Author's response to reviews

Title: Identifying specific non-attending groups in breast cancer screening - population-based registry study of participation and socio-demography

Authors:

Line F Jensen (line.jensen@alm.au.dk)
Anette F Pedersen (AFP@ALM.AU.DK)
Berit Andersen (Berit.Andersen@stab.rm.dk)
Peter Vedsted (p.vedsted@alm.au.dk)

Version: 2 Date: 23 August 2012

Author's response to reviews: see over
Dear Editor,


Thank you for the opportunity to revise this manuscript. We are very pleased to learn that you found the article of relevance in its field. Please find below a detailed description of our response to each of the points made by the referees.

All authors declare no competing interests. This is included as a separate section in the revised manuscript after the conclusion.

We hope that this cover letter and the revised manuscript will answer the raised issues and will be suitable for publication. Do not hesitate to contact us if there are any further issues to be handled.

On behalf of the authors,

Yours sincerely,

Line Flytkjær Jensen
PhD Fellow

RE. Referee 1:

1) I suggest that the aims are merged as "........characteristics of participants and active- and passive non-participants."

We agree with the referee that this is a good idea, and the paragraph describing the objective of the article has been reworded following the suggestion of the referee.

2) I am not very fond of this step-wise analysis. Why not perform a classical analysis starting with a bivariate analysis and then include the significant (p<0.05 or p<0.10 depending on how many variables you want in the final model and how you plan to use the model) variables in the multivariate analysis.

There are several reasons for the chosen analytical strategy we have chosen. First of all, because of the large sample size and because of the very clear associations between almost all variables and participation in the unadjusted analysis, all unadjusted associations were highly significant (p<0.0001 – except the variable ‘previous cancer diagnosis’ which was statistical significant with a p-value of 0.0083). Because of this, the approach suggested by the referee would result in a full model omitting the important information gained from adding new variables to the model.

Secondly, within social epidemiology, directed acyclic graphs (DAGs) are a well known approach and are recommended by several of the leading social epidemiologists (e.g. Dr Kenneth Rothman and Dr Michael
DAGs are useful to guide the statistical analysis when dealing with the complex associations often characterizing socio-demographic variables. However, we do acknowledge that our chosen method was not sufficiently described and we appreciate the referee for directing our attention to this. The referee’s comment on our statistical approach made us reconsider our models and because of this, the following changes to the manuscript have been made:

- The DAG figure was included in the manuscript so that the reader can follow our method better
- The models were altered a bit and some of the variables were included together. Now the results comprise an unadjusted analysis, a basis model and a model 2 and 3.
- The analysis section (pp 6-7) has been revised so that we explain more thoroughly what the purpose of DAGs is and how we plan to use it in the article
- Following the changes outlined above, the results section were altered as well
- Finally, we included a section in the discussion where we discuss the use of DAG.

Hopefully, this will make it clearer to the reader why this theory-based method was used for this study.

3) **Table 2 and 3 must be included in the paper and not as supplementary files. The tables should always include number of non-participants and number of participants**

Table 2 and 3 are now included in the paper as requested. Also we have incorporated the total number of respondents in each table together with number of non-participants (in Table 2) and passive non-participants and active non-participants (in Table 3). This enables the reader to assess the distributions of the groups included in the analysis.

4) **Minor comments: Material and Methods**

The ‘s’ in material has now been removed.

5) **Registries and variables: Specify western countries**

We have used the definition of western countries proposed by Statistics Denmark. Instead of listing all western countries we have incorporated our source as a reference, so that the reader can look up the definition used.

RE. Referee 2:

1) **In Table 2 and 3, the reason why these 6 variables, not others, were selected for Model 1 should be explained more properly.**

As described above we acknowledge that our method of using DAG and the subsequently regression models was not properly described. We describe this more clearly in the revised version.

The criterion for variables to be included in the basic model was that they should not be intermediate variables. If we take ‘age’ as an example: No other variable could ‘lead’ to the women’s age, which subsequently could affect the women’s participation. Contrary, age could be associated with e.g. the women’s income (e.g. older women often make less money than younger women), and income can affect women’s participation. Hence, income is considered an possible intermediate variable and was not included in the basis mode.
The variables we assessed not to be intermediate variables were included in the basic mode as displayed in the DAG (Figure 1 – in light blue color).

2) There might be other confounding variables which are not analyzed in the study, such as family history of breast cancer, so the authors should state this kind of limitation.

This is a very good point - residual confounding cannot be excluded. This has been incorporated in the revised manuscript in the Discussion section.