 'cases' to 'controls'?

What response rate do you expect to Q2?

The piece about using the questionnaire at different times of the year is quite confusing - when will the study be done? Do you expect a difference between answers to the screening questionnaire (Q1) at different times of the year? This needs to be actively managed, and not simply let happen!

Q2 could be further described - could you attach it to this paper?

Your terminology for variables is a little confusing. I suggest you distinguish between outcome variables - these are what you are trying to explain e.g. high scores on the ESS/PSQI; explanatory variables these are what you think might cause your outcomes that is of interest to you e.g. self-reported psychiatric symptoms; confounding variables these are alternative explanations of what you see, or other explanations that are not of interest to you e.g. age, ethnic group etc....

You must do a formal power calculation - what is the smallest Odds Ratio you are likely to be reliably able to distinguish form unity?

2. Are sufficient details provided to allow replication of the work or comparison with related analyses: if not, what is missing?

It would be useful to provide a set of pre-specified associations to study, with a justification for each from the literature. This would be an expanded Table 2.

3. Is the planned statistical analysis appropriate?

Yes, to a point.

First, the main outcome variables need to be operationally defined, and you need to explain why you elect to analyse what are basically continuous variables, for example the ESS goes from '0' to '24', as categorical variables, for example, excessive sleepiness 'yes' or 'no'. This can be perfectly reasonable but you need to justify it.

Your analysis strategy is a bit confusing. You are suggesting using both regression models and the older pre-regression techniques like the Mantel-Haenszel. Note that if you do logistic regression you can either do matched analysis, or you can break the matching, and do ordinary logistic regression with the match variables added in to every single analysis regardless of 'statistical significance'. 
I suggest you review your strategy like this :-

First we will carry out a series of tabular analyses, tabulating our outcome variables against potential explanatory and confounding variables. (For continuous outcome variables you need to do plots instead). These tabulation will guide further analysis. A matched regression analysis will be done using conditional logistic regression in SPSS. For each model we will compare the results of univariate analyses to those of multivariate analyses. Our goal will be to construct credible parsimonious explanatory model. Models will be critiqued using residual analysis, and the Hosmer-Lemeshow goodness of fit test.

Missing values the usual approach to deal with missing values is some form of multiple imputation. SPSS has some facilities for this, but I am not familiar with them. The approach you describe seems quite unsuitable.

Variable selection - it is not appropriate to use p-values below 0.05 for variable selection for multi-variate models. The relevant criterion is does the additional variable make a difference to the model or to any of the other variables?

4.Is the writing acceptable?

Yes quite clear. There are one or two odd phrases to the native-speakers ear, but nothing of great importance.

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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'