Estimating Bias From Loss to Follow-up in the Danish National Birth Cohort

Naomi Greene,[a] *Sander Greenland,*[a,b] *Jørn Olsen,*[a,c] *and Ellen Aagaard Nohr*[d]

Abstract: Loss to follow-up in cohort studies may result in biased association estimates. Of 61,895 women entering the Danish National Birth Cohort and completing the first data-collection phase, 37,178 (60%) opted to be in the 7-year follow-up. Using national registry data to obtain end point information on all members of the cohort, we estimated associations in the baseline and the 7-year follow-up participant populations for 5 exposure-outcome associations: (*a*) size at birth and childhood asthma, (*b*) assisted reproductive treatment and childhood hospitalizations, (*c*) prepregnancy body mass index and childhood infections, (*d*) alcohol drinking in early pregnancy and childhood developmental disorders, and (*e*) maternal smoking in pregnancy and childhood attention-deficit hyperactivity disorder (ADHD). We estimated follow-up bias in the odds or rate ratios by calculating relative ratios. For all but one of the above analyses, the bias appeared to be small, between −10% and +8%. For maternal smoking in pregnancy and childhood ADHD, we estimated a positive bias of approximately 33% (95% bootstrap limits of −30% and +152%). The presence and magnitude of bias due to loss to follow-up depended on the nature of the factors or outcomes examined, with the most pronounced contribution in this study coming from maternal smoking. Our methods and results may inform bias analyses in future pregnancy cohort studies.

(*Epidemiology* 2011;22: 815–822)

Lifecourse cohort studies aim to follow individuals from conception into late adulthood and can thus contribute to understanding how health status may be affected by causes interacting over the span of life. Events that occur during fetal life and early childhood can be recorded nearly as they happen, reducing recall errors, improving accuracy, and making differential misclassification unlikely for later endpoints. Such studies may collect a broad spectrum of information on all subjects at baseline, providing a rich source of data for research.

Unfortunately, loss to follow-up may be substantial and can result in biased association estimates if follow-up is related to both the exposure and the outcome in a given analysis and if proper adjustment is not made.[1] This bias may occur even if losses are marginally independent of exposure and outcome; bias is not identifiable unless the relative within-cell losses across exposure-outcome categories are known.[2]

We estimated follow-up bias for selected exposure-outcome associations in a large, ongoing, lifecourse cohort study. We studied associations between constitutional, behavioral, and sociodemographic characteristics and childhood outcomes of varying severity, which we expected to produce loss from different mechanisms.

METHODS

The Danish National Birth Cohort is a nationwide cohort study that, between 1996 and 2002, recruited just over 100,000 women during early pregnancy (n = 100,419 pregnancies; fewer than 100,000 individual women, as some women had more than 1 pregnancy in the cohort).[3] Information regarding the cohort's aims, structure, and progress can be found at the study website (http://www.dnbc.dk) and in several publications listed there. Briefly, the aim was to assemble a large database of information about early life exposures (conception to early childhood) that may influence risk of disease across the lifecourse, as well as contextual information regarding lifestyle choices, socioeconomic factors, dietary intake, and emotional and mental states to aid in accounting for systematic biases. Women enrolling in the birth cohort agreed to complete 4 computer-assisted telephone interviews and a food-frequency questionnaire at 25 weeks' gestation, and to give two blood samples during pregnancy and cord blood of the newborn at birth. In addition, women agreed to be invited to participate in subsequent data collection waves throughout childhood. The children born into the cohort will be given the opportunity at their 18th birthday to continue participation.

As part of the consent process, women were assured that they were free to leave the study at any time, and they were

Submitted 11 August 2010; accepted 3 June 2011; posted 14 September 2011.
From the [a]Department of Epidemiology, School of Public Health, University of California, Los Angeles, CA; [b]Department of Statistics, University of California, Los Angeles, CA; [c]Danish Epidemiology Science Centre, Department of Epidemiology, Aarhus, Denmark; and [d]Department of Epidemiology, Institute of Public Health, Aarhus University, Aarhus, Denmark.
The Danish National Research Foundation has established the Danish Epidemiology Science Centre that initiated and created the Danish National Birth Cohort. The cohort is furthermore a result of a major grant from this Foundation. Additional support for the Danish National Birth Cohort is obtained from the Pharmacy Foundation, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Augustinus Foundation, and the Health Foundation. The DNBC 7-year follow-up is supported by the Lundbeck Foundation (195/04) and the Danish Medical Research Council (SSVF 0646).
Correspondence: Naomi Greene, Department of Epidemiology, UCLA School of Public Health, Box 951772, Los Angeles, CA 90095. E-mail: ngreene@ucla.edu.

ISSN: 1044-3983/11/2206-0815
DOI: 10.1097/EDE.0b013e31822939fd

encouraged not to enroll if they were in doubt about staying for the duration of the study. When their children turned 7 years of age, participating women were invited to complete a self-administered questionnaire. This follow-up phase concluded in June 2010 when the last children born in the birth cohort reached the age of 7. The response rate at 7 years was 60%–65%. Women who did not participate in the 7-year follow-up did not necessarily leave the birth cohort permanently, and will be invited to participate in subsequent data collections. Fewer than 0.5% have formally withdrawn from the birth cohort.

Information in several national administrative registries has been linked to cohort-study participants through Danish personal identification numbers (Central Person Register numbers).[4] From the National Medical Birth Registry, information on maternal age, pregnancy-related smoking status, birth date, sex, birth weight and length, parity, and multiple births has been extracted.[5] The National Hospital Discharge Register contains data on all hospital admissions, and (since 1995) information on outpatient and emergency room events.[6] Variables include the date and type of hospital admission and up to 20 diagnoses for each admission (both primary and secondary) according to International Classification of Diseases (ICD-10) codes. At regular intervals, the birth cohort is linked with the National Hospital Discharge Register, providing outcome status for subjects in the birth cohort. This information allowed us to estimate exposure-outcome associations for all birth cohort participants, not just those who participated in the 7-year follow-up.

The Figure illustrates the flow of study participation, study selection, and losses in our study. Of the 100,419 birth cohort-enrolled pregnancies, the first interview was conducted for 92,889 (93%) and these were considered eligible for the baseline cohort of the present study. We then excluded all spontaneous and induced abortions and stillbirths (n = 3646; 4%), all subsequent births after a woman's first live-born birth in the cohort (n = 8720; 9%), and all multiple births (n = 1899; 2%). Of the remaining 78,581 singleton live births, the baseline population for the present study consisted of the 61,895 children who were born by 28 March 2002 and were invited to the first wave of the 7-year follow-up.

Within the baseline population, we identified children whose mothers completed the 7-year follow-up (n = 37,178). Thus, there were 24,717 mothers who were invited to participate when the child was age 7 years but who did not respond (hereafter called lost to follow-up), leading to a participation rate of 60%.

We studied loss to follow-up by comparing exposure-outcome associations in the baseline population with those in the follow-up participant population. We chose 5 exposure-outcome associations, each of interest in previous literature and thought to involve different selection mechanisms: (1) Small-for-gestational-age at birth (SGA) and childhood asthma,[7] (2) assisted reproductive treatment (ART) and hospital utilization rates during childhood,[8] (3) prepregnancy body mass index (BMI) and childhood infections,[9] (4) maternal alcohol consumption and childhood disorders of psychologic development,[10] and (5) maternal

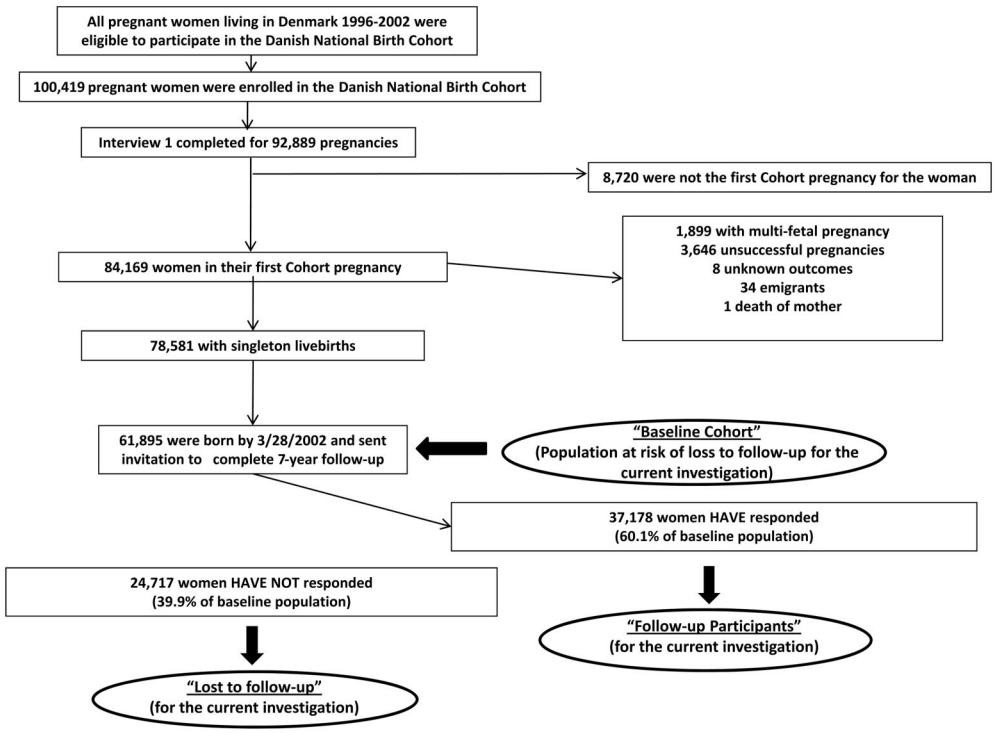

FIGURE. Flowchart of subject participation and loss to follow-up in the Danish National Birth Cohort (1996–2008).

smoking in pregnancy and childhood attention-deficit hyperactivity disorder (ADHD).[11]

Outcome data were obtained from the National Hospital Discharge Register by linkage with the birth cohort. The following variables were binary: asthma (ICD-10 codes J45, J450, J451, J458, and J459), infection (ICD-10 codes A09, G00-G03, H10, H60, H65-H66, J00-J06, J35, J10-J18, J20, K35-K37, N10-N12, N30, L00-L08, and M00-M03), psychologic developmental disorders (ICD-10 codes F80-F89), and ADHD (ICD-10 codes F90 and F98). Hospital utilization was defined as the total number of hospital encounters (inpatient, outpatient, and emergency room) listed for each child in the National Hospital Discharge Register. The number of hospital visits per child was divided by the person-time measured from birth to the end of follow-up to construct a hospital utilization rate.

The exposure variable SGA was defined as birth weight less than the 10th percentile for sex and gestational age, using the reference table suggested by Kramer et al.[12] The variable for assisted reproductive treatment was constructed using responses to a question posed in the first interview to women who planned to become pregnant (and so were not taking contraceptives). If women responded positively to the question, "Were you treated for infertility prior to this pregnancy?" they were considered exposed to ART. Prepregnancy BMI was calculated from Interview 1 (prepregnancy) weight (in kg) and height (in m) as kg/m^2. Alcohol consumption was taken from Interview 1 pertaining to drinking during the pregnancy; the possible responses were as follows: (1) no drinks, (2) less than 1 drink/week, (3) 1 or more drinks/week.

The National Medical Birth Registry collected categorized information on maternal smoking in pregnancy. Using this information and specifying category values in terms of average packs per day (assuming 20 cigarettes/pack), we coded smoking as a continuous variable according to the schedule shown in Table 1. Pack-per-day category codes were assigned before examining follow-up participation rates.

Preterm birth was categorized as very preterm if a live-born child was delivered at fewer than 224 completed gestational days (<32 weeks), preterm if delivered at 224–237 days (32–33.9 weeks), late preterm if delivered at 238–258 days (34 to 36.9 weeks), and term if 259–315 days (37–45 weeks). Socio-occupational status was based on information from the first interview and defined according to the mother's and father's most recent job (or type of education, if still in school). Those in managerial positions or attending higher education were categorized as "high," office or skilled workers and those in military service were classified as "medium," and unskilled or unemployed workers were classified as "low"; we used the highest status within the couple.[13]

TABLE 1. Coding Schedule for Smoking During Pregnancy Using Danish Medical Birth Registry Data

Category	Meaning	Assignment
0	No smoking during pregnancy	0
10	Stopped smoking during first trimester	Assigned lowest value among smokers (category 20)[a]
11	Stopped smoking after first trimester	Assigned next highest value among smokers (category 21)[b]
20	Smoked 1–5 cigarettes/day	Used mid point of 1–5[a]
21	Smoked 6–10 cigarettes/day	Used mid point of 6–10[b]
22	Smoked 11–20 cigarettes/day	Used mid point of 11–20[c]
23	Smoked >20 cigarettes/day	>1 pack/day[d]
29	Smoked, amount not specified	Average of categories 20–23 excluding missing[e]
99	Smoking status unspecified	Average of categories 0, 10, 11, 20–23, 29 excluding missing[f]
	Missing	Missing

[a]Coded as 3/20 = 0.15.
[b]Coded as 8/20 = 0.40.
[c]Coded as 15.5/20 ≈ 0.80.
[d]Coded as 30/20 = 1.5.
[e]0.45.
[f]0.10.

We tabulated the marginal frequencies comparing the baseline, follow-up, and lost-to-follow-up populations. Due to the use of registry data for the study endpoints, as well as computer-assisted telephone interview methods for baseline covariate data collection, missing information was kept to minimal levels; 0.3% was missing for smoking and for preterm birth, 0.4% for SGA, 1.7% for prepregnancy BMI, and 4.3% for socio-occupational status.

To compare the exposure-outcome associations in follow-up participants and in the baseline population, we first carried out regression analyses of each of the exposure-outcome pairs in each of these populations. To take into account that follow-up participant children were, on average, born into the cohort earlier than those lost, indicators for follow-up time (7 years, 8 years, etc) were added to all regression models except for the relationships between ART and hospitalization. For these analyses, log follow-up time was used as an offset in the Poisson regression. As our goal was not to examine the causal mechanisms of these previously studied relationships, the models were kept fairly simple with minimal covariate control, chosen with guidance from published studies.

To estimate bias due to loss to follow-up, the adjusted relative odds ratios (relative OR) for SGA-Asthma, BMI-Infection, Alcohol-Developmental Disorders, and Smoking-ADHD were calculated as follows:

$$\text{adjusted OR}_{\text{follow-up}}/\text{adjusted OR}_{\text{baseline}}$$

For ART-Hospitalization, the adjusted relative rate ratio (relative RR) was calculated as above, substituting rate ratios (RRs) for

odds ratios (ORs). These relative ratios are equivalent to selection bias factors, which are cross-products of participation.[14,15]

The association measures in the baseline and follow-up participant populations are highly dependent on each other, which complicates testing and estimation. We therefore employed a nonparametric bootstrapping method to construct confidence intervals (CIs) around each ROR. After resampling (with replacement) the baseline cohort of 61,895 5000 times, the ln(ROR) in each replicate was calculated as the difference in exposure coefficients from the baseline and follow-up participant populations. A 95% bootstrap interval was constructed around the bias-corrected ln(ROR) estimate

$$2 \cdot \ln(\mathrm{ROR})_{\mathrm{observed}} - \mathrm{mean}(\ln(\mathrm{ROR})_{\mathrm{replicates}}$$

using the standard deviation of the $\ln(\mathrm{ROR})_{\mathrm{replicates}}$ to estimate the standard error.[16]

To describe the relationship between each study covariate and participation in the 7-year follow-up, participation patterns were analyzed using logistic regression of loss to follow-up on study covariates, both unadjusted and adjusted for the other model-specific covariates. When using hospitalization count as a predictor of loss to follow-up, counts above 100 (0.15% of the cohort) were shrunk to 100 to minimize the influence of outliers. We evaluated whether there was increasing loss to follow-up with increasing smoking (in pack-days) using logistic regression of loss to follow-up on smoking (smoking in units of a 1-pack-per-day increase). We further investigated this association by computing risk ratios that compared the loss to follow-up risk in each level of smoking with loss to follow-up risk in nonsmokers.

The study was approved by the Data Inspectorate in Denmark. All analyses were carried out using SAS 9.1 (Cary, NC).

RESULTS

Table 2 presents characteristics of the baseline group, the follow-up participants, and those lost to follow-up. Follow-up participants were, on average, slightly older than those lost (63% of follow-up participants were 30 years or older vs. 58% of those lost to follow-up). Those lost were more often overweight prior to pregnancy, from the lowest socio-occupational group, smokers (and heavier smokers) during pregnancy, and with a history of prior preterm birth or small-for-gestational-age baby. In addition, these women were slightly more likely to have reported that their pregnancy was unplanned or a first birth, or to have no one but their partners to ask for help with financial problems (data not shown).

Table 3 provides the relative association estimates (ROR or RRR) comparing the odds ratios or rate ratios among follow-up participants with those in the baseline population (relative ratio = 1 if the 2 ratios are equal). For SGA-asthma and ART-hospitalization, the bootstrap limits were consistent with small positive bias away from the null. For BMI infection and alcohol-developmental disorders, the bootstrap limits were consistent with small negative bias away from the null. The smoking-ADHD ROR estimate was 1.33 (95% bootstrap limits = 0.70–2.52). In our analyses, inclusion of the follow-up time indicator had no practical impact on the results.

The logistic regression analysis of loss to follow-up on smoking resulted in an OR per 1-pack/day smoking increase of 2.15 (95% CL = 1.99–2.33), with P value for trend <0.001. The change in lost-to-follow-up proportions with each increase in smoking (in portions of packs/day) is demonstrated in Table 4, along with the risk ratio comparing each level of smokers with nonsmokers. There were insufficient numbers of children with ADHD or developmental disorders to separate the trend in loss to follow-up across smoking between affected and unaffected children. When considering asthma cases and noncases, the trend in cases was not as marked as it was in noncases (OR for a one pack/day increase in smoking was 1.52 in cases and 2.19 in noncases, $P = 0.01$ for the product term between smoking and asthma). The trend in loss to follow-up between cases and noncases of childhood infection were indistinguishable (product term $P = 0.23$).

As demonstrated in Table 5, the addition of smoking to the separate logistic regressions of loss to follow-up on SGA, asthma, hospitalization, and infection slightly reduced the magnitude of each of their coefficients, but each remained a predictor of loss to follow-up. The addition of alcohol consumption to the above regression models did not produce changes more than 0.02 in the odds ratio estimates or their 95% limits, with or without smoking.

DISCUSSION

Covariate distributions among the follow-up participants differed from those in the baseline population, and the confounder structure may well have changed over time—related, at least in part, to selection. The mothers who continued participation for at least 7 years were somewhat older, more likely to be in the highest socio-occupational group, and perhaps healthier (based on lower smoking and overweight prevalence and lower proportions of small or preterm babies). This is consistent with other reports on loss to follow-up.[17,18]

Because it is the child's mother/caregiver who is continuing participation on behalf of the child, loss to follow-up may be influenced primarily by maternal characteristics. In another large lifecourse cohort study of pregnant women and their offspring, loss to follow-up when children were 8–9 years of age was associated with lower socio-occupational group and maternal smoking,[19] as in our study.

There was a modest 8% higher SGA-Asthma association among the follow-up participants. The ART-Hospitalization associations were essentially identical in the baseline and

TABLE 2. Characteristics of Baseline, Follow-up, and Lost-to-follow-up Populations, Danish National Birth Cohort (1996–2008)

	Baseline Population (n = 61,895) No. (%)[a]	Follow-up Population (n = 37,178) No. (%)[a]	Lost-to-follow-up Population (n = 24,717) No. (%)[a]
Maternal age (years)			
<20	388 (1)	152 (<1)	236 (1)
20–24	3618 (6)	1824 (5)	1794 (7)
25–29	20,245 (33)	11,920 (32)	8325 (34)
30–34	24,533 (40)	15,105 (41)	9428 (38)
35–39	10,400 (17)	6462 (17)	3938 (16)
≥40	2711 (4)	1715 (5)	996 (4)
SGA percentile[b]			
<10th	4146 (7)	2349 (6)	1797 (7)
≥10th	57,523 (93)	34,689 (94)	22,834 (93)
Asthma			
Yes	3558 (6)	1959 (5)	1599 (7)
No	58,337 (94)	35,219 (95)	23,118 (94)
ART			
Yes	3523 (6)	2121 (6)	1402 (6)
No	58,372 (94)	35,057 (94)	23,315 (94)
Child hospitalizations			
None	148 (1)	92 (1)	56 (1)
1–20	57,267 (93)	34,548 (93)	22,719 (92)
21–40	3898 (6)	2237 (6)	1661 (7)
>40	582 (1)	301 (1)	281 (1)
BMI[c]			
≥25	16,851 (28)	9383 (25)	7468 (30)
<25	44,013 (72)	27,205 (73)	16,808 (68)
Infection			
Yes	15,741 (25)	9138 (25)	6603 (27)
No	46,154 (25)	28,040 (75)	18,114 (73)
Alcohol consumption[b]			
None	33,855 (55)	20,118 (54)	13,737 (56)
<1 drink/week	12,943 (21)	7702 (21)	5241 (21)
≥1 drink/week	15,082 (24)	9350 (25)	5732 (23)
Disorder of psychologic development			
Yes	157 (<1)	100 (<1)	57 (<1)
No	61,738 (>99)	37,078 (>99)	24,660 (>99)
Smoked during pregnancy[b]			
Did not smoke	48,743 (79)	30,174 (81)	18,569 (75)
Quit in first trimester	1064 (2)	620 (2)	444 (2)
Quit after first trimester	245 (<1)	138 (<1)	107 (<1)
Up to 5 cigarettes/day	3121 (5)	1654 (5)	1467 (6)
6–10 cigarettes/day	3561 (6)	1872 (5)	1689 (7)
11–20 cigarettes/day	2406 (4)	1134 (3)	1272 (5)
>20 cigarettes/day	255 (<1)	109 (<1)	146 (1)
Unstated amount	245 (<1)	126 (<1)	119 (1)
Unspecified status	2093 (3)	1250 (3)	843 (3)
ADHD in child			
Yes	80 (<1)	54 (<1)	26 (<1)
No	61,667 (>99)	37,032 (>99)	24,635 (>99)
Socio-occupational group[d]			
Low	5448 (9)	2877 (8)	2571 (11)
Medium	22,939 (39)	13,240 (37)	9699 (41)
High	30,849 (52)	19,611 (55)	11,238 (48)
Preterm birth[b]			
Very preterm	280 (1)	163 (<1)	117 (1)
Preterm	332 (1)	186 (1)	146 (1)
Late preterm	2023 (3)	1167 (3)	856 (4)
Term	59,088 (96)	35,564 (96)	23,524 (96)

[a]May not add to 100% due to rounding.
[b]Missing in <0.5%.
[c]Missing in 1.7%.
[d]Missing in 4.3%.
SGA indicates small-for-gestational-age at birth (<10th percentile); ART, assisted reproductive treatment; BMI, prepregnancy body mass index (kg/m^2); ADHD, attention-deficit hyperactivity disorder.

follow-up participant groups as were the prepregnancy BMI-Infection associations. The Alcohol-Developmental Disorders associations were slightly lower at all levels of drinking in the follow-up participants compared with the baseline population.

The smoking-ADHD association was estimated with considerable imprecision, with the ratio of bootstrap limits equal to about 3.6, compared with ratios between 1.1 and 2.0 for the other 4 relative association estimates. This is due to the rarity of hospitalized ADHD cases. Other recent studies of ADHD/hyperkinetic disorder from Denmark that have also relied on registry-based hospitalizations for ADHD[20,21] show similar prevalences to ours. However, only the most severe cases reach the hospital, and it is possible that the selection forces we found are related to comorbidities, disease severity, or the diagnostic process.

We studied malleable lifestyle factors that could be associated with various selection mechanisms. Smoking during pregnancy is an established risk factor and may have influenced women's decision to discontinue participation. Birth cohort women in higher socio-occupational groups reported drinking alcohol during pregnancy more often than women in the lowest group. This is consistent with recent work suggesting that the highest average drinking levels occur in the most highly educated Danish men and women.[22] In our study, the women in higher socio-occupational groups may also be more likely to have had the time and resources to continue participation in the birth cohort. We did not have access to information on the number of alcohol drinks per week, however, and therefore could not examine these relationships in greater detail.

Assisted reproductive treatment, being overweight, and having SGA babies are preexisting conditions that could also

TABLE 3. Relative Ratios Based on Adjusted Ratios in the Baseline and Follow-up Populations, Danish National Birth Cohort (1996–2008)

	Baseline Population AOR or ARR (95% CI)	Follow-up Population AOR or ARR (95% CI)	Ratio Follow-up/Baseline Ratio[a] (95% CI)[b]
1. SGA and asthma[c]			
AGA[d]	1.00	1.00	1.00
SGA	1.34 (1.18–1.51)	1.45 (1.23–1.70)	1.08 (0.97–1.21)
2. ART and hospitalization[c]			
No ART[d]	1.00	1.00	1.00
ART	1.07 (1.05–1.09)	1.08 (1.05–1.10)	1.01 (0.99–1.03)
3. BMI[e] and infection[c]			
BMI[e]	1.07 (1.02–1.12)	1.03 (0.97–1.10)	0.97 (0.93–1.01)
4. Alcohol and disorders of psychologic development[f]			
Nondrinker[d]	1.00	1.00	1.00
<1 drink/week	0.84 (0.55–1.27)	0.80 (0.48–1.34)	0.97 (0.69–1.33)
≥1 drink/week	0.74 (0.49–1.12)	0.66 (0.39–1.10)	0.90 (0.65–1.25)
5. Smoking (increase of 1 pack/day) and ADHD[g]			
Smoking	1.61 (0.71–3.66)	2.13 (0.79–5.74)	1.33 (0.70–2.52)

[a]Bias-corrected ln(relative ratio) (see text for details).
[b]95% bootstrap confidence intervals with 5000 resamplings (see text for details).
[c]Adjusted for socioeconomic status, child's sex, time, preterm birth, alcohol drinking, smoking (packs/day).
[d]Reference category.
[e]Prepregnancy body mass index in units of 10.
[f]Adjusted for socioeconomic status, child's sex, time, smoking (packs/day).
[g]Adjusted for socioeconomic status, child's sex, time, alcohol drinking.
AGA indicates appropriate size for gestational age at birth (≥10th percentile).

TABLE 4. Smoking (Packs/Day) in Lost-to-follow-up and Follow-up Participants, Danish National Birth Cohort (1996–2008)

Category	Pack/Day Codes[a]	Lost to Follow-up No.	Follow-up Participants No.	Total No.[b]	Loss to Follow-up %	Risk Ratio Relative to Nonsmokers (95% CI)
No smoking during pregnancy	0	18,569	30,174	48,743	38	1.00[c]
Unspecified smoking status	0.10	843	1250	2093	40	1.06 (1.01–1.12)
1–5 cigarettes/day or stopped during 1st trimester	0.15	1911	2274	4185	46	1.20 (1.16–1.24)
6–10 cigarettes/day or stopped after 1st trimester	0.40	1796	2010	3806	47	1.24 (1.20–1.29)
Smoked unspecified amount	0.45	119	126	245	49	1.28 (1.13–1.46)
11–20 cigarettes/day	0.80	1272	1134	2406	53	1.39 (1.34–1.45)
More than 20 cigarettes/day	1.50	146	109	255	57	1.50 (1.35–1.67)
Test for trend					$P < 0.001$	

[a]Codes were assigned before examining participation rates.
[b]Missing n = 162.
[c]Reference category.

be related to the decision to continue participation in the birth cohort, perhaps by different mechanisms. The inability to conceive without assistance might have caused women to doubt their own fecundity, thus encouraging these women to continue participation in the birth cohort (the Danish name for which translates as "Better Health for Mother and Child"). High BMI may be a proxy for lower levels of education/income and poor lifestyle and therefore lower participation. For women who gave birth to SGA babies, their concern regarding what they might have done to cause this might have motivated continued participation.

Our results address loss to follow-up in a situation where mothers know the outcome when deciding whether or not to continue participation. In some cases, knowing the outcome may prompt a woman to continue participating, so she may learn more about why a particular condition occurred in the child. In other cases, the extra time needed to care for a child with special needs may prevent a woman from continuing participation even if she so desired. Although we cannot know which of these was the predominant factor influencing the decision, we would expect to see less follow-up bias due to either of these forces when the study outcomes do not occur until later in life.

TABLE 5. Logistic Regression Analyses of Loss to Follow-up on SGA, Asthma, ART, Hospitalization, BMI, Infection, Alcohol, Developmental Disorders, and Smoking. Danish National Birth Cohort (1996–2008)

Variable	Controlling for	OR (95% CI)
SGA	None	1.16 (1.09–1.24)
	Smoking	1.09 (1.02–1.16)
	Asthma	1.16 (1.09–1.23)
	Asthma, smoking	1.08 (1.01–1.15)
Asthma	None	1.24 (1.16–1.33)
	Smoking	1.20 (1.12–1.29)
	SGA	1.24 (1.16–1.33)
	SGA, smoking	1.20 (1.12–1.28)
ART	None	0.99 (0.93–1.07)
	Smoking	1.01 (0.94–1.08)
	Hospitalization	0.99 (0.92–1.06)
	Hospitalization, smoking	1.00 (0.94–1.08)
Hospitalization[a]	None	1.08 (1.06–1.11)
	Smoking	1.07 (1.05–1.10)
	ART	1.08 (1.06–1.11)
	ART, smoking	1.07 (1.05–1.10)
BMI[b]	None	1.19 (1.15–1.24)
	Smoking	1.20 (1.15–1.24)
	Infection	1.19 (1.15–1.24)
	Infection, smoking	1.20 (1.15–1.24)
Infection	None	1.12 (1.08–1.17)
	Smoking	1.10 (1.06–1.15)
	BMI	1.12 (1.07–1.16)
	BMI, smoking	1.10 (1.05–1.14)
Alcohol (drinks/week)	None	
<1 vs. none		1.00 (0.96–1.04)
≥1 vs. none		0.90 (0.86–0.93)
	Smoking	
<1 vs. none		1.01 (0.97–1.05)
≥1 vs. none		0.90 (0.87–0.94)
	Developmental disorders	
<1 vs. none		1.00 (0.96–1.04)
≥1 vs. none		0.90 (0.86–0.93)
	Developmental disorders, smoking	
<1 vs. none		1.01 (0.97–1.05)
≥1 vs. none		0.90 (0.87–0.94)
Developmental disorders[c]	None	0.87 (0.63–1.20)
	Smoking	0.84 (0.60–1.17)
	Alcohol	0.85 (0.62–1.18)
	Smoking, alcohol	0.83 (0.60–1.15)
Smoking[d]	None	2.15 (1.99–2.33)
	SGA	2.12 (1.96–2.29)
	Asthma	2.12 (1.96–2.30)
	SGA, asthma	2.09 (1.93–2.27)
	ART	2.15 (1.99–2.33)
	Hospitalization	2.12 (1.96–2.30)
	ART, hospitalization	2.12 (1.96–2.30)
	BMI	2.16 (1.99–2.34)
	Infection	2.13 (1.96–2.30)
	BMI, infection	2.14 (1.98–2.32)
	Alcohol	2.15 (1.99–2.33)
	Developmental disorders	2.14 (1.98–2.32)
	Alcohol, developmental disorders	2.15 (1.96–2.33)

[a]Hospitalization has been rescaled by dividing the number of hospitalizations by 10.
[b]Prepregnancy body mass index measured in increases of 10.
[c]Disorders of psychologic development.
[d]Smoking in packs/day.

Participation proportions in Table 4 suggest smoking during pregnancy was an important factor affecting follow-up. Baseline factors that influence loss to follow-up can produce bias if uncontrolled.[1,2,14] Because prenatal smoking is associated with many health-related factors and is a risk factor for many conditions, as well as a strong predictor of loss to follow-up, its control in pregnancy cohort studies over follow-up time may reduce follow-up bias as well as confounding. If any factors that affect a mother's decision to participate in the follow-up are known and adequately measured on all members of the source population, bias may be reduced by controlling for them or their surrogates in the analyses. We identified smoking as one such surrogate. In studies in which these covariates are unmeasured, other methods to account for follow-up bias will be needed[15,23,24]; these methods may use estimates of relative ratios (eg, Table 3) as a starting point for sensitivity analysis or prior distributions.

The principal strengths of our investigation are a large sample size, nearly complete covariate information on all subjects, covariate information collected prior to the outcome occurrence, and outcomes that were registry-based, and therefore available for all subjects in our baseline population, allowing us to estimate the relation of losses to exposures and outcomes. Limitations of our study include misclassification of self-reported measures such as smoking, alcohol drinking, ART, prepregnancy BMI, and socioeconomic status. Misclassification of outcome measures can also occur with the use of registries to ascertain cases; only the most severe occurrences are presumably listed. In addition, we examined only 5 associations among all that could have been considered. Finally, we could study only those who entered the study. Our results might not extend to those who had been invited to participate but declined to join. However, an earlier Danish National Birth Cohort study found no evidence of serious bias related to the initial recruitment.[25]

In conclusion, bias from loss to follow-up in a life-course cohort study may be quite modest for medical factors whereas for behavioral factors it may be large. In particular, maternal smoking appeared strongly related to loss and out-

come. Alcohol consumption did not appear to have a large effect, although as with our other results this finding may be specific to Nordic populations. The trade-off between broad recruitment and minimizing loss to follow-up may seem to favor enrolling a subset of motivated participants who are likely to participate in the study long term. Unfortunately, the results may not then be generalizable to people who practice the most risky behaviors and may thus be in the greatest need of study. We had access to outcomes for all baseline cohort members, regardless of their eventual follow-up participation status. Our study offers support for the notions that (1) the ultimate uses of a study, especially in terms of exposures of interest, should play a role in recruitment strategies and (2) detailed measurement of high-risk behaviors may facilitate adjustment for loss to follow-up as well as control of confounding.

REFERENCES

1. Rothman KJ, Greenland S, Lash TL. Validity in epidemiologic studies. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008:128–147.
2. Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol*. 1977;106:184–187.
3. Olsen J, Melbye M, Olsen SF, et al. The Danish National Birth Cohort: its background, structure and aim. *Scand J Public Health*. 2001;29:300–307.
4. Pedersen CB, Gøtzsche H, Møller JO, Mortensen PB. The Danish Civil Registration System. *Dan Med Bull*. 2006;53:441–449.
5. Knudsen LB, Olsen J. The Danish National Birth Registry. *Dan Med Bull*. 1998;45:320–323.
6. Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish National Hospital Register. A valuable source of data for modern health sciences. *Dan Med Bull*. 1999;46:263–268.
7. Nepomnyaschy L, Reichman NE. Low birthweight and asthma among young urban children. *Am J Public Health*. 2006;96:1604–1610.
8. Basatemur E, Sutcliffe A. Follow-up of children born after ART. *Placenta*. 2008;S135–S140.
9. Yuan W, Basso O, Sørensen HT, Olsen J. Maternal prenatal lifestyle factors and infectious disease in early childhood: a follow-up study of hospitalization within a Danish Birth Cohort. *Pediatrics*. 2001;107:357–362.
10. Streissguth A, Barr H, Carmichael OH, Sampson P, Bookstein F, Burgess D. Drinking during pregnancy decreases word attack and arithmetic scores on standardized tests: adolescent data from a population-based prospective study. *Alcohol Clin Exp Res*. 1994;18:248–254.
11. Kotimaa AJ, Moilanen I, Taanila A, et al. Maternal smoking and hyperactivity in 8-year-old children. *J Am Acad Child Adolesc Psychiatry*. 2003;42:826–833.
12. Kramer MS, Platt RW, Wen SW, et al. A new and improved population-based Canadian reference for birth weight for gestational age. *Pediatrics*. 2001;108:E35.
13. Nohr EA, Bech BH, Davies MJ, Frydenberg M, Henriksen TB, Olsen J. Prepregnancy obesity and fetal death: a study within the Danish National Birth Cohort. *Obstet Gynecol*. 2005;106:250–259.
14. Austin MA, Criqui MH, Barrett-Connor E, Holdbrook MJ. The effect of response bias on the odds ratio. *Am J Epidemiol*. 1981;114:137–143.
15. Greenland S, Lash TL. Bias analysis. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2008:362.
16. Greenland S. Interval estimation by simulation as an alternative to and extension of confidence intervals. *Int J Epidemiol*. 2004;33:1389–1397.
17. Deeg DJ. Attrition in longitudinal population studies: does it affect the generalizability of the findings? *J Clin Epidemiol*. 2002;55:213–215.
18. Powers J, Loxton D. The impact of attrition in an 11-year prospective longitudinal study of younger women. *Ann Epidemiol*. 2010;20:318–321.
19. Kotecha SJ, Watkins WJ, Heron J, Henderson J, Dunstan FD, Kotecha S. Spirometric lung function in school age children: effect of intrauterine growth retardation and catch-up growth. *Am J Respir Crit Care Med*. 2010;181:969–974.
20. Atladottir HO, Parner ET, Schendel D, Dalsgaard S, Thomsen PH, Thorsen P. Time trends in reported diagnoses of childhood neuropsychiatric disorders. *Arch Pediatr Adolesc Med*. 2007;161:193–198.
21. Linnet KM, Wisborg K, Secher NJ, et al. Coffee consumption during pregnancy and the risk of hyperkinetic disorder and ADHD: a prospective cohort study. *Acta Paediatr*. 2009;98:173–179.
22. Johnson W, Ohm Kyvik K, Mortensen E, et al. Does education confer a culture of healthy behavior? smoking and drinking patterns in Danish twins. *Am J Epidemiol*. 2011;173:55–63.
23. Lash TL, Fox MP, Fink AK. *Applying Quantitative Bias Analysis to Epidemiologic Data*. New York: Springer; 2009;142–144.
24. Greenland S. Bayesian perspectives for epidemiologic research: III. Bias analysis via missing-data methods. *Int J Epidemiol*. 2009;38:1662–1673.
25. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? *Epidemiology*. 2006;17:413–418.

