Author's response to reviews

Title: Breast density in birth cohorts of Danish women: A longitudinal study

Authors:

Sophie S Hellmann (sohe@sund.ku.dk)
Elsebeth Lynge (elsebeth@sund.ku.dk)
Walter Schwartz (walter.schwartz@ouh.regionsyddanmark.dk)
Ilse Vejborg (ilse.vejborg@rh.regionh.dk)
Sisse H Njor (sissenj@sund.ku.dk)

Version: 3 Date: 7 February 2013

Author's response to reviews: see over
February 07, 2013

Editor-in-Chief
BMC Cancer
Point-by-point response letter

Dear Editor

Thank you for your kind interest in our manuscript entitled “Breast Density in Birth Cohorts of Danish Women: A Longitudinal Study”. We have addressed point-by-point the concerns of the reviewers’ reports and hope that you will find all concerns carefully addressed.

--oo0oo--

Point-by-point response to review comments by Giske Ursin

1. **Concern:** The authors use mammographic density data obtained at different time periods to determine whether the changes over time are predominantly due to age or cohort effects. The problem is that the density is not assessed all at one time, but at each screening. Therefore changes in film, technique, equipment and assessment over time could have affected the mammographic density readings. How this was in Denmark during this time period is not discussed. This can be a challenge because if there were systematic changes then this could have caused some of the findings that are attributed to cohort effects.

   **Answer:** Throughout the planning of the study, analysis, and interpretation of the findings, we have been very aware of the question addressed by the reviewer. Even though we did have screening data available until 2008, we decided in close collaboration with the radiologists in charge of the two independent screening programs, to take into account the concern raised by the reviewer by restricting our analysis to the time period from 1991-2001. We restricted the analysis to the period 1991-2001, because there were no systematic changes in assessment practice or screening techniques in the two independent screening programs during this time period that could cause the birth cohort effects found in our study.

   We have made this important concern even more clear in the Discussion (page 14), 1st paragraph, line 3-5: “We therefore restricted the analysis to the period 1991-2001, where no systematic changes in screening practice or assessment took place in the two programs that could have caused the findings attributed to the birth cohort effects.”.

   However, to further ensure that we accounted for the reviewers concern, we conducted sensitivity analysis in subgroups of women with mammographic density assessments conducted within the same screening rounds of the screening program. The sensitivity analysis confirmed within all five screening rounds analyzed separately the findings of the study of larger proportions of mixed/dense compared to fatty mammographic density in younger compared to older birth cohorts of women. We have addressed this issue in the Results (page 12), 2nd paragraph, line 9-11: “Sensitivity analysis in subgroups of women with mammographic density assessments conducted within the same screening period confirmed the age- and birth cohort patterns found in the overall data (data not shown)”. 

Page 1 of 10
We also kindly pay attention to the study by Anne Helene Olsen et al entitled “Breast density and outcome of mammography screening: a cohort study” published in the British Journal of Cancer 2009 analyzing the same dichotomized outcome from the Copenhagen mammography screening program, where an increased risk of breast cancer RR 2.45 (2.14-2.81) was found in women with mixed/dense compared with fatty mammographic density, which is in agreement with existing literature.

In our opinion, it is therefore not likely that the birth cohort effects found in our study could be attributed to systematic changes in the screening program over time in the two independent screening programs in Copenhagen and Funen within the restricted time period 1991-2001 analyzed in our study.

2. **Concern:** An added problem is that the assessment (whether a woman had a fatty breast or not) appears to have had clinical implications, i.e. if a woman was found to have mixed or dense breasts, then she would be called in for a two-view mammography at the subsequent screening, otherwise only one view. This may be why there are essentially no declines in % of women with mixed/dense breasts from their first image in their early 50s to ten years later (diagonal in Table 2). This is a problem with the design and may be why there are no changes with age within one cohort, which we would have expected based on other studies.

It also seems possible that a radiologist’s previous assessment of each woman could have affected the current assessment – again this reflects the challenge with this design. There was no way of doing this blinded, because obviously either the radiologist had two views (i.e. previously this was assessed as mixed/dense) or there was one view (previously this was assessed as fatty). Thus, this appears to be another challenge.

**Answer:** The radiologists assessed women’s mammographic density as either fatty or mixed/dense at each mammographic screening as an individual assessment, which was not by default equal to previous assessments. Previous screening mammograms were used to ensure good clinical practice according to international standards, to ensure correct diagnostics of clinical parenchymal changes and evaluation of potentially previous suspicious findings.

To ensure the validity of the binary outcome in mixed/dense versus fatty mammographic density, we evaluated the binary outcome against the BI-RADS distribution in a subsample of 118 women with very high accuracy. The evaluation against the BI-RADS distribution was an attempt to evaluate the validity of our outcome against an international scale of mammographic breast density, and is described in the Methods (page 7), 2nd paragraph, line 1-8: “The density coding was re-evaluated using data from a Copenhagen study on long-term breast cancer risk in women with false-positive screening test24. In total, 118 negative screening mammograms taken prior to the false-positive screening were re-evaluated. Among 31 women with fatty mammographic density, 32% (n=10) had BI-RADS code 1, 61% (n=19) BI-RADS code 2, and 7% (n=2) BI-RADS code 3 at re-evaluation. Among 87 women with mixed/dense mammographic density, 1% (n=1) had BI-RADS code 1, 31% (n=27) BI-RADS code 2, 62% (n=54) BI-RADS code 3, and 6% (n=5) BI-RADS code 4 at re-evaluation. Manual control of the 3 misclassified women revealed they were borderline cases with changed density status over time“.

Based on the evaluation study we found a high accuracy between the dichotomous outcome of our study and the BI-RADS distribution.
The sensitivity analysis, analyzing separately subgroups of women with mammographic density assessments conducted within the same screening round (period of time) of the screening program, confirmed the findings of our study of larger proportions of mixed/dense compared to fatty mammographic density in younger compared to older birth cohorts of women for all five screening rounds (1991-2001) included in our study, which is now addressed in the Results (page 12), 2nd paragraph, line 9-11: “Sensitivity analysis in subgroups of women with mammographic density assessments conducted within the same screening period confirmed the age- and birth cohort patterns found in the overall data (data not shown)”.

14% and 24% of women in Copenhagen and Funen, respectively, changed mammographic density status over time within the 10-years of follow-up in our study. The 10% difference between the two programs might be explained by a known larger proportion of women in the ages of menopausal transition 50-55 years in the Funen compared to the Copenhagen mammographic screening program. The size of the proportions of women with changing mammographic density status in Copenhagen and Funen is therefore evidence that women’s mammographic density is not by default equal to previous assessments. We find that the proportion with changing mammographic density was as expected within the 10-years of follow-up in our study, and too large to confirm the concerns of the reviewer that the radiologists’ previous assessment of each woman by default would have affected later assessments. We have addressed this issue in the Discussion (page 13), 2nd paragraph, line 4-10: “14% and 24% of women in Copenhagen and Funen, respectively, changed mammographic density status over time within the 10-years of follow-up in our study. The 10% difference between the two programs might be explained by a known larger proportion of women in the ages of menopausal transition 50-54 years in the Funen compared to the Copenhagen mammographic screening program. The proportions of women with changed mammographic density status in Copenhagen and Funen were as expected considering the 10-years follow-up period”.

In our opinion it is therefore very unlikely, that there should be such systematic inaccuracies in the dichotomous outcome that would explain the birth cohort pattern in mammographic breast density in table 2 and table 3.

3. **Concern:** In Copenhagen there was an added radiologist towards latter part of the period. It is not clear how similar the two radiologists were, nor on whether there was any attempt to keep the definitions the same over time.

**Answer:** Mammograms in Copenhagen were evaluated by consensus reading by the senior radiologists to ensure consistency in assessment practice. We have made this clearer in the Methods (page 6) and Discussion (page 13). In the Methods (page 6), 2nd paragraph, line 2-3 the sentence (first manuscript): “Independent double reading was performed by highly trained radiologists who assessed mammographic density qualitatively into fatty or mixed/dense breasts.” has been changed to (revised manuscript): “Independent double reading by consensus reading was performed by highly trained radiologists, who assessed mammographic density qualitatively into fatty or mixed/dense mammographic density”. Further, In the Discussion (page 13), 2nd paragraph, line 2-4, the sentence (first manuscript): “To ensure consistency, the analysis was restricted to a study period where the same radiologists were in charge of the density assessment” has been changed to (revised manuscript): “To ensure consistency, the analysis was restricted to a study period where the same radiologists by consensus reading were in charge of the density assessment”.

Page 3 of 10
Also, in the Methods (page 7), Data subsection, 1st paragraph, line 4-7, the sentence (first manuscript): “In Copenhagen, all mammograms were assessed by the same senior radiologist from 1991-1998, supplemented by an added senior radiologist from the autumn of 1996, who assessed all mammograms for the rest of the study period [23]” has been changed to the sentence (revised manuscript): “In Copenhagen, all mammograms were assessed by the same senior radiologist from 1991-1998, supplemented by an added senior radiologist from the autumn of 1996, who by consensus reading assessed all mammograms for the rest of the study period [23].”

4. **Concern**: Copenhagen coverage decreased from 70 to 61% during study period. Was it predominantly higher or lower socioeconomic status women that stopped attending? Would the ones that dropped out be expected to have lower SES in general (higher BMI), and thus could it have been the low mammographic density women that dropped out, which compounded the problem with this design?

**Answer**: We cannot evaluate effects of BMI on screening adherence in the current data, since we do not have information available on women’s body mass index, and we have not been able to find published studies with relative risk estimates for the importance of level of BMI on screening participation. However, effects of educational level and socioeconomic status on screening adherence were evaluated in the Copenhagen and Funen mammography screening databases in three published studies by My von Euler-Chelpin and colleagues entitled: 1) Socioeconomic status as determinant for participation in mammography screening: assessing the difference between using women’s own versus their partner’s (2010), 2) Does educational level determine screening participation? (2008), 3) Socio-demographic determinants of participation in mammography screening (2008). With reference to these three papers, we have in the manuscript addressed the reviewer’s concern about potential socioeconomic differences in screening participation. The decreased coverage in Copenhagen from 70 to 61% during 1991-2001 was mainly due to the urban-rural gradient as discussed by My von Euler-Chelpin and colleagues in the paper “Socio-demographic determinants of participation in mammography screening” (2008). The results of this article was, that the urban-rural gradient prevailed even after adjustment for differences in population characteristics, and that women’s non-participation was determined by structural factors of the mammography screening program such as easy access to mammographic screening, and not only factors of the individual woman, because the strong urbanity factor in screening participation prevailed after taking account of demographic and health care use variables. It was found, that screening participation did not increase with increasing educational level, but that the association between educational level and non-participation was U-shaped. Several studies on educational level and screening participation have found similar results as discussed by My von Euler-Chelpin et al. in the above mentioned article from 2008.

We have in the Discussion (page 13) in the 2nd paragraph, line 12-14, added the sentence: “The association between non-participation and women’s educational level or socio-economic status was found to be U-shaped with a strong urban-rural gradient in non-participation”.

5. **Concern**: The authors do not discuss the linear versus non-linear components of these models displayed in Table 4. Some more space could have been allocated to modeling and
interpreting the non-linear parts of these various effects. In particular, it looks as if it is particular the younger cohorts that are different.

**Answer:** To describe more clearly the modeling of Table 4, we have moved the sentence: “As age, period, and birth cohort are correlated, only two of these variables were included in each regression model to avoid multicollinearity.” from the Results section in the first manuscript (page 12), 1st paragraph, line 2-3, to the Methods section, Statistical analysis, (page 9, revised manuscript), 1st paragraph, line 2-4.

Further, we address the concerns about the linear versus non-linear effects of the modeling in ad 6b.

6. **Concern:** There is no explanation on how the modeling in Table 4 are done:

a. **How were the adjustments made?**

   **Answer:** In the Methods, Statistical analysis, (page 8, revised manuscript), 2nd paragraph, line 4-9, we have clearly added that the variables are included as categorical variables. The modeling of the variables in the analysis is now described by the following sentence:”Age, period, and birth cohort were analysed as categorical variables. Age was classified into two-year age groups (50-51, 52-53….70-71). Period was defined by the two-year invitation rounds for Copenhagen (April 1991 to April 1993,…April 1999 to April 2001) and for Funen (January 1994 to December 1995…January 2000 to December 2001) to have periods of equal length. Birth cohort was defined in the diagonals by the linear relation birth cohort = period-age for Copenhagen (1921-1922….1949-1950) and Funen (1924-1925….1948-1949).”

b. **With linear terms? Categorical terms?**

   **Answer:** We find that analyzing data in two-year intervals will most finely explore the time trends in our data since this was possible with the large scale data we had available. Linear terms in the statistical analysis would not contribute with further information in the current data and would only give us a single point estimate in the odds ratio per unit increase in birth cohort. We find it important to have a reference category for the birth cohort to be able to quantify the risk in the younger birth cohorts compared to the referenced older birth cohort. Since there are no previous analysis done on birth cohort effects and mammographic density, we did not find it appropriate to assume that a single parameter would describe the association appropriately across all birth cohorts or age groups.

c. **What happens if the age effects are modeled only with the linear cohort effects? This analysis should be done and discussed. How did various age adjustments affect the birth cohort results? Only as a linear term? Why would this be adequate given the menopause issue? The authors should discuss this and the inherent assumptions, results and challenges with various adjustments.**

   **Answer:** We hope that the reviewers will find that our explanation above (ad 6b) adequately explains why linear birth cohort effects would not be appropriate considering the purpose of our study and the distribution of the data. We find that it is relevant to
have a reference category and that we cannot assume a linear association explained by one point estimate. From the changes in the manuscript described in ad 6a, we hope that it seems clearer to the reader how the variables were modeled in the statistical analysis.

Adjustment between birth cohort and age as continuous variables did not affect the results of our study, which we have added in the Results (page 12), 2nd paragraph, line 6-7: “Mutually linear adjustments did not alter the results in the multivariate analysis (data not shown).”

7. **Concern:** The part on perinatal exposures having an effect should be omitted or modified. This evidence is tenuous at best, and even if there are such effects, it is not clear how strong they might be or if they are even close to effects of other lifestyle factors. The focus in the introduction on whether these (of all) factors have changed over time and caused the observed changes is therefore somewhat strange. This should be rewritten, and the abstract, discussion and conclusion should be modified accordingly. The authors could alter it to state that changes throughout a woman’s life can have caused any cohort effects, to better fit with what we currently know about mammographic density.

**Answer:** We have modified the manuscript according to the concern as the purpose of the content was on line with the reviewer’s remarks. In the Background (page 5, first manuscript), 2nd paragraph, line 4-5, we have erased the sentence “If prenatal exposures have changed over time then the proportion of women with dense breasts would be expected to vary across birth cohorts”. Instead we have in the Background (page 5, revised manuscript), 2nd paragraph, line 1-2, added the sentence: “Suggestive evidence exists of an association between perinatal exposures and adult mammographic density [11-13]”.

We have as suggested by the reviewer in the Background (page 5, revised manuscript), 2nd paragraph, line 4-6, added the sentence “Women’s mammographic density might vary across birth cohorts by changes in women’s exposures over time”.

Further in the Abstract under Conclusion (page 3), we have changed the sentence in the first manuscript: “…suggesting that postmenopausal mammographic density is linked to exposures accumulated earlier in life.” to the sentence in the revised manuscript: “…suggesting that postmenopausal mammographic density could be linked to changing exposures accumulated over time in women’s lives”.

8. **Concern:** Breast density is an unfortunate term that implies something with the consistency of the breast, and is often misunderstood by women and investigators outside of the field. Please use mammographic density throughout.

**Answer:** We agree and have made the requested changes, now using the term mammographic density throughout the revised manuscript.

**MINOR:**

1. **Concern:** In the description of the Copenhagen reevaluation study – when N is less than 100 – it would probably be better with both % and N.
Answer: In the Methods (page 7), 2nd paragraph, line 3-7 we have added the absolute numbers in the parenthesis to the relative numbers as requested by the reviewer.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.
Declaration of competing interests: I declare that I have no competing interests

--oo0oo--

Point-by-point response to review comments by Andre LM Verbeek:

1. Concern: Is the question posed by the authors well defined?

The research question to be addressed are the trends in breast density as observed on screening mammograms in birth cohorts as causal approximation for the 'historical increase in breast cancer incidence.' The authors are taking hormone use into consideration. The research is original in the sense that there are no previous longitudinal studies on mammographic density across different birth cohorts available.

Answer: We agree with the reviewer

2. Concern: Are the methods appropriate and well described?

Authors did not use a reproducible measure for breast density, rather a crude type of measure for this type of study. Very high percentages of dense patterns were observed., but based on the way it appears to be associated with the BI-RADS density classification, findings seem to be valid.

Answer: The outcome measure was a dichotomized outcome based on assessments from highly experienced senior radiologists in charge of the two independent mammographic screening programs, who assessed the mammograms by consensus reading. The outcome was validated against the BI-RADS distribution with high accuracy as described in ad 2 in our reply to reviewer Giske Ursin. Also, in the study by Anne Helene Olsen et al entitled “Breast density and outcome of mammography screening: a cohort study” (British Journal of Cancer 2009) conducted on the same data from the Copenhagen mammography screening program, an increased risk of breast cancer RR 2.45 (2.14-2.81) was found in women with mixed/dense compared with fatty mammographic density, which is in agreement with existing literature.

3. Concern: Are the data sound?
The data are collected from population-based large scale screening programmes, and mammographic density scored according to BI-RADS methodology.

Answer: We agree with the reviewer

4. Concern: Does the manuscript adhere to the relevant standards for reporting and data deposition?
Yes, adequately done.
**Answer:** We agree with the reviewer

**5. Concern:** Are the discussion and conclusions well balanced and adequately supported by the data? Although references to other well-known studies are taken into account, some information/interpretation on the background and relevance of the findings are missing to fully appreciate the differences across birth cohorts potentially caused by early-life exposures. The focus now is on the changes with age within birth cohorts.

**Answer:** We acknowledge that we could have elaborated more on a possible interpretation of the results of the study. We have therefore expanded the Discussion (page 15, 2nd paragraph, line 4-8 to page 16, 1st paragraph, line 1-3): “This difference may be explained by a higher background burden of hormone related exposures in younger compared with older birth cohorts of women. An increased burden of cohort borne breast cancer risk factors in younger compared with older birth cohorts of Danish women has previously been linked to the increasing breast cancer incidence across birth cohorts of women over time in Denmark [5]. Changes over time in native Danish women’s reproductive pattern, increasing obesity, and sedentary lifestyle could be proxies for increased sex- and growth hormone levels potentially influencing changes in women’s mammographic density [4,6].”

**6. Concern:** Are limitations of the work clearly stated? One of the main strengths of this study is that authors were able to look into birth cohorts. The fact that density was assessed by the same radiologists could be seen as a strength as well. The non-reproducible measure for density and the (somewhat) short period of follow-up (10 years) can be considered as weaknesses. Strictly spoken, the study design is not longitudinal (cross-sectional not individual data).

**Answer:** We acknowledge that we analyze the data cross-sectionally in Table 2 and Table 3, where the analysis is conducted within defined screening periods in time. We defined the study as longitudinal because of the data structure with repeated mammographic longitudinal data, and because the analysis of the data in Table 4 is best described as a longitudinal regression in which the variables are considered to be associated with a sequence of points in time from the same woman, and where we therefore take account of the correlation in measurements by use of the GEE modeling. Based on the structure of the data and statistical analysis by use of the GEE modeling, we found that the study design was best described as longitudinal even though we acknowledge that the study is also analyzed cross-sectionally in Table 2 and Table 3.

We have not changed the study design to cross-sectional based on the above considerations, however, if decided in further review, we accept a change of the manuscript title to cross-sectional.

We validated our dichotomous outcome in mixed/dense and fatty mammographic density against the BI-RADS distribution with very high accuracy among 118 women in our cohort. To ensure that the outcome was not affected by structural changes in the mammography screening program, we decided in close collaboration with the radiographers in charge of the two mammography screening programs to restrict the period of follow-up between 1991-2001 even though we had mammographic data available until 2008. We find that this restriction is clearly stated in the revised manuscript in the Methods, Data subsection, page 7, 1st paragraph, line 1-4: “Data were collected on women participating in any of the first five invitation rounds in Copenhagen from April 1991 to March 2001, and from Funen any of the first four invitation rounds from November 1993 to
December 2001. The analysis was restricted to this period to ensure consistency in the assessment of mammographic density.”, and in the Discussion (page 14), 1st paragraph, line 3-5: “We therefore restricted the analysis to the period 1991-2001 where no systematic changes in screening practice or mammographic density assessment took place in the two programs that could have caused the findings attributed to the birth cohort effects.”

7. Concern: Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

Yes.

Answer: We agree with the reviewer and have modified the introduction about perinatal risk factors as requested by reviewer Giske Ursin.

8. Concern: Do the title and abstract accurately convey what has been found?

Title and abstract are covering the work done.

Answer: We agree with the reviewer and are willing to change the title of the manuscript to cross-sectional study instead of longitudinal study if decided by the editorial board based on our reply in ad 6.

9. Concern: Is the writing acceptable?

Quality in general is quite good. Some small revisions are needed (also in the references). Table 2 and 3 are perhaps not that easily comprehensible (mainly the birth cohorts). Nevertheless, a nicely conducted and written study, that published may prompt the reading audience to replicate the findings based on larger observation periods and individual data

Answer: We find that the design of Table 2 and Table 3 most efficiently summarize the results of the three-dimensionality of the age-period-birth cohort analysis. We have considered a revision of the in-text description of the birth cohort effects in Table 2 and Table 3. The birth cohort effect in Table 2 and Table 3 is described in the Methods, Statistical Analysis, (page 8), 2nd paragraph, line 8-9: “Birth cohort was defined in the diagonals by the linear relation birth cohort = period-age for Copenhagen (1921-1922….1949-1950) and Funen (1924-1925….1948-1949). Further, an example of the birth cohort effects is given in the Results (page 10), 3rd paragraph, line 1-7: “The proportions of women with mixed/dense mammographic density for each birth cohort in Copenhagen are found in the diagonals of Table 2, where each step-wise diagonal chain of squares represents a birth cohort. In Table 2, women born in 1940-1942 were 50-51 years of age in 1991-1993, 52-53 years in 1993-1995, 54-55 years in 1995-1997, 56-57 years in 1997-1999, and 58-59 years in 1999-2001. This birth cohort had a relatively constant probability of mixed/dense mammographic density at 64% in 1991-1993, 62% in 1993-1995, 61% in 1995-1997, 61% in 1997-1999, and 63% in 1999-2001...”

We have revised the conclusions of the study to clearly state the need for future studies on the association between age, period, and birth cohort effects and mammographic density by adding the sentence in the end of the Discussion (page 16), 2nd paragraph, line 1-2: “The results of the current study warrant future replication in larger individual level data with longer follow-up.”

--oo0oo--
The results of our study based on unique Danish data from two independent large scale mammographic screening programs with individual level birth cohort and mammographic density data are novel and have not been investigated in previous studies. Future studies are warranted to replicate the findings of this study.

We are very grateful for your interesting comments on our manuscript. We hope that you will find our response to the reviewers’ comments satisfactory, and that you will find our manuscript relevant for publication.

Correspondence should be addressed to Sophie Sell Hellmann, Department of Public Health, University of Copenhagen, Denmark.

Yours sincerely,

Sophie Sell Hellmann, MSc
Department of Public Health
University of Copenhagen,
Øster Farimagsgade 5,
Copenhagen K, DK-1014, Denmark.
Tel: +45 35327733  Fax: +4535327383, E-mail: sohe@sund.ku.dk