Reviewer's report

Title: Determinants of participation in a longitudinal two-stage study of the health consequences of the Chornobyl nuclear power plant accident

Version: 1 Date: 3 December 2007

Reviewer: Linda Grievink

Reviewer's report:

This is an interesting paper that studies the determinants of participation in an Eastern Europe population. Different populations groups (western and non-western) have different reasons to participate. The results are helpful in the future when preparing longitudinal health studies after disasters.

Major compulsory revisions

1. From the data-analyses section and presentation of the results I conclude that the author's assume that the evacuees and (classmate) controls have the same response predictors, i.e. that there is no effect modification whether or not you are evacuated. I conclude this because they are pooling the two groups in the analyses and are adjusting for the group status. However, in this way you can not study effect modification. You loose some additional information that is worth looking at. Additional analyses are needed. Each table should be analyzed for the evacuees and for the controls separately. If the OR's are similar for both groups, you are allowed to pool the analyses. This should be described in the data-analyses section and in the results section. The last column in each table is redundant.

2. The analyses of variance are used for the continuous variables. This is a good analysis for descriptive analyses. However, I could argue if this is the best way to study the predictors of follow-up. For the multi-variable regression model, this analysis can not be used. Unclear from the paper is how these continuous variables are treated in the multivariable regression model. I assume that the variables were dichotomized. Why not use the dichotomized variables in the crude logistic regression analyses instead of the analyses of variance?

3. The analyses do not show the multi-variable analyses in the table. This way the reader does not know the independent variables that predict the follow-up. Only the significant predictor is mentioned in the text (unknown is whether statistical significance is meant or a clinical significance). Since the last column in the table is redundant (see point 1), I would like to see the multivariate logistic regression analyses in the last column. Personally I am not a fan of step-wise backward-elimination techniques. Since this is a matter of taste, I leave it up to the authors how they do the multivariate analyses.

Minor essential revisions

1. From the abstract it is unclear to the reader whether the response figures
mentioned were from both stages and from evacuees/controls.

2. Background, third paragraph, second sentence: "#studies of toxic exposure are often implemented". Often? To my knowledge I do not know other studies after disasters that choose for such design. Maybe the authors mean that this is done often after other studies. Clarify this to the reader by adding references. In the whole paragraph the studies among the general population and among survivors of disasters are mixed. Do the authors suggest that these studies can be compared?

3. Sample. Clarify to the reader why this a high risk group for thyroid cancer.

Discretionary revisions

1. Background, first sentence. Although we could argue until what time long-term follow-up is needed, the reader likes some clarification on why long-term follow-up is needed.

2. Second paragraph, first sentence is unclear to the reader.

3. Sample and procedure. Which ethical constraints are there on collecting data of non-participants?

4. The description of the design is not very clear on page 6. Who were the participants at both times? Were both the mother and child at both times interviewed and had both a medical examination? What was the length of the interviews? Were they real interviews with an interviewer or did the participants fill in questionnaires themselves (either on the computer or on paper)? The description of the study design is important since the objective of this paper is to improve the design to increase the response in future studies.

5. Results, first paragraph. Since this journal does not have a space limit, a table in which the non-response groups are compared can be included in the paper. The reader can look at the results themselves.

6. Results, page 10. The paragraph with all the numbers of response is hard to read. My suggestion is either include the numbers in Figure 1 or in a table. In addition, why describe both the conditional participation and unconditional participation rates? This differences is not explained later on in the discussion.

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

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.