Examining Non-Participation to the Maternal Follow-up within the Danish National Birth Cohort

Mette Bliddal, Zeyan Liew, Anton Pottegård, Helene Kirkegaard, Jørn Olsen, and Ellen A. Nohr

Correspondence to Dr. Mette Bliddal, Odense Patient Data Explorative Network (OPEN), J.B. Winsløw Vej 9a, 3, 5000 Odense C, Denmark (e-mail: mette.bliddal@rsyd.dk)

Author affiliations: Odense Patient Data Explorative Network (OPEN), Department of Clinical Research, University of Southern Denmark and Odense University Hospital, Odense, Denmark (Mette Bliddal); Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA), California, USA (Zeyan Liew); Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark, Odense, Denmark (Anton Pottegård); Research Unit for Gynaecology and Obstetrics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark (Helene Kirkegaard, and Ellen A. Nohr); Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark (Jørn Olsen).

The Danish National Research Foundation established the Danish Epidemiology Science Centre, which initiated and created the Danish National Birth Cohort. The cohort is a result of a major grant from this Foundation. Additional support for the Danish National Birth Cohort was obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation, and the Augustinus Foundation. Finally, support for the Maternal Follow-up was granted by The Danish Council for Independent Research (0602-01042B). No funding was granted specifically for this work. Zeyan Liew was supported by the NIH/NIEHS career development award (K99ES026729).

Conflict of interest: none declared.

Running head: Non-participation in a follow-up of a birth cohort
Abbreviations

BMI = Body Mass Index

CI = Confidence Intervals

DNBC = the Danish National Birth Cohort

ICD10 = International Classification of Diseases, version 10

IPW = Inverse Probability Weighting
Abstract

A follow-up questionnaire on maternal health has been distributed within the Danish National Birth Cohort (established 1996-2002) 14 years after the index birth. Answers were obtained for 41,466 of 78,010 (53.2%) eligible mothers. To ensure the appropriate use of these data, the possibility of selection bias due to non-participation must be evaluated. We estimated four selected exposure-outcome associations (pre-pregnancy weight – depression; exercise – degenerative musculoskeletal conditions; smoking – heart disease; and alcohol consumption – breast cancer). We adjusted for several factors associated with participation and applied inverse probability weighting. To estimate the degree of selection bias, we calculated relative odds ratios between the baseline cohort and the subset participating in the follow-up. Participating women were in general healthier, of better social status, and older than the baseline cohort. However, selection bias was limited in the chosen scenarios with ratios of the odds ratios ranging from -14% to +5% after adjustment for age, parity, social status, and if not the exposure, pre-pregnancy body mass index, exercise, smoking, and alcohol consumption. Applying inverse probability weighting technique did not further reduce bias. In conclusion, while participants differed somewhat from the baseline cohort, selection bias was limited after factors that associated with participation status were accounted for.

Key words
Birth cohort; inverse probability weighting; longitudinal cohort studies; non-participation; selection bias
The Danish National Birth Cohort (DNBC) is one of the largest birth cohorts in the world with initial participation of more than 100,000 pregnancies (See www.dnbc.dk) (1). While most birth cohorts have focused on the outcomes of the children, the DNBC also grants the opportunity to study women’s health after giving birth with detailed information collected during their pregnancies and early motherhood. From the inception of the DNBC, regular prospective follow-ups of the children were planned and until now, several have been conducted. From December 2013 to December 2014, on an average 14 years after childbirth, the first follow-up focusing on the health of the mothers, the Maternal Follow-up, was conducted. With time passing, the incentive to participate in the DNBC may vary, due to changes in life situations, social or health conditions as well as lifestyle. Differences in participation likely correlate with exposures and health outcomes under study and, thus, selection bias may occur (1–3), making it imperative to evaluate the possibility of bias due to non-participation. Fortunately, the extensive nationwide health registers in Denmark allow formal assessment of the influence of potential selection as they hold information also about DNBC participants who chose not to participate in the Maternal Follow-up.

In this study, we aimed to describe selection on maternal characteristics in the Maternal Follow-up within the DNBC. We selected four maternal exposure - outcome pairs and examined the direction and magnitude of potential selection bias due to non-participation. Further, we examined if using the Inverse Probability Weighting (IPW) technique could reduce the possible influence of selection bias.
METHODS

Danish National Birth Cohort

This study is based on the DNBC. From 1996-2002, a total of 100,421 pregnancies among 91,389 women in Denmark were recruited to the cohort corresponding to approximately 30% of all pregnancies in Denmark during the study recruitment period (4) (Figure 1). The baseline information on demographics and lifestyle was collected using computer-assisted telephone interviews (1). The Maternal Follow-up was conducted during 2013 and 2014 with the majority of women being in their forties, having ended childbearing, but not yet reached menopause. A questionnaire was developed by researchers familiar with the DNBC and relevant experts within mental and physical health, occupational health and lifestyle, reproduction, and strains in motherhood. From January 2013 to November 2014, an invitation to fill in a web-based questionnaire was sent to all eligible mothers in the DNBC. Mothers who in connection with previous follow-ups regarding the children provided an e-mail address were contacted by e-mail (54%), the rest were contacted by mail (46%). The women were reminded twice by e-mail or letter (dependent on group) with a fortnight interval had they not respond to the initial invitation. From the initial cohort, women were excluded if they had only unsuccessful pregnancies in the cohort (n=3,152), only unknown outcomes in the cohort (n=58), had emigrated (n=43) or died during pregnancy (n=3). This left us with a sample of 88,136 mothers. If a woman had more than one pregnancy in the cohort leading to a liveborn child, the first pregnancy served as index pregnancy.

A total of 5,567 mothers were not eligible the Maternal Follow-up due to death (n=449), emigration, or withdrawal of consent to participate in future data collections. In all, 82,569 mothers were invited (46 % by mailed invitation, 54% invited by e-mail) and 43,641 completed the Maternal Follow-up questionnaire. The overall response rate was 52.9% (26% by mailed invitation, 68% by e-mail invitation). The questionnaire was completed at a
median of 13.8 years after childbirth (IQR 12.7-14.6 years) and the data collection concluded in February 2015. For this study, we excluded mothers who did not answer the first interview in the DNBC (n=4,559), leaving a study population of 79,010 mothers (baseline cohort) of whom 53.2% (n=41,466) responded to the Maternal Follow-up (Maternal Follow-up subset).

Other data sources

By use of the unique individual person identification number assigned to all Danish individuals (5), the baseline cohort including mothers who did not answer the Maternal Follow-up was linked to the Danish National Patient Register (6), the Danish Civil Registration System (5), and the National Medical Birth Registry (7,8). Due to the individual linkage and the Danish Health registers, the linkage rate was virtually complete (9). The Danish National Patient Registry contains information on all inpatient contacts from 1977 and outpatient contacts and emergency room events from 1995 in Danish hospitals (6). The diagnostic codes used in the Patient registry are classified according to the International Classification of Diseases version 8 (1977 - 1993) and version 10 (since 1994). These data allowed us to identify diseases diagnosed in hospital settings for each woman in the cohort. The Danish Civil Registration System enabled us to retrieve information on death and migration (5). Data from the National Medical Birth Registry informed on parity and birth outcomes (7).

Exposure-outcome associations

We studied if selection by non-participation in the follow-up affected relative risk estimates by comparing four different exposure-outcome associations in the baseline cohort and Maternal Follow-up subset. The four selected associations were: (I) pre-pregnancy body mass index (BMI) – anxiety and depression disorders (10–12), (II) leisure time exercise in pregnancy – degenerative musculoskeletal disorders (13–15), (III) smoking in pregnancy – stroke/ischemic heart disease (16), and (IV) alcohol consumption prior to pregnancy – breast
cancer (17,18). These associations were chosen because they have previously been subjects of interest in the literature, and each of the selected exposures and outcomes could influence participation in the Maternal Follow-up. Further, different types of diseases may affect selection differently and the outcomes chosen cover four major areas of disease that are all relevant to study in women in midlife and sufficiently common to allow meaningful assessment in this still fairly young population. Finally, the chosen diseases represent both common and rare diseases, which again may affect the impact of selection differently.

Study variables

All exposure information was self-reported from the first interview at a median 17 weeks of gestation in the DNBC. Pre-pregnancy BMI was calculated based on pre-pregnancy weight (kg) and height (m) as kg/m² and categorized according to the World Health Organization (19) definition as underweight (BMI <18.5 kg/m²), normal-weight (18.5-24.99 kg/m²), overweight (25-29.99 kg/m²), and obese (≥30 kg/m²). Leisure time exercise in pregnancy was categorized as no exercise, 1-180 minutes per week, or >180 minutes per week. Smoking in pregnancy was categorized according to smoking status at first interview (No smoking, smoking cessation in early pregnancy, or smoking). Alcohol consumption before pregnancy was categorized as no alcohol intake/less than 1 drink per week, 1-4 drinks per week, or 5+ drinks per week.

Outcome data was obtained from the Danish registers and all outcome variables were dichotomous. Anxiety and depression disorders were defined by a diagnosis of one or more of the following International Classification of Diseases version 10 (ICD10) codes: F30-39 “All affective mental disorders” or F40-48 “All nervous and stress-related disorders/disorders with physical symptoms”. Degenerative musculoskeletal disorders were identified as any ICD10 code M related to degenerative musculoskeletal conditions. Cardiovascular diseases were identified as diagnosis of any ischemic heart disease (ICD10 I20-21 or I24-25), or
stroke (ICD10 I60-I64). Breast cancer cases were identified by ICD10 codes C50. Before focusing on each outcome, we excluded women with any record of the actual outcome before the day of conception by either ICD8 or ICD10 codes (see Web Table 1 for ICD8 and ICD10 specification for definition of exclusions and outcomes).

Other variables were defined at the index pregnancy and included age (continuous), parity (0, 1, ≥2 children), and social status defined by type of job or type of education if still attending school (low, middle, high) (20). Finally, parity after the index child birth until June 2011 (0, 1, ≥2 children) was presented.

**Statistical methods**

Exposures and background characteristics at baseline as well as parity during follow-up and selected outcomes were described by marginal frequencies for the baseline cohort and for the Maternal Follow-up subset. We used multiple logistic regression analyses with 95% confidence intervals (CI) to estimate the odds ratios for each of the exposure-outcome pairs in the baseline cohort and in the Maternal Follow-up subset, respectively. Adjustments were made for age, parity, and social status at baseline, and if not the exposure, pre-pregnancy BMI, leisure time exercise in pregnancy, smoking in pregnancy, and alcohol consumption prior to pregnancy. We chose traditional models often applied to examine causal relations, but abstained from applying sophisticated model to maintain the focus on the bias analyses.

Adjusting for the chosen factors may block confounding and selection paths via these factors (21).

To evaluate the magnitude and direction of selection bias, we compared the distribution of the exposures, covariates, and the outcomes in the baseline cohort and the Maternal Follow-up subset computing the relative differences (prevalence ratios) between the baseline cohort and the Maternal Follow-up subset (22). We also obtained selection bias estimates by use of
the ratio of the odds ratios for each of the exposure-outcome pairs by dividing adjusted odds ratio
Mater - Follow-up subset with adjusted odds ratioBaseline cohort (23). Bias estimates, i.e. relative odds ratio values, below 1 indicate an under-estimation of the association in the Maternal Follow-up subset and conversely, estimates above 1 indicate an over-estimation. Adjusted relative odds ratio was used for the evaluation of selection bias as crude relative odds ratio would be a mix of both selection and confounding bias. To calculate confidence intervals on the prevalence ratios and relative odds ratios between the two dependent study populations (the Maternal Follow-up participants inherently being a subset of the baseline cohort), we used an equation method presented by Nohr et al (22) as:

\[
se(\hat{\theta}_{\text{sub}} - \hat{\theta}_{\text{tot}}) = \sqrt{se(\hat{\theta}_{\text{sub}})^2 - se(\hat{\theta}_{\text{tot}})^2}
\]

which was found in a simulation study to be valid, especially when the expected bias is modest (22). The thetas represent odds ratios.

In addition, we performed weighted regression analysis using IPW (24) by estimating the probability of participation in the Maternal Follow-up based on the women’s information collected at baseline to account for potential selection bias in analyses of women participating in the Maternal Follow-up. A participating woman is thus assigned a weight so that she account not only for herself in the analyses but also for those who in characteristics are similar to her but did not answer the follow-up (24). We first used logistic regression to predict the odds of participation in the Maternal Follow-up using a wide range of baseline factors, i.e., all four exposure variables of interest, the selected covariates, and the number of children (1 or 2+) enrolled in the DNBC. These factors were all associated with participation both univariate and mutually adjusted (See Web Table 2). An IPW variable for each woman was then computed and included in the regression model for the Maternal Follow-up subset
with a robust error estimator to obtain the 95% CI. All analyses were performed using
STATA 13.0 (StataCorp, College Station, TX, USA).

**Ethics**

Participants in the DNBC initially gave written consent to participate in the longitudinal
collection of data and allowed use of their data for research in maternal and child health.
Permission to use data was granted by the Danish Data Protection Agency (2008-58-0035).

**RESULTS**

Characteristics of the baseline cohort and the participants in the Maternal Follow-up are
presented in Table 1. The Maternal Follow-up subset differed from the baseline cohort as
mothers who were older at baseline or had more than one child in the DNBC were over-
represented in the Maternal Follow-up. Also, they were healthier in regard to weight and
exercise than all participants at baseline. Some subgroups were less likely to participate in the
Maternal Follow-up. Thus, young mothers were under-represented with a prevalence ratio of
0.55 (95% CI: 0.50, 0.61), women of low social status with a prevalence ratio of 0.74 (95%
CI: 0.71, 0.76), and women who smoked in pregnancy with a prevalence ratio of 0.75 (95%
CI: 0.74, 0.77) (Table 1). Parity during follow-up was similar in the baseline cohort and the
sub-sample.

Women with incident disease during the follow-up period tended to be less likely to
participate in the Maternal Follow-up. Under-representation was most pronounced for women
with depression/anxiety and for women with stroke/ischemic heart disease (prevalence ratio
0.80 (95% CI: 0.77, 0.82) and 0.83 (95% CI: 0.78, 0.89), respectively) (Table 2). Women
who had received a diagnosis of breast cancer were over-represented with a prevalence ratio
of 1.13 (95% CI: 1.07, 1.20).
Table 3 shows the crude and adjusted odds ratios for the baseline cohort and the Maternal Follow-up subset for each of the four exposure-outcome pairs along with the adjusted relative odds ratios comparing the adjusted odds ratios for the baseline cohort and the Maternal Follow-up subset. Associations in the chosen exposure-outcome pairs were as expected. Risk of a diagnosis of depression was higher for both underweight and overweight/obese than normal-weight women. Risk of degenerative musculoskeletal conditions was slightly increased with increasing leisure time exercise in pregnancy. Smoking in pregnancy was associated with higher risk of cardiovascular disease than non-smoking only in the baseline cohort. Finally, results indicated no association between alcohol consumption and breast cancer. Overall, selection bias was generally limited in the chosen scenarios, with relative odds ratios ranging from -0.86 to 1.05. For pre-pregnancy BMI and depression, the excess risk on underweight women was underestimated in the Maternal Follow-up subset compared to the baseline cohort (relative odds ratio 0.90 (95% CI: 0.73, 1.07)), whereas for overweight and obese women, there was a slight overestimation (relative odds ratio 1.04 (95% CI: 0.96, 1.13) and 1.05 (95% CI: 0.93, 1.20), respectively). Point estimates were slightly higher in the Maternal Follow-up subset for exercise and MSD and slightly lower for smoking and CVD compared with those in the baseline cohort. When examining smoking and cardiovascular disease, women indicating smoking cessation in the Maternal Follow-up had a divergence of adjusted odds ratio of 0.86 (95% CI: 0.65, 1.14) compared to the baseline cohort. It should be noted, however, that for all associations CIs were largely overlapping. When applying IPW to the adjusted analyses, the bias estimates were largely unchanged.
DISCUSSION

More than half of the baseline cohort in the DNBC replied to the Maternal Follow-up 14 years after childbirth. Of all invited mothers, those who chose to participate were in general older and healthier at baseline and at follow-up. The only exception was an overrepresentation of women with breast cancer, which is noteworthy for future studies of cancer in the cohort. Also, we found that maternal social status and several lifestyle factors at baseline were associated with participation status. However, in the four selected exposure-outcome associations that we evaluated, the possible influence of selection bias in the effect estimates was limited after we adjusted for factors that may influence selection. Additionally applying IPW had virtually no impact on the bias estimates.

Although non-participation in the Maternal Follow-up was as high as 47.1%, it was not substantially larger than the 39.9% non-participation in the 7-year follow-up focusing on the health of the children within the same cohort (21), and 40.2% of mothers participated in both follow-ups indicating that women once willing to participate in a follow-up also adhere to subsequent follow-ups. This is also supported by the fact that women who once had informed their email-address had a much higher response rate (68%) than women invited by mailed letter (26%). That the most healthy and well educated mothers are more willing to participate in follow-ups is consistent with findings from other large longitudinal cohorts of younger women (25–27).

Estimates of associations in the four exposure-outcome pairs were as expected except for alcohol consumption prior to pregnancy and breast cancer where we found no association in contrast to others who observed that alcohol intake was significantly related to breast cancer risk (18,28). This may be explained by the fact that that the alcohol consumers in the Maternal Follow-up are “healthy alcohol consumers”, i.e., most women in this category having a glass of wine at dinner several times a week but with few heavy drinkers. Our upper
category of alcohol consumers are set at >5 drinks per week – a fairly low cut-off. This no
association of an increased risk of breast cancer in light users of alcohol is supported by
others who do not find an association between alcohol intake of up to 6 drinks per week and
breast cancer in pre-menopausal women (29). Even though we have adjusted for other
lifestyle factors, the finding may also be due to residual confounding.

Despite differences in prevalence estimates and risk of disease between the baseline cohort
and the subset, the chosen associations between exposure and disease were only slightly
affected by selection after adjustment for few factors associated with participation. This
corresponds with other studies examining effects of selection on measures of association in
longitudinal studies (21,25,30). In the DNBC cohort, Greene et al found selection bias to be
small with relative ratios of -10% - +8% when examining childhood outcomes 7 years
postpartum (21). Also in other large longitudinal cohort studies such as the Norwegian
Mother and Child Cohort Study (30) and the Avon Longitudinal Study of Parents And
Children (22) drop-out or self-selection was systematic yet it only biased selected exposure –
outcome associations marginally. We found little indication of selection bias in the four
associations evaluated, and we cannot rule out that simple stochastic variation drives some or
all of our findings. As the sample size and number of cases were smaller in the Maternal
Follow-up subset, the CIs were wider and the variance of estimates increased. Hence, odds
ratios were not different but more accurate in the larger baseline cohort. Interestingly, the
largest difference in odds ratios was found for women indicating smoking cessation in
pregnancy and stroke/ischemic heart disease. Smoking is a time-varying factor and it is likely
that some mothers start to smoke again after their pregnancy, and our finding may indicate
that these women were less likely to participate in the Maternal Follow-up.

We present the adjusted relative odds ratios well aware that by adjusting for potential
confounders we also remove some selection bias when the covariates are also associated with
selection, which is the case here. Adjusting for a sufficient set of measured covariates that influence selection could appropriately close the open collider path that would otherwise induce spurious association between the exposure and the outcome (24,31). In our chosen examples, we included a few important covariates in the regression model, and only minor selection bias was present. Regression adjustment is convenient and easy to implement. Additionally applying the IPW technique did not reduce the bias estimates notably. This again indicates that the IPW may just slightly add to the regression model in terms of addressing possible direct influence of the exposure on the selection that cannot be removed by adjusting for other covariates (24). If, however, it requires a large number of covariates to predict selection, the regression model becomes ineffective having to include all these independent variables in the outcome regression and using IPW may then be an advantage (24).

The principal strengths of the study are the large sample size and almost full information on covariates collected at baseline in the baseline cohort. Further, diseases were register-based and nearly complete for the baseline cohort, which allowed us to estimate the effect of non-participation related to both exposures and outcomes (6).

A limitation was that some of the outcomes of interest were rare leading to uncertainty in bias estimates and we cannot rule out some selection bias. We chose a limited set of measured covariates to predict participation, and clearly, many of these factors may vary over time. However, looking at for instance parity, childbirth during follow-up did not seem to be associated with participation, and our selected baseline covariates appeared to be sufficient to predict participation. We recognize that time-varying factors’ status at the time of the Maternal Follow-up would likely be more related to MF participation than only using the status at baseline. However, we did not have such information on for instance social status for
the full cohort but only among those who participated in the Maternal Follow-up study and provided this information.

Further, we expect any uncontrolled confounding might affect the estimates in the same direction in the full cohort and the MF cohort, allowing us to tease apart the bias due to selection effects. The associations investigated involve potential fatale outcomes (e.g. breast cancer or stroke) but only 449 of the baseline cohort members were not eligible for the MF study due to death (0.51%); thus potential influence from survival bias is likely minimal. Notably, we studied four exposure-outcome associations and found no evidence of strong selection bias. While generally reassuring, we cannot exclude for other associations that selection bias may differ substantially from the -14 to +5% range found in our analysis. Future studies using data from the Maternal Follow-up should always reflect this possibility. Further, residual selection bias may be present due to factors not accounted for both when selecting covariates to control for as well as performing IPW (24).

Conclusion

Several exposure and outcome factors that we evaluated appear to be associated with the participation in the Maternal Follow-up within the DNBC, hence, mothers with favorable baseline social and lifestyle factors were the most likely to adhere to long-term participation. Reassuringly, influence of selection bias in the exposure-outcome effect estimates was limited after factors that affect participation were accounted for in the analysis, and applying IPW techniques did not decrease this bias any further. Our findings add to previous literature that suggests that despite a systematic non-participation according to baseline characteristics in large population-based birth cohorts, the resultant selection bias is often relatively small if these factors can be accounted for in analysis. Our results may inform bias analyses for longitudinal studies of women’s health that are prone to selective participation in follow-ups.
Acknowledgment

Authors affiliations

Odense Patient Data Explorative Network (OPEN), Department of Clinical Research, University of Southern Denmark and Odense University Hospital, Odense, Denmark (Mette Bliddal), Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA), California, USA (Zeyan Liew), Clinical Pharmacology and Pharmacy, Department of Public Health, University of Southern Denmark, Odense, Denmark (Anton Pottegård), Research Unit for Gynaecology and Obstetrics, Department of Clinical Research, University of Southern Denmark, Odense, Denmark (Helene Kirkegaard), Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark (Jørn Olsen), Research Unit for Gynaecology and Obstetrics, Odense University Hospital, and Department of Clinical Research, University of Southern Denmark, Odense, Denmark (Ellen Aagaard Nohr)

Grants and/or financial support

The Danish National Research Foundation established the Danish Epidemiology Science Centre, which initiated and created the DNBC. The cohort is a result of a major grant from this Foundation. Additional support for the DNBC was obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation, and the Augustinus Foundation. Finally, support for the Maternal Follow-up was granted by The Danish Council for Independent Research (0602-01042B).

No funding was granted specifically for this work. Zeyan Liew was supported by the NIH/NIEHS career development award (K99ES026729).

Conflict of interest

None declared.
References


Table 1. Characteristics of the Baseline Cohort and the Maternal Follow-up Subset within the DNBC 1996 – 2014

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline cohort</th>
<th>Maternal Follow-up</th>
<th>Participation</th>
<th>Prevalence ratio&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>study population (%)</td>
<td>78,010</td>
<td>41,466</td>
<td>53.2</td>
<td></td>
</tr>
<tr>
<td>age of index conception (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 20</td>
<td>843</td>
<td>247</td>
<td>29.3</td>
<td>0.55</td>
</tr>
<tr>
<td>20 – 24</td>
<td>9,655</td>
<td>4134</td>
<td>42.8</td>
<td>0.81</td>
</tr>
<tr>
<td>25 – 29</td>
<td>32,512</td>
<td>17,321</td>
<td>53.3</td>
<td>1.00</td>
</tr>
<tr>
<td>30 – 34</td>
<td>26,107</td>
<td>14,543</td>
<td>55.7</td>
<td>1.05</td>
</tr>
<tr>
<td>35 – 39</td>
<td>8,113</td>
<td>4741</td>
<td>58.4</td>
<td>1.10</td>
</tr>
<tr>
<td>40+</td>
<td>779</td>
<td>479</td>
<td>61.5</td>
<td>1.16</td>
</tr>
<tr>
<td>parity at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>39,771</td>
<td>21,423</td>
<td>53.9</td>
<td>1.01</td>
</tr>
<tr>
<td>1</td>
<td>26,376</td>
<td>13,901</td>
<td>52.7</td>
<td>0.99</td>
</tr>
<tr>
<td>≥2</td>
<td>11,862</td>
<td>6141</td>
<td>51.8</td>
<td>0.97</td>
</tr>
<tr>
<td>children during follow-up&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>35,119</td>
<td>18,843</td>
<td>53.7</td>
<td>1.01</td>
</tr>
<tr>
<td>1</td>
<td>31,622</td>
<td>16,794</td>
<td>53.1</td>
<td>1.00</td>
</tr>
<tr>
<td>≥2</td>
<td>9,993</td>
<td>5227</td>
<td>52.3</td>
<td>0.98</td>
</tr>
<tr>
<td>children enrolled in DNBC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 child</td>
<td>70,309</td>
<td>36,997</td>
<td>52.6</td>
<td>0.99</td>
</tr>
<tr>
<td>&gt;1 child</td>
<td>7,701</td>
<td>4469</td>
<td>58.0</td>
<td>1.09</td>
</tr>
<tr>
<td>index pre-pregnancy BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>underweight</td>
<td>3,420</td>
<td>1692</td>
<td>49.5</td>
<td>0.93</td>
</tr>
<tr>
<td>normal-weight</td>
<td>52,110</td>
<td>28,809</td>
<td>55.3</td>
<td>1.04</td>
</tr>
<tr>
<td>overweight</td>
<td>14,909</td>
<td>7600</td>
<td>51.0</td>
<td>0.96</td>
</tr>
<tr>
<td>obese</td>
<td>6,267</td>
<td>2745</td>
<td>43.8</td>
<td>0.82</td>
</tr>
<tr>
<td>social status at baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>low</td>
<td>6,766</td>
<td>2647</td>
<td>39.1</td>
<td>0.74</td>
</tr>
<tr>
<td>middle</td>
<td>28,391</td>
<td>13,590</td>
<td>47.9</td>
<td>0.90</td>
</tr>
<tr>
<td>high</td>
<td>39,668</td>
<td>23,872</td>
<td>60.2</td>
<td>1.13</td>
</tr>
<tr>
<td>smoking in index pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non smoking</td>
<td>57,102</td>
<td>32,290</td>
<td>56.5</td>
<td>1.06</td>
</tr>
</tbody>
</table>

<sup>a</sup> Prevalence ratio estimated as the rate of participation in the Maternal Follow-up subset divided by the rate in the Baseline cohort. 95% CI: 95% confidence interval.
### Smoking cessation

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking</td>
<td>7,603</td>
<td>9.7</td>
<td>3857</td>
<td>9.3</td>
<td>50.7</td>
<td>0.95</td>
<td>0.93</td>
<td>0.97</td>
</tr>
</tbody>
</table>

### Exercise in pregnancy (min/wk)

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>49,001</td>
<td>62.8</td>
<td>24,911</td>
<td>60.1</td>
<td>50.9</td>
<td>0.96</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td>1 – 179</td>
<td>22,779</td>
<td>29.2</td>
<td>13,060</td>
<td>31.5</td>
<td>57.3</td>
<td>1.08</td>
<td>1.07</td>
<td>1.09</td>
</tr>
<tr>
<td>180+</td>
<td>6,122</td>
<td>7.8</td>
<td>3442</td>
<td>8.3</td>
<td>56.2</td>
<td>1.06</td>
<td>1.03</td>
<td>1.08</td>
</tr>
</tbody>
</table>

### Alcohol consumption prior to index pregnancy

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
<th>Count</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>None/≤1 drink/week</td>
<td>17,480</td>
<td>22.4</td>
<td>8200</td>
<td>19.8</td>
<td>46.9</td>
<td>0.88</td>
<td>0.87</td>
<td>0.90</td>
</tr>
<tr>
<td>1 – 4 drinks/week</td>
<td>42,634</td>
<td>54.7</td>
<td>23,184</td>
<td>55.9</td>
<td>54.4</td>
<td>1.02</td>
<td>1.02</td>
<td>1.03</td>
</tr>
<tr>
<td>&gt;4 drinks/week</td>
<td>17,487</td>
<td>22.4</td>
<td>9906</td>
<td>23.9</td>
<td>56.6</td>
<td>1.07</td>
<td>1.05</td>
<td>1.08</td>
</tr>
</tbody>
</table>

**DNBC = Danish National Birth Cohort, CI = confidence intervals, BMI = body mass index, min = minutes, wk = week.**

* Ratio Maternal-Follow-up/Ratio Baseline cohort, values above 1 indicate that women with this characteristic are overrepresented in the Maternal Follow-up, values below 1 indicate underrepresentation.

b Percentages do not add up to 100 due to missing information. Missing information in baseline cohort (%): Age 1 (<0.01), parity 1 (<0.01), children enrolled in DNBC: 1 (<0.01), pre-pregnancy BMI: 1,304 (1.7), social status: 3,185 (4.1), smoking in pregnancy: 30 (0.04), exercise in pregnancy: 108 (0.1), alcohol consumption: 409 (0.5).

c Computed using equation, see text for details.

d Information on subsequent births only available until June 2011.

e Weight (kg)/height(m)$^2$. 
Table 2. Distribution of Diseases in the Baseline Cohort and the Maternal Follow-up Subset within the DNBC 1996 – 2014.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Baseline cohort</th>
<th>Maternal Follow-up</th>
<th>Participation</th>
<th>Prevalence ratio*a</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>Rate</td>
</tr>
<tr>
<td>Depression/anxiety</td>
<td>76,668</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5,458</td>
<td>7.1</td>
<td>2319</td>
<td>5.7</td>
<td>42.5</td>
</tr>
<tr>
<td>No</td>
<td>71,210</td>
<td>92.9</td>
<td>38,606</td>
<td>94.3</td>
<td>54.2</td>
</tr>
<tr>
<td>Degenerative musculo-skeletal</td>
<td>72,922</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conditions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>16,408</td>
<td>22.5</td>
<td>8367</td>
<td>21.4</td>
<td>51.0</td>
</tr>
<tr>
<td>No</td>
<td>56,514</td>
<td>77.5</td>
<td>30,659</td>
<td>78.6</td>
<td>54.3</td>
</tr>
<tr>
<td>Stroke or ischemic heart disease</td>
<td>77,902</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>981</td>
<td>1.3</td>
<td>435</td>
<td>1.1</td>
<td>44.3</td>
</tr>
<tr>
<td>No</td>
<td>76,921</td>
<td>98.7</td>
<td>40,977</td>
<td>98.9</td>
<td>53.3</td>
</tr>
<tr>
<td>Breast Cancer</td>
<td>77,987</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>757</td>
<td>1.0</td>
<td>455</td>
<td>1.1</td>
<td>60.1</td>
</tr>
<tr>
<td>No</td>
<td>77,230</td>
<td>99.0</td>
<td>40,997</td>
<td>98.9</td>
<td>53.1</td>
</tr>
</tbody>
</table>

DNBC = Danish National Birth Cohort, CI = confidence intervals.
*aRatio_maternal-follow-up/Ratio_baseline cohort; values above 1 indicate that women with this characteristic are overrepresented in the Maternal Follow-up, values below 1 indicate underrepresentation.
*bComputed using equation, see text for details.
*cCases prior to conception have been excluded.
### Table 3. Odds Ratios and Adjusted Relative Odds Ratios Comparing the Baseline Cohort and the Maternal Follow-up Subset within the Danish National Birth Cohort

<table>
<thead>
<tr>
<th>Associations</th>
<th>Crude Baseline cohort</th>
<th>Crude Maternal Follow-up</th>
<th>Adjusted Baseline cohort</th>
<th>Adjusted Maternal Follow-up with IPW ROR</th>
<th>Adjusted Maternal Follow-up with IPW ROR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
<td>OR</td>
</tr>
<tr>
<td></td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
<td>95% CI</td>
</tr>
<tr>
<td>Pre-pregnancy BMI - Depression&lt;sup&gt;3&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 18.5</td>
<td>1.43</td>
<td>1.21</td>
<td>1.19</td>
<td>1.05, 1.36</td>
<td>1.07, 1.33</td>
</tr>
<tr>
<td>18.5 – 24.9</td>
<td>1.00</td>
<td>1.00</td>
<td>Referent</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>25 – 29.9</td>
<td>1.16</td>
<td>1.22</td>
<td>1.08</td>
<td>1.00, 1.16</td>
<td>1.13, 1.26</td>
</tr>
<tr>
<td>30+</td>
<td>1.35</td>
<td>1.41</td>
<td>1.14</td>
<td>1.03, 1.26</td>
<td>1.20, 1.41</td>
</tr>
<tr>
<td>Exercise in pregnancy - Degenerative musculoskeletal conditions&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1.00</td>
<td>1.00</td>
<td>Referent</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>1 – 179 min/wk</td>
<td>0.95</td>
<td>0.99</td>
<td>1.06</td>
<td>1.02, 1.10</td>
<td>1.10, 1.16</td>
</tr>
<tr>
<td>180+ min/wk</td>
<td>0.96</td>
<td>0.99</td>
<td>1.10</td>
<td>1.03, 1.18</td>
<td>1.15, 1.26</td>
</tr>
<tr>
<td>Smoking in pregnancy - Stroke or ischemic heart disease&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No smoking</td>
<td>1.00</td>
<td>1.00</td>
<td>Referent</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>Cessation in first trimester</td>
<td>1.01</td>
<td>0.86</td>
<td>1.17</td>
<td>0.98, 1.50</td>
<td>1.01, 1.47</td>
</tr>
<tr>
<td>Smoking</td>
<td>2.30</td>
<td>2.05</td>
<td>2.16</td>
<td>1.85, 2.51</td>
<td>2.02, 2.57</td>
</tr>
<tr>
<td>Alcohol consumption prior to conception - breast cancer&lt;sup&gt;4&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>none/≤ 1 drink/wk</td>
<td>1.00</td>
<td>1.00</td>
<td>Referent</td>
<td>1.00</td>
<td>Referent</td>
</tr>
<tr>
<td>1 – 4 drinks/wk</td>
<td>1.10</td>
<td>1.07</td>
<td>0.94</td>
<td>0.77, 1.15</td>
<td>0.95, 1.23</td>
</tr>
<tr>
<td>5+ drinks/wk</td>
<td>1.39</td>
<td>1.27</td>
<td>0.99</td>
<td>0.79, 1.25</td>
<td>0.98, 1.32</td>
</tr>
</tbody>
</table>

ROR = Relative Odds Ratios, IPW = Inverse Probability Weighted. Generated based on the variables used for adjustment and numbers of children enrolled in DNBC, OR = Odds Ratio.
Adjusted for age, social status, parity, and if not the exposure, pre-pregnancy BMI, exercise in pregnancy, smoking in pregnancy, and alcohol consumption prior to conception. Computed using equation, see text for details. Weight (kg)/height (m)$^2$

Baseline population = 76,668, Maternal Follow-up = 40,925.

Baseline population = 72,922 Maternal Follow-up = 39,026.

Baseline population = 77,902, Maternal Follow-up = 41,412.

Baseline population = 77,987, Maternal Follow-up = 41,452.
Figure legend

Figure 1. Flowchart of Baseline Population in the Danish National Birth Cohort, Participants, and Non-participants in the Maternal Follow-up (1996-2014).

Supplemental Digital Content

Web Table 1. Specification of Exclusion ICD Codes, Version 8 and 10 and Outcome Definition (ICD10 codes)

Web Table 2. Crude and adjusted odds ratios for participation according to selected variables in the Maternal Follow-up within the Danish National Birth Cohort.
Enrolled in the Danish National Birth Cohort
(n = 91,389 women)

Mothers Excluded Because of:
Unsuccessful pregnancies (n = 3,149)
Unknown outcomes (n = 58)
Emigration during pregnancy (n = 43)
Death during pregnancy (n = 3)

Mothers (n = 88,136)

Mothers Excluded Because of:
Unsuccessful pregnancies (n = 3,149)
Unknown outcomes (n = 58)
Emigration during pregnancy (n = 43)
Death during pregnancy (n = 3)

Mothers Excluded Because of:
Unsuccessful pregnancies (n = 3,149)
Unknown outcomes (n = 58)
Emigration during pregnancy (n = 43)
Death during pregnancy (n = 3)

Mothers (n = 88,136)

Not eligible for Maternal Follow-up:
Death (n = 449)
Emigration or withdrawal of consent (n = 5,118)

Eligible for the Maternal Follow-up
(n = 82,569 mothers)

Mothers Who Participated in the Maternal Follow-up
(n = 43,641 mothers; 52.9%)

Mothers Lost to Follow-up in the Maternal Follow-up
(n = 38,928; 47.1% mothers)

Mothers not Answering First Pregnancy Interview in the Danish National Birth Cohort
(n = 4,559)

Mothers who Participated in the First Pregnancy Interview and the Maternal Follow-up
(n = 41,446; 53.2% mothers)

Mothers who Participated in the First Pregnancy Interview but not the Maternal Follow-up
(n = 36,544; 46.8% mothers)