NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

# Analysis of Self-Selection Bias in a Population-Based Cohort Study of Autism Spectrum Disorders

**Roy M. Nilsen**[a], **Pål Surén**[b], **Nina Gunnes**[b], **Elin R. Alsaker**[b], **Michaeline Bresnahan**[c], **Deborah Hirtz**[d], **Mady Hornig**[c], **Kari Kveim Lie**[b], **W. Ian Lipkin**[c], **Ted Reichborn-Kjennerud**[b,e], **Christine Roth**[b], **Synnve Schjølberg**[b], **George Davey Smith**[f], **Ezra Susser**[c], **Stein Emil Vollset**[a,b], **Anne-Siri Øyen**[g], **Per Magnus**[b], and **Camilla Stoltenberg**[a,b]

[a]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

[b]Norwegian Institute of Public Health, Oslo, Norway

[c]Mailman School of Public Health, Columbia University, New York, NY, USA

[d]National Institute of Neurological Disorders and Stroke, Bethesda, MD, USA

[e]Institute of Psychiatry, University of Oslo, Oslo, Norway

[f]MRC Centre for Causal Analysis in Translational Epidemiology, University of Bristol, Bristol, United Kingdom

[g]Nic Waals Institute, Lovisenberg Hospital, Oslo, Norway

## Abstract

**Background—**This study examined potential self-selection bias in a large pregnancy cohort by comparing exposure-outcome associations from the cohort to similar associations obtained from nationwide registry data. The outcome under study was specialist-confirmed diagnosis of autism spectrum disorders.

**Methods—**The cohort sample (n = 89,836) was derived from the population-based prospective Norwegian Mother and Child Cohort Study and its sub-study of autism spectrum disorders, the Autism Birth Cohort study. The nationwide registry data were derived from the Medical Birth Registry of Norway (n = 507,856). The children were born in 1999-2007, and seven prenatal and perinatal exposures were selected for analyses.

**Results—**Autism spectrum disorders were reported for 234 (0.26%) children in the cohort and 2,072 (0.41%) in the nationwide population. Compared with the nationwide population, the cohort had an underrepresentation of the youngest women (<25 years), those who had single status, mothers who smoked during pregnancy, and nonusers of prenatal folic acid supplements. The ratios of the adjusted odds ratios in the cohort over the adjusted odds ratios in the nationwide population were as follows; primipara pregnancy: 1.39/1.22, prenatal folic acid use: 0.85/0.86, prenatal smoking: 1.20/1.17, preterm birth (<37 weeks): 1.48/1.42, low birthweight (<2,500 g): 1.60/1.58, male sex: 4.39/4.59 (unadjusted only); and cesarean section history: 1.03/1.04.

**Conclusions—**Associations estimated between autism spectrum disorders and perinatal and prenatal exposures in the cohort are close to those estimated in the nationwide population. Self-selection does not appear to compromise validity of exposure-outcome associations in the Autism Birth Cohort study.

Corresponding Author Roy Miodini Nilsen, Department of Global Public Health and Primary Care, University of Bergen, Kalfarveien 31, 5018 Bergen, Norway, Telephone: (+47) 55 97 70 75, Fax number: (+47) 55 97 60 88, roy.nilsen@uib.no.

# BACKGROUND

Self-selection and low participation proportion is a well-known challenge in epidemiologic studies.[1] When eligible participants choose not to take part in a study, the nonparticipation may introduce a bias in the effect estimates of exposure-outcome associations, a situation usually referred to as selection bias.[2] In conventional effect estimates (i.e., odds ratio or risk ratio), this type of bias occurs when participation depends on both the exposure and the outcome under study, or indirectly when participation is influenced by underlying factors that are also associated with both the exposure and the outcome under study.[2, 3]

Self-selection and its consequences in epidemiologic studies have been investigated in association with various health outcomes, but for many outcomes this remains an uncharted area. Over the past two decades, several population-based epidemiologic studies (both prospective studies and case-control studies) have been or are being conducted to detect potential genetic and environmental causes of autism spectrum disorders (ASDs) in the offspring.[4-7] While offering important contributions to public health and ASD research, the general knowledge of systematic bias in exposure-outcome associations due to self-selection in these and other ASD studies is limited.

Various methods can be used to investigate selection bias in epidemiologic studies. Most recently, one attractive method is to quantify bias by comparing exposure-outcome associations from the study participants to similar associations obtained from the source population using registry data that are available for both populations.[8-10] Because registry data are independently collected without knowing which individuals will participate in a future epidemiologic study or not, the presence of bias in effect estimates is likely to reflect selection bias. Further, since the study is intended to reflect the source population, the method may also be regarded as a direct measure of generalizability of the study results.

In the present study, we aimed to examine potential self-selection bias in the population-based Norwegian Mother and Child Cohort Study (MoBa) and its sub-study of ASD, the Autism Birth Cohort (ABC) study.[6] The associations of ASD with seven selected prenatal and perinatal exposures in the cohort sample were compared with similar associations obtained from nationwide registry data consisting of all liveborn children in Norway in the same time period. The nationwide registry data were derived from the Medical Birth Registry of Norway (MBRN).

# METHODS

## Cohort sample

MoBa is a prospective pregnancy cohort that includes 109,020 children.[11] Mothers were recruited to the study through a postal invitation after they had signed up for the routine ultrasound examination at their local hospital, approximately around gestational week 18 (participation proportion 38.5%). The children were born between August 1999 and July 2009, and data are linked to the nationwide MBRN to obtain additional registered pregnancy and birth data (see description below). The analyses in the present study included children born in 1999-2007 who were recorded to be alive and living in Norway past the age of three years. Since follow-up on ASD diagnosis was in 31 December 2010 (see description below), all children in our study were three years or older. A total of 89,836 children in the MoBa cohort were included in the analyses.

## Nationwide population

The nationwide registry data were derived from the MBRN.[12] This registry was established in 1967 and contains data for all pregnancies lasting more than 16 weeks (12 weeks since

2001). Reporting to the MBRN is mandated by law. A standardized notification form is used to record essential information about the pregnancy and the delivery, such as demographic information, maternal reproductive history, maternal health before and during pregnancy, pregnancy and birth complications, and pregnancy outcomes. The nationwide population included children born in the same period as the cohort sample (1999-2007) who were recorded to be alive and living in Norway past the age of three years. A total of 507,856 children were included in the nationwide population.

### Diagnosis of ASD

Information about ASD diagnoses for the cohort sample and the nationwide population was obtained by linkage to the Norwegian Patient Registry (NPR).[13] The NPR is an administrative database containing activity data from all Norwegian specialist health services, i.e., all hospitals and outpatient clinics in Norway. All ASD diagnoses are specialist-confirmed by paediatricians, child psychiatrists, or specialists in clinical psychology. Reporting of data to the NPR is mandatory and linked to the governmental reimbursement system for funding of health services. Individual-level research data are available from 2008 onwards. The data files provided for the present study contained NPR diagnoses until December 31st, 2010. Diagnoses are reported according to *International Statistical Classification of Diseases and Related Health Problems*, Tenth Revision. The ASD case definition in the present study included F84.0 Childhood Autism, F84.1 Atypical Autism, F84.5 Asperger's Syndrome, F84.8 Other Pervasive Developmental Disorder, and F84.9 Pervasive Developmental Disorder, Unspecified. Diagnoses of F84.2 Rett's Syndrome and F84.3 Childhood Disintegrative Disorder were not included.

### Validation of ASD diagnoses

All MoBa children identified with ASD diagnoses in the NPR are invited to participate in an assessment that includes the research-standard instruments for diagnosis of ASD, the Autism Diagnostic Interview - Revised (ADI-R)[14] and the Autism Diagnostic Observation Schedule (ADOS).[15] The NPR diagnoses have a high validity for ASD; of 60 NPR cases clinically assessed through the ABC study to date, 58 were found to meet the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) criteria for ASD, resulting in a positive predictive value of 97% [95% confidence interval 88%,100%]. The ABC study also identifies ASD cases through screening and referrals of potential ASD cases,[6] but we only included ASD diagnoses from the NPR in this study in order to make the cohort sample and the nationwide population comparable. The overlap of ASD cases between the NPR and the ABC study among children born in 1999-2007 is shown in Figure 1. The 51 children that were diagnosed with ASD in the ABC study but yet had no diagnosis in the nationwide population (NPR) were treated as non-ASD cases in the cohort sample. Excluding these 51 ASD cases from the analyses altogether did not alter any of the effect estimates of exposure-outcome associations.

### Baseline data

Information about background characteristics and the seven selected exposures was obtained from the MBRN and included maternal age (<25, 25-29, 30-34,  35 years) and paternal age (<30, 30-34, 35-39,  40 years) at delivery, maternal marital status (single, cohabiting/married), parity (0, 1,  2 previous deliveries), cesarean section history (no, yes), hospital size (<500, 500-1,499, 1,500-2,999,  3,000 births per year), prenatal folic acid use (no, yes), prenatal smoking (no, yes), preeclampsia (no, yes), preterm birth (no, yes), low birthweight (no, yes), offspring sex (female, male), and multiple pregnancies (no, yes). Prenatal smoking was defined as any maternal smoking during pregnancy and was recorded in the MBRN by check boxes for daily and occasionally smoking at the beginning of and at the end of

pregnancy.[16] Likewise, prenatal folic acid use was defined as any use of maternal folic acid supplements before and/or during pregnancy and was recorded by check boxes on regular use before or during pregnancy.[16] Low birthweight was defined as <2,500 g, whereas preterm birth was defined as <37 completed weeks of gestation. Gestational age was based on second trimester ultrasound measurements or, if missing, the first day of the last reported menstrual period. Hospital size was used as a proxy for geographical area. In Norway, smaller hospitals are generally found in the rural areas, outside the large cities.

For the overall nationwide population, information was missing for 4 (0.0%) children on maternal age, 3,433 (0.7%) children on paternal age, 83,843 (16.5%) children on prenatal smoking, 3,270 (0.6%) children on preterm birth, and 635 (0.1%) children on low birthweight. For the cohort sample, information was missing for 261 (0.3%) children on paternal age, 14,134 (15.7%) on prenatal smoking, 352 (0.4%) children on preterm birth, and 59 (0.1%) children on low birthweight.

We estimated the risk of ASD associated with the following seven prenatal and perinatal exposures: primipara pregnancy (no, yes), prenatal folic acid use (no, yes), prenatal smoking (no, yes), low birthweight (no, yes), preterm birth (no, yes), offspring sex (female, male), and cesarean section history (no, yes). Except for cesarean section history, all variables were chosen because they have been associated with ASD in previous studies.[17-24] A history of cesarean section was expected not to be associated with ASD, and was included as a control exposure. All associations were estimated separately for the cohort sample and the nationwide population.

### Statistical analyses

Statistical analyses were performed by using SAS version 9.2 (SAS Institute, Inc., Cary, North Carolina) software for Windows.

Distributions of background characteristics and exposures were calculated as relative frequencies for the cohort sample and the nationwide population separately. To determine whether a variable category was overrepresented or underrepresented in the cohort sample, we estimated the ratio of relative frequencies by dividing the relative frequency for the variable category in the cohort sample by the corresponding relative frequency in the nationwide population. A ratio below 1 indicates an underrepresentation of that category in the cohort, whereas a ratio above 1 indicates an overrepresentation. As the cohort sample is a subsample of the nationwide population, uncertainties in the ratio estimates were quantified with 95% confidence intervals using a non-parametric bootstrap method of size 4,000.[9]

The aforementioned procedure was also used to quantify the magnitude and direction of relative differences in odds ratios (ORs) for the associations between ASD and the seven selected exposures. ORs with 95% confidence intervals were estimated separately for the cohort sample and the nationwide population from logistic regression models. If an exposure increases the risk of outcome, a ratio of ORs ($OR_{cohort}/OR_{nationwide}$) below 1 indicates an underestimation of the exposure-outcome association in the cohort sample, whereas a ratio of ORs above 1 indicates an overestimation. This is reversed when the exposure is protective against the outcome. Very few women had two or more children (either siblings or twin sets) with an ASD diagnosis. Given the large sample size in both the cohort sample and the nationwide population, potential intra-individual correlation due to clustering within families was therefore not accounted for in the logistic regression analyses.

To investigate how potential over- and underestimation occurred, we also calculated the proportions of participants in the cohort sample within cells of contingency tables of ASD

and each of the seven selected exposures using the numbers from the nationwide population as denominators. A cross-product ratio of these proportions provides a quantity that is similar to that of the ratio of the unadjusted ORs.[25] The inverse of this cross-product ratio is referred to as the "selection bias odds ratio" and can be used to correct the OR estimated in the cohort sample.[2]

It has been proposed that sample selection may alter the confounding patterns originally present in the general population.[10] The comparison of ORs in our study was therefore also performed after adjustment for the following covariates: year of birth (as a linear term), maternal age (<25, 25-29, 30-34, 35 years), paternal age (<30, 30-34, 35-39, 40 years), marital status (single, cohabiting/married), and parity (0, 1, 2 previous deliveries), hospital size (<500, 500-1,499, 1,500-2,999, 3,000 births per year). Because sex was poorly associated with any of the covariates (i.e., 0.90 OR 1.10; Supplemental Table 1), only the unadjusted OR for the association between sex and ASD was presented. For the same reason, we omitted marital status as a covariate in the analysis of cesarean section history, paternal age in the analysis of preterm birth, and year of birth in the analysis of low birthweight. In the analyses of preterm birth and low birthweight, we instead adjusted the ORs for prenatal smoking (no, yes), which was associated with both exposures (Supplemental Table 1).

To avoid unnecessary loss of statistical power in regression analyses, missing data in exposures and covariates were imputed by using a multiple imputation method.[26] We assumed data were missing at random, i.e., missing data can be recovered by the help of observed data of the exposures, covariates, and other baseline data. All effect estimates are presented based on multiple imputations. Simple imputations (i.e., missing data coded as an individual category), showed almost similar effect estimates (not shown).

## RESULTS

The proportion of participants in the cohort sample among all children in the nationwide registry data was 17.7% (89,836/507,856). The overall cumulative incidence of ASD was 0.26% (234/89,836) in the cohort sample and 0.41% (2,072/507,856) in the nationwide population (Table 1). In both the cohort sample and the nationwide population, the cumulative incidence decreased by year of birth. The average age of the children was 5.5 years for the cohort sample and 7.0 years for the nationwide population.

The mothers participating in the cohort sample differed from the mothers in the nationwide population on several background characteristics and exposures (Table 2). Overall, the cohort sample had an underrepresentation of the youngest women (<25 years), women who had single status, and women who delivered at hospitals with 1500 to 2999 births per year (relative deviations from 28% to 49%). Mothers who smoked during pregnancy were also underrepresented (relative deviation of 35%), whereas mothers using prenatal folic acid supplements were overrepresented (relative deviation of 47%). For other variables, the relative deviation was less than 16% between the cohort sample and the nationwide population (Table 2).

Despite differences in the distribution of ASD, background characteristics, and exposures, the effect estimates for exposure-outcome associations were essentially the same in the cohort sample and the nationwide population (Table 3). The ratios of the adjusted ORs in the cohort sample over the adjusted ORs in the nationwide population were as follows; primipara pregnancy: 1.39/1.22, prenatal folic acid use: 0.85/0.86, prenatal smoking: 1.20/1.17, preterm birth (<37 weeks): 1.48/1.42, low birthweight (<2,500 g): 1.60/1.58, male sex: 4.39/4.59 (unadjusted); and cesarean section history: 1.03/1.04. The largest differences

in unadjusted and adjusted ORs of ASD between the cohort sample and the nationwide population were observed for primipara pregnancy (relative deviations of 10% and 14%, respectively).

Year of birth was an important adjustment variable for the association of prenatal folic acid use and prenatal smoking with ASD (not shown). In addition to being inversely related to ASD risk during the study period (Table 1), year of birth was also strongly associated with increased prenatal folic acid use and decreased prenatal smoking in both the cohort sample and the nationwide population (Supplemental Table 1).

As expected, the cross-product ratio of individual cell proportions of participants within each contingency table of the dichotomous exposure and outcome was equal to the corresponding ratio of unadjusted ORs in Table 3 (Table 4). Furthermore, the ratio of participation of ASD cases over non-cases appeared to be similar across exposure categories for all seven exposures (overall ratio 0.64; range: 0.60 – 0.70), suggesting that the associations between participation and the exposures were not modified by ASD outcomes.

We also performed analyses of bias in the seven exposure-outcome associations after excluding multiple births. In these analyses, relative differences in ORs of ASD between the cohort sample and the nationwide population were slightly increased for preterm birth and low birthweight compared with the analyses that included multiple births (Table 5).

## DISCUSSION

This study examined the potential effects of self-selection on seven exposure-outcome associations in the MoBa cohort and its sub-study of ASD, the ABC study. Despite under- and overrepresentation of several exposures in the cohort, all effect estimates of exposure-outcome associations deviated less than 16% from those estimated in the nationwide population comprising all children born in Norway in the same time period. Even for prenatal folic acid use and prenatal smoking, which were strongly over- and underrepresented in the cohort, bias in risk estimates was minimal. These findings are consistent with those of previous prospective cohort studies[8-10] and demonstrate that large prospective studies are generally robust against bias arising from initial participation and self-selection.

The unique contribution of the present study is the ability to compare effect estimates between cohort participants and children in a whole country, using exposure and diagnosis data from the exact same sources. Many studies addressing self-selection and low participation in epidemiologic studies usually assess difference in prevalence estimates of exposures and outcome between participants and non-participants, and not the differences in the estimates of exposure-outcome associations. Although differences in prevalence may point to the direction of non-representativeness, such findings are generally not adequate to reliably determine whether selection bias has affected the effect estimates.[8-10]

The findings of the present study may be generalizable to other prospective ASD studies, but some caution is warranted. Since ORs in the cohort sample are validated against ORs in the nationwide population, we assume that the nationwide population provides the results closest to the truth. This assumption may not hold if exposures are reported differently for cohort participants than for the nationwide population overall. Furthermore, for small studies, random sampling variation will be greater than that seen in this study. Therefore, precision of estimates will be lower, and deviations in effect estimates may be due to random error and not necessarily due to selection bias. Additionally, our analyses pertain to self-selection into the cohort, and all of the exposures are ascertained before or at birth. Continued participation may involve a different set of selection factors, and exposures

ascertained at a later age may or may not follow this pattern. Therefore, potential bias may still be present in other exposure-outcome associations than those studied here.

Data on ASD diagnoses were only available from 2008 onwards. If a child was diagnosed with ASD before 2008 and had not been in contact with specialist health services in 2008 or later, the ASD diagnosis would not have been recorded for that child in the NPR.[13] However, given that ASD data were collected from the joint NPR database and we only considered NPR diagnoses for this study, the proportions of undiagnosed cases are likely to be equal for both the cohort sample and the nationwide population and may thus not have influenced the comparison of effect estimates between them. Notably, if the ASD diagnosis was first made before 2008, contact with specialist health services subsequent to the initial diagnosis will provide additional opportunities to capture ASD diagnostic codes in the NPR.

The average age of children at the end of follow-up was 5.5 years in the cohort sample and 7.0 years in the nationwide population. In Norway, ASDs and particularly Asperger's syndrome are often diagnosed after the age of 6 years.[13] Consequently, the cumulative incidence of ASD in our study was lower among the younger children, and the overall cumulative incidence was lower in the cohort sample than in the nationwide population (0.26% versus 0.41%). The lower mean age of children in MoBa was attributable to the fact that only a few hospitals were included during the initial phase of recruitment in 1999-2001.[9, 11]

Mothers participating in MoBa have healthier lifestyle patterns and higher socio-economic status than the nationwide population. [9, 11] If these factors also are strongly associated with both the exposure and outcome under study, such selection might induce selection bias and alter confounding patterns originally present in the nationwide population.[10] In the present analysis, we adjusted the effect estimates for several available covariates that were associated with the exposure of interest (Supplemental Table 1). Year of birth influenced the effect estimates of prenatal smoking and prenatal folic acid use, but adjustment for this and other covariates did not change the relative difference in any of the estimated ORs, suggesting that confounding patterns were approximately the same in the cohort sample and the nationwide population.[10]

Consistent with previous reports, results from the cohort sample and the nationwide population showed that the risk of ASD was increased for boys,[17, 18] firstborn children, [19] children born with birthweight below 2,500 g,[20, 21] children born before gestational week 37, [22, 23] and children whose mothers smoked during pregnancy. [24] In the cohort sample, confidence intervals of ORs overlapped with 1 for prenatal smoking, preterm birth, and low birthweight, possibly because of low statistical power.

We further observed that maternal prenatal folic acid supplement use was associated with a 14-17% adjusted risk reduction for ASD (although confidence intervals overlapped with 1 in the cohort sample). It should be noted that the MBRN does not comprise precise information on timing, dose or frequency of folic acid supplement use, and that there is also underreporting of folic acid use in the registry.[27] Studies using detailed supplementation data from MoBa specific questionnaires have found substantially stronger associations between maternal folic acid use and the children's risk of ASD,[27, 28] particularly when folic acid supplements are used within the time interval lasting from four weeks before to eight weeks after the start of pregnancy (approximately 45% risk reduction). Associations of ASD with prenatal folic acid use and maternal folate status during pregnancy have also been reported by others,[29] but the results are not conclusive.[30]

In conclusion, this is the first study known to compare effect estimates of exposure-outcome associations from participants in a large prospective study of ASD to similar effect estimates

calculated for all liveborn children born in the same country and in the same time period. Our results show that unbiased estimates of associations can be achieved even though the marginal distributions of ASD and exposures variables differ between cohort participants (the ABC study) and the source population from which the cohort was recruited.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## References

1. Galea S, Tracy M. Participation rates in epidemiologic studies. Annals of Epidemiology. 2007; 17:643–653. [PubMed: 17553702]

2. Lash, TL.; Fox, MP.; Fink, AK. Applying Quantitative Bias Analysis to Epidemiologic Data. New York: Springer; 2009.

3. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004; 15:615–625. [PubMed: 15308962]

4. Landrigan PJ, Trasande L, Thorpe LE, Gwynn C, Lioy PJ, D'Alton ME, et al. The National Children's Study: a 21-year prospective study of 100,000 American children. Pediatrics. 2006; 118:2173–2186. [PubMed: 17079592]

5. Hertz-Picciotto I, Croen LA, Hansen R, Jones CR, van de Water J, Pessah IN. The CHARGE study: an epidemiologic investigation of genetic and environmental factors contributing to autism. Environmental Health Perspectives. 2006; 114:1119–1125. [PubMed: 16835068]

6. Stoltenberg C, Schjolberg S, Bresnahan M, Hornig M, Hirtz D, Dahl C, et al. The Autism Birth Cohort: a paradigm for gene-environment-timing research. Molecular Psychiatry. 2010; 15:676–680. [PubMed: 20571529]

7. Schendel DE, Diguiseppi C, Croen LA, Fallin MD, Reed PL, Schieve LA, et al. The Study to Explore Early Development (SEED): a multisite epidemiologic study of autism by the Centers for Autism and Developmental Disabilities Research and Epidemiology (CADDRE) network. Journal of Autism and Developmental Disorders. 2012; 42:2121–2140. [PubMed: 22350336]

8. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? Epidemiology. 2006; 17:413–418. [PubMed: 16755269]

9. Nilsen RM, Vollset SE, Gjessing HK, Skjaerven R, Melve KK, Schreuder P, et al. Self-selection and bias in a large prospective pregnancy cohort in Norway. Paediatric and Perinatal Epidemiology. 2009; 23:597–608. [PubMed: 19840297]

10. Pizzi C, De Stavola BL, Pearce N, Lazzarato F, Ghiotti P, Merletti F, et al. Selection bias and patterns of confounding in cohort studies: the case of the NINFEA web-based birth cohort. Journal of Epidemiology and Community Health. 2011; 65:407–411. [PubMed: 20881022]

11. Magnus P, Irgens LM, Haug K, Nystad W, Skjaerven R, Stoltenberg C, et al. Cohort profile: the Norwegian Mother and Child Cohort Study (MoBa). International Journal of Epidemiology. 2006; 35:1146–1150. [PubMed: 16926217]

12. Irgens LM. The Medical Birth Registry of Norway. Epidemiological research and surveillance throughout 30 years. Acta Obstetricia et Gynecologica Scandinavica. 2000; 79:435–439. [PubMed: 10857866]

13. Suren P, Bakken IJ, Aase H, Chin R, Gunnes N, Lie KK, et al. Autism spectrum disorder, ADHD, epilepsy, and cerebral palsy in Norwegian children. Pediatrics. 2012; 130:e152–158. [PubMed: 22711729]

14. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders. 1994; 24:659–685. [PubMed: 7814313]

15. Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. Journal of Autism and Developmental Disorders. 2000; 30:205–223. [PubMed: 11055457]

16. Nilsen RM, Vollset SE, Rasmussen SA, Ueland PM, Daltveit AK. Folic acid and multivitamin supplement use and risk of placental abruption: a population-based registry study. American Jouornal of Epidemiology. 2008; 167:867–874.

17. Williams E, Thomas K, Sidebotham H, Emond A. Prevalence and characteristics of autistic spectrum disorders in the ALSPAC cohort. Developmental Medicine & Child Neurology. 2008; 50:672–677. [PubMed: 18754916]

18. Scott FJ, Baron-Cohen S, Bolton P, Brayne C. Brief report: prevalence of autism spectrum conditions in children aged 5-11 years in Cambridgeshire, UK. Autism. 2002; 6:231–237. [PubMed: 12212915]

19. Gardener H, Spiegelman D, Buka SL. Prenatal risk factors for autism: comprehensive meta-analysis. The British Journal of Psychiatry. 2009; 195:7–14. [PubMed: 19567888]

20. Maimburg RD, Vaeth M. Perinatal risk factors and infantile autism. Acta Psychiatrica Scandinavica. 2006; 114:257–264. [PubMed: 16968363]

21. Lampi KM, Lehtonen L, Tran PL, Suominen A, Lehti V, Banerjee PN, et al. Risk of Autism Spectrum Disorders in Low Birth Weight and Small for Gestational Age Infants. Journal of Pediatrics. 2012; 161:830–836. [PubMed: 22677565]

22. Buchmayer S, Johansson S, Johansson A, Hultman CM, Sparen P, Cnattingius S. Can association between preterm birth and autism be explained by maternal or neonatal morbidity? Pediatrics. 2009; 124:e817–825. [PubMed: 19841112]

23. Larsson HJ, Eaton WW, Madsen KM, Vestergaard M, Olesen AV, Agerbo E, et al. Risk factors for autism: perinatal factors, parental psychiatric history, and socioeconomic status. American Journal of Epidemiology. 2005; 161:916–925. discussion 926-918. [PubMed: 15870155]

24. Hultman CM, Sparen P, Cnattingius S. Perinatal risk factors for infantile autism. Epidemiology. 2002; 13:417–423. [PubMed: 12094096]

25. Austin MA, Criqui MH, Barrett-Connor E, Holdbrook MJ. The effect of response bias on the odds ratio. American Journal of Epidemiology. 1981; 114:137–143. [PubMed: 7246521]

26. Yuan Y. Multiple imputation using SAS software. Journal of Statistical Software. 2011; 45:1–25.

27. Suren P, Roth C, Bresnahan M, Haugen M, Hornig M, Hirtz D, et al. Association between maternal use of folic acid supplements and risk of autism spectrum disorders in children. Journal of the American Medical Association. 2013; 309:570–577. [PubMed: 23403681]

28. Roth C, Magnus P, Schjolberg S, Stoltenberg C, Suren P, McKeague IW, et al. Folic acid supplements in pregnancy and severe language delay in children. Journal of the American Medical Association. 2011; 306:1566–1573. [PubMed: 21990300]

29. Schmidt RJ, Tancredi DJ, Ozonoff S, Hansen RL, Hartiala J, Allayee H, et al. Maternal periconceptional folic acid intake and risk of autism spectrum disorders and developmental delay in the CHARGE (CHildhood Autism Risks from Genetics and Environment) case-control study. American Journal of Clinical Nutrition. 2012; 96:80–89. [PubMed: 22648721]

30. Main PA, Angley MT, Thomas P, O'Doherty CE, Fenech M. Folate and methionine metabolism in autism: a systematic review. American Journal of Clinical Nutrition. 2010; 91:1598–1620. [PubMed: 20410097]

**FIGURE 1.**
Cases of Autism Spectrum Disorders in the Autism Birth Cohort Study and the Medical Birth Registry of Norway. Cases in the Medical Birth Registry of Norway are obtained from the Norwegian Patient Registry (n = 2,072), whereas cases in the Autism Birth Cohort study are obtained from the Norwegian Patient Registry (n = 234) as well as from screening and referrals of potential cases in the cohort study (n = 51). The present study only included children with reported diagnoses in the Norwegian Patient Registry (white area). Abbreviations: ASDs, autism spectrum disorders; MBRN, Medical Birth Registry of Norway; MoBa, Norwegian Mother and Child Cohort Study; ABC study, Autism Birth Cohort study; NPR, Norwegian Patient Registry.
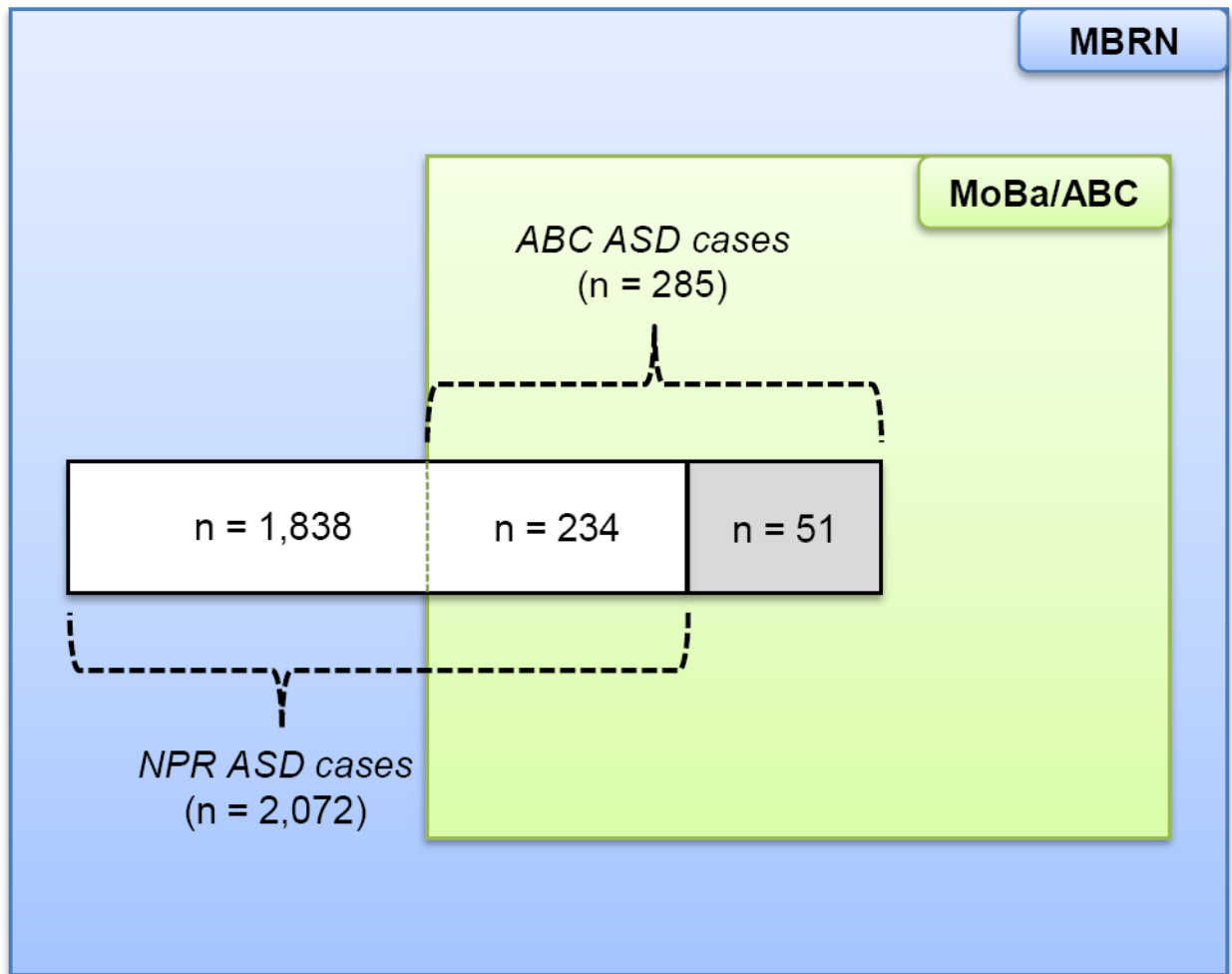
**TABLE 1**

Cumulative Incidence of Autism Spectrum Disorders in the Autism Birth Cohort Study and the Medical Birth Registry of Norway

| Year of birth | Medical Birth Registry of Norway | | | Autism Birth Cohort Study | | |
|---|---|---|---|---|---|---|
| | Children | ASDs (%)[a] | Age (range)[b] | Children | ASDs (%)[a] | Age (range)[b] |
| All years | 507,856 | 2,072 (0.41) | 7.0 (3-11) | 89,836 | 234 (0.26) | 5.5 (3-11) |
| 1999-2001 | 170,521 | 999 (0.59) | 10.0 (9-11) | 6,396 | 26 (0.41) | 9.4 (9-11) |
| 2002-2004 | 165,960 | 731 (0.44) | 7.0 (6-8) | 34,736 | 140 (0.40) | 6.9 (6-8) |
| 2005-2007 | 171,375 | 342 (0.20) | 4.0 (3-5) | 48,704 | 68 (0.14) | 4.0 (3-5) |

Abbreviations: ASDs, autism spectrum disorders.

[a] Diagnoses from the Norwegian Patient Registry were last updated on December 31st, 2010.

[b] Age of children was estimated as the mean difference between 2010 and year of birth, weighted by the relative year-of-birth frequencies.

**TABLE 2**

Comparison of Baseline Data in the Autism Birth Cohort Study with the Same Baseline Data in the Medical Birth Registry of Norway

| Baseline Data | Medical Birth Registry of Norway (%)[a] | Autism Birth Cohort Study (%)[b] | Ratio of Relative Frequencies [95% CI][c] |
|---|---|---|---|
| All Children | 507,856 (100) | 89,836 (100) | |
| Maternal age (years) | | | |
| <25 | 86,212 (17.0) | 10,413 (11.6) | 0.68 [0.67, 0.69] |
| 25-29 | 167,802 (33.0) | 29,592 (32.9) | 1.00 [0.99, 1.01] |
| 30-34 | 170,721 (33.6) | 34,506 (38.4) | 1.14 [1.13, 1.15] |
| 35 | 83,117 (16.4) | 15,325 (17.1) | 1.04 [1.03, 1.06] |
| Paternal age (years) | | | |
| <30 | 160,265 (31.6) | 24,939 (27.8) | 0.88 [0.87, 0.89] |
| 30-34 | 177,362 (34.9) | 34,663 (38.6) | 1.10 [1.10, 1.11] |
| 35-39 | 109,512 (21.6) | 21,026 (23.4) | 1.09 [1.07, 1.10] |
| 40 | 57,284 (11.3) | 8,947 (10.0) | 0.88 [0.87, 0.90] |
| Marital status | | | |
| Cohabiting/married | 467,999 (92.2) | 86,232 (96.0) | 1.04 [1.04, 1.04] |
| Single/other | 39,857 (7.8) | 3,604 (4.0) | 0.51 [0.50, 0.53] |
| Parity (no. of previous births) | | | |
| 0 | 207,454 (40.8) | 39,298 (43.7) | 1.07 [1.06, 1.08] |
| 1 | 181,469 (35.7) | 32,309 (36.0) | 1.01 [1.00, 1.01] |
| 2 | 118,933 (23.4) | 18,229 (20.3) | 0.87 [0.86, 0.88] |
| Cesarean section history[d] | | | |
| Yes | 41,279 (13.7) | 7,116 (14.1) | 1.02 [1.01, 1.04] |
| Hospital size (births per year) | | | |
| <500 | 56,583 (11.1) | 10,823 (12.0) | 1.08 [1.07, 1.10] |
| 500-1,499 | 121,880 (24.0) | 22,456 (25.0) | 1.04 [1.03, 1.05] |
| 1,500-2,999 | 143,060 (28.2) | 18,327 (20.4) | 0.72 [0.72, 0.73] |
| 3,000 | 186,333 (36.7) | 38,230 (42.6) | 1.16 [1.15, 1.17] |
| Prenatal smoking | | | |
| Missing | 83,843 (16.5) | 14,134 (15.7) | 0.95 [0.94, 0.97] |
| No | 335,749 (66.1) | 65,568 (73.0) | 1.10 [1.10, 1.11] |
| Yes | 88,264 (17.4) | 10,134 (11.3) | 0.65 [0.64, 0.66] |
| Prenatal folic acid use | | | |
| Yes | 193,315 (38.1) | 50,409 (56.1) | 1.47 [1.47, 1.48] |
| Preeclampsia | | | |
| Yes | 21,656 (4.3) | 3,688 (4.1) | 0.96 [0.94, 0.99] |
| Preterm birth (<37 weeks) | | | |
| Yes | 34,337 (6.8) | 5,736 (6.4) | 0.94 [0.92, 0.97] |
| Low birthweight (<2,500 g) | | | |
| Yes | 24,189 (4.8) | 3,852 (4.3) | 0.90 [0.88, 0.93] |
| Offspring sex | | | |

| Baseline Data | Medical Birth Registry of Norway (%)[a] | Autism Birth Cohort Study (%)[b] | Ratio of Relative Frequencies [95% CI][c] |
|---|---|---|---|
| Male | 260,296 (51.3) | 45,907 (51.1) | 1.00 [0.99, 1.00] |
| Multiple birth | | | |
| Yes | 18,232 (3.6) | 3,160 (3.5) | 0.98 [0.95, 1.01] |

Abbreviations: CI, confidence interval.

[a] Column frequencies; information was missing for 4 children on maternal age, 3,433 children on paternal age, 3,270 children on preterm birth, and 635 children on low birthweight.

[b] Column frequencies; information was missing for 261 children on paternal age, 352 children on preterm birth, and 59 children on low birthweight.

[c] The ratio of relative frequencies in the Autism Birth Cohort study to the relative frequencies in the Medical Birth Registry of Norway; 95% confidence interval was corrected for inter-dependency between samples by using a non-parametric bootstrap method.

[d] Analyses included only women with one or more previous births (Medical Birth Registry of Norway: 1,113 autism spectrum disorders cases among 300,402 children; Autism Birth Cohort study: 113 autism spectrum disorders cases among 50,538 children).

**TABLE 3**

Comparison of Unadjusted and Adjusted Odds Ratios of Autism Spectrum Disorders in the Autism Birth Cohort Study with the Same Unadjusted and Adjusted Odds Ratios in the Medical Birth Registry of Norway

| Exposures[a] | Medical Birth Registry of Norway | | Autism Birth Cohort Study | | Ratio of Unadjusted OR [95% CI][c] | Ratio of Adjusted OR [95% CI][c] |
|---|---|---|---|---|---|---|
| | Unadjusted OR [95% CI][b] | Adjusted OR [95% CI][b] | Unadjusted OR [95% CI][b] | Adjusted OR [95% CI][b] | | |
| Primipara pregnancy | 1.25 [1.15, 1.36] | 1.22 [1.11, 1.34] | 1.38 [1.07, 1.78] | 1.39 [1.05, 1.85] | 1.10 [0.86, 1.40] | 1.14 [0.87, 1.50] |
| Prenatal folic acid use | 0.69 [0.62, 0.75] | 0.86 [0.78, 0.95] | 0.72 [0.56, 0.93] | 0.85 [0.65, 1.11] | 1.05 [0.83, 1.33] | 0.99 [0.77, 1.27] |
| Prenatal smoking | 1.29 [1.16, 1.45] | 1.17 [1.04, 1.31] | 1.35 [0.96, 1.91] | 1.20 [0.84, 1.71] | 1.05 [0.74, 1.41] | 1.03 [0.71, 1.39] |
| Preterm birth (<37 weeks) | 1.48 [1.28, 1.72] | 1.42 [1.23, 1.65] | 1.59 [1.03, 2.45] | 1.48 [0.96, 2.29] | 1.07 [0.66, 1.54] | 1.04 [0.64, 1.48] |
| Low birthweight (<2,500 g) | 1.70 [1.44, 1.99] | 1.58 [1.34, 1.86] | 1.75 [1.07, 2.87] | 1.60 [0.97, 2.63] | 1.03 [0.59, 1.57] | 1.01 [0.57, 1.55] |
| Male sex | 4.59 [4.10, 5.15] | | 4.39 [3.14, 6.13] | | 0.96 [0.72, 1.36] | |
| Cesarean section history[d] | 1.01 [0.85, 1.20] | 1.04 [0.87, 1.23] | 1.01 [0.59, 1.71] | 1.03 [0.60, 1.74] | 1.00 [0.55, 1.59] | 0.99 [0.55, 1.59] |

Abbreviations: OR, odds ratio; CI, confidence interval.

[a] Reference categories for odds ratios were multipara pregnancy, no prenatal folic acid use, no prenatal smoking, term birth ( 37 weeks), normal birthweight ( 2,500 g), female sex, and no cesarean section history.

[b] The odds ratios were estimated using logistic regression models and adjusted for the covariates as described in the Methods section. Missing exposure and covariate data were imputed by using a multiple imputation method.

[c] The ratio of the odds ratio in the Autism Birth Cohort study over the odds ratio in the Medical Birth Registry of Norway; 95% confidence interval was corrected for inter-dependency between samples by using a non-parametric bootstrap method.

[d] Analyses included only women with one or more previous births (Medical Birth Registry of Norway: 1,113 autism spectrum disorders cases among 300,402 children; Autism Birth Cohort Study: 113 autism spectrum disorders cases among 50,538 children).

**TABLE 4**

Proportions of Participation in the Autism Birth Cohort Study as a Fraction of the Medical Birth Registry of Norway Within Contingency Tables of Autism Spectrum Disorders and Seven Selected Exposures

| Exposures | Autism Birth Cohort Study | | Ratio of Row Proportions[a] | Cross Product Ratio[b] |
|---|---|---|---|---|
| | Non-ASDs | ASDs | | |
| All children | 17.7% | 11.3% | 0.64 | |
| Primipara pregnancy | | | | 1.10 |
| No | 16.8% | 10.2% | 0.61 | |
| Yes | 19.0% | 12.7% | 0.67 | |
| Prenatal folic acid use | | | | 1.05 |
| No | 12.6% | 8.37% | 0.66 | |
| Yes | 26.1% | 18.2% | 0.70 | |
| Prenatal smoking | | | | 1.05 |
| No | 19.2% | 12.4% | 0.64 | |
| Yes | 12.1% | 8.17% | 0.67 | |
| Preterm birth (<37 weeks) | | | | 1.07 |
| No | 17.8% | 11.3% | 0.63 | |
| Yes | 16.7% | 11.4% | 0.68 | |
| Low birthweight (<2,500 g) | | | | 1.03 |
| No | 17.8% | 11.4% | 0.64 | |
| Yes | 16.0% | 10.5% | 0.66 | |
| Offspring sex | | | | 0.96 |
| Female | 17.8% | 11.8% | 0.66 | |
| Male | 17.7% | 11.2% | 0.63 | |
| Cesarean section history [c] | | | | 1.00 |
| No | 16.8% | 10.1% | 0.60 | |
| Yes | 17.3% | 10.4% | 0.60 | |

Abbreviations: ASDs, autism spectrum disorders.

[a] Ratio of proportions between cases and non-cases of autism spectrum disorders within exposure categories.

[b] Cross-product ratio of individual cell proportions within the contingency table of the dichotomous exposure and outcome. The cross-product ratio is equivalent to the ratio of unadjusted odds ratios in Table 3. Missing exposure data were imputed by using a multiple imputation method.

[c] Analyses included only women with one or more previous births (Medical Birth Registry of Norway: 1,113 autism spectrum disorders cases among 300,402 children; Autism Birth Cohort study: 113 autism spectrum disorders cases among 50,538 children).

**TABLE 5**

Comparison of Unadjusted and Adjusted Odds Ratios of Autism Spectrum Disorders Among Singleton Pregnancies in the Autism Birth Cohort Study with the Same Unadjusted and Adjusted Odds Ratios in the Medical Birth Registry of Norway

| Exposures[a] | Medical Birth Registry of Norway | | Autism Birth Cohort Study | | Ratio of Unadjusted OR [95% CI][c] | Ratio of Adjusted OR [95% CI][c] |
| --- | --- | --- | --- | --- | --- | --- |
| | Unadjusted OR [95% CI][b] | Adjusted OR [95% CI][b] | Unadjusted OR [95% CI][b] | Adjusted OR [95% CI][b] | | |
| Primipara pregnancy | 1.27 [1.16, 1.39] | 1.23 [1.11, 1.35] | 1.40 [1.08, 1.82] | 1.41 [1.06, 1.89] | 1.10 [0.86, 1.42] | 1.15 [0.88, 1.51] |
| Prenatal folic acid use | 0.66 [0.60, 0.73] | 0.84 [0.76, 0.93] | 0.70 [0.54, 0.90] | 0.83 [0.63, 1.10] | 1.06 [0.82, 1.34] | 1.00 [0.77, 1.28] |
| Prenatal smoking | 1.31 [1.17, 1.46] | 1.17 [1.04, 1.33] | 1.37 [0.97, 1.95] | 1.20 [0.84, 1.72] | 1.05 [0.74, 1.41] | 1.02 [0.71, 1.38] |
| Preterm birth (<37 weeks) | 1.57 [1.34, 1.85] | 1.49 [1.27, 1.76] | 1.82 [1.14, 2.91] | 1.67 [1.04, 2.68] | 1.16 [0.68, 1.71] | 1.11 [0.66, 1.66] |
| Low birthweight (<2,500 g) | 1.85 [1.53, 2.22] | 1.70 [1.41, 2.05] | 2.08 [1.19, 3.64] | 1.84 [1.05, 3.24] | 1.12 [0.58, 1.77] | 1.08 [0.56, 1.73] |
| Male sex | 4.63 [4.12, 5.21] | | 4.72 [3.33, 6.69] | | 1.02 [0.75, 1.48] | |
| Cesarean section history[d] | 0.95 [0.79, 1.13] | 0.98 [0.82, 1.17] | 0.99 [0.57, 1.70] | 1.00 [0.58, 1.73] | 1.04 [0.54, 1.67] | 1.02 [0.53, 1.64] |

Abbreviations: ASD, autism spectrum disorders; OR, odds ratio; CI, confidence interval.

[a] Reference categories for odds ratios were multipara pregnancy, no prenatal folic acid use, no prenatal smoking, term birth ( 37 weeks), normal birthweight ( 2,500 g), female sex, and no cesarean section history.

[b] The odds ratios were estimated using logistic regression models and adjusted for the covariates as described in the Methods section. Missing exposure and covariate data were imputed by using a multiple imputation method.

[c] The ratio of the odds ratio in the Autism Birth Cohort study over the odds ratio in the Medical Birth Registry of Norway; 95% confidence interval was corrected for inter-dependency between samples by using a non-parametric bootstrap method.

[d] Analyses included only women with one or more previous births (Medical Birth Registry of Norway: 1,060 ASD cases among 290,311 children; Autism Birth Cohort study: 108 ASD cases among 48,883 children).