nature medicine

Article

https://doi.org/10.1038/s41591-024-03479-5

Familial confounding in the associations between maternal health and autism

Received: 17 March 2024

Accepted: 18 December 2024

Published online: 31 January 2025

Check for updates

Vahe Khachadourian[1], Elias Speleman Arildskov[2], Jakob Grove[2], Paul F. O'Reilly[3], Joseph D. Buxbaum[3,4], Abraham Reichenberg[4,5], Sven Sandin[4,6], Lisa A. Croen[7], Diana Schendel[8,9,10], Stefan Nygaard Hansen[11] & Magdalena Janecka[1,12]

Evidence suggests that maternal health in pregnancy is associated with autism in the offspring. However, most diagnoses in pregnant women have not been examined, and the role of familial confounding remains unknown. Our cohort included all children born in Denmark between 1998 and 2015 (n = 1,131,899) and their parents. We fitted Cox proportional hazard regression models to estimate the likelihood of autism associated with each maternal prenatal ICD-10 diagnosis, accounting for disease chronicity and comorbidity, familial correlations and sociodemographic factors. We examined the evidence for familial confounding using discordant sibling and paternal negative control designs. Among the 1,131,899 individuals in our sample, 18,374 (1.6%) were diagnosed with autism by the end of follow-up. Across 236 maternal diagnoses we tested (prevalence ≥0.1%), 30 were significantly associated with autism after accounting for sociodemographic factors, disorder chronicity and comorbidity, and correction for multiple testing. This included obstetric, cardiometabolic and psychiatric disorders (for example, diabetes in pregnancy (hazard ratio (HR) 1.19, 95% confidence interval (CI) 1.08–1.31) and depression (HR 1.49, 95% CI 1.27–1.75)), previously shown to be associated with autism. Family-based analyses provided strong evidence for familial confounding in most of the observed associations. Our findings indicate pervasive associations between maternal health in pregnancy and offspring autism and underscore that these associations are largely attributable to familial confounding.

Autism is a developmental condition typically diagnosed in early childhood and characterized by differences in social communication and restrictive and repetitive behaviors. Due to its early neurodevelopmental origins and the gestational link between the mother and the fetus, research on nongenetic factors associated with autism has focused on perinatal exposures[1]. For brevity, throughout the article, we use 'autism' to refer to a diagnosis of autism spectrum disorder (ASD).

Epidemiological studies have demonstrated associations between autism in the child and several maternal conditions around the time of pregnancy, including depression[2], diabetes[3], immune system diseases[4] and infections[5]. However, multiple gaps in the knowledge remain; for example, pregnant women experience many other health problems, most of which have not been studied in relation to autism in the offspring and whose co-occurrence[1,6,7] with each other during pregnancy has not been accounted for to date. Furthermore, even for the established associations between certain maternal diagnoses and autism, causal interpretations remain challenging.

Maternal health conditions may be associated with autism in the offspring due to their direct effects on the fetus (for example, physiological changes related to disease, medication use). Such direct

A full list of affiliations appears at the end of the paper. e-mail: magdalena.janecka@nyulangone.org

effects are compatible with the early neurodevelopmental origins of autism and are biologically plausible—multiple factors from the maternal circulation can pass through the placenta[8], and the blood–brain barrier develops gradually during gestation[9]. Nevertheless, these observational associations can also arise because of familial factors, both genetic (genetic variation shared by family members) and nongenetic (for example, family residence in polluted areas, socioeconomic status). Evidence for potential genetic confounding in the associations between maternal health and offspring autism comes from the established genetic overlap between autism and other conditions (both psychiatric and somatic[10]), documented through a higher prevalence of those disorders in relatives of individuals with autism[11] and genotype data from unrelated individuals[10]. Importantly, such genetic confounding can arise from the transmission of maternal alleles to the child and the indirect genetic effects of maternal genotype on the in utero environment[12]. Attributing the observational associations to either direct effects of the disease on the fetus or familial confounding is critical to understanding the causes of autism and informing the public about the true role of maternal health factors in autism etiology.

Various study designs can help elucidate the role of familial confounding in observational evidence. Sibling designs can control for family-level and sibling-invariant confounders without the need to explicitly define or measure them[13]. Additionally, the use of negative controls (for example, paternal exposures during pregnancy) can identify unmeasured familial confounding linked to maternal exposures and other types of biases[14]. Complementing standard observational studies with the use of these specialized study designs is critical to providing insights into the mechanisms underlying autism etiology and discerning the degree to which the identified factors represent noncausal associations.

Here, we leveraged a large, population-based cohort from Denmark with extensive health information and family linkage recorded through national registers. We used this resource to assess in a comprehensive manner the associations between maternal diagnoses around pregnancy and offspring autism, with adjustments for a broad range of demographic and socioeconomic factors. We accounted for the comorbidity and chronicity of maternal conditions and explored their associations with autism according to the offspring's sex and intellectual disability (ID) status. Finally, we complemented our cohort study with designs aimed at assessing the impact of unmeasured familial confounding on the observed associations and performed extensive sensitivity analyses to assess the robustness of our conclusions. Together, our results provide a comprehensive atlas of the associations between the full breadth of maternal ICD-10 (International Classification of Diseases, tenth revision) diagnoses and offspring autism and systematically reveal the potential mechanisms underlying these associations.

Results

The study sample included 1,131,899 children born (to 648,901 mothers) in Denmark between 1998 and 2015, of whom 1,074,756 (95%) had full covariate data (see Table 1 for the missingness percentage of each covariate). A total of 18,374 (1.6%) children received an ASD diagnosis during the study follow-up period. The median duration of follow-up was 9.7 years (interquartile range (IQR) 5.3–14.3 years). The median age at the first ASD diagnosis was 8.3 years (IQR 5.4–11.8 years) in the full sample and 5.6 years (IQR 4.2–6.8 years) in the subset of children under 8 years old with full follow-up data (consistent with methods reported in surveillance studies[15–17]). Over the cohort years, the age at diagnosis decreased, reflecting both the temporal trends toward an earlier diagnosis[16,18,19] and the shorter follow-up durations available for children born later in the cohort years (see Supplementary Table 1 for age at diagnosis by birth year). Table 1 presents the sample characteristics and rate of covariate missingness by ASD diagnosis status.

Associations between maternal diagnoses and offspring autism

There were 1,613 and 1,702 distinct ICD-10 registry-reported level 3 codes (diagnoses) in the mothers of cohort children during the 12 and 48 months preceding childbirth, respectively. Based on the prevalence of autism in our sample and the frequency thresholds for the diagnosis to be included in the analyses (20 for chronic and 10 for nonchronic, in each individual with or without autism), we analyzed chronic and nonchronic maternal diagnoses with a prevalence of at least 0.1% and 0.05%, respectively—amounting to 168 nonchronic and 68 chronic diagnoses (Extended Data Fig. 1; see Supplementary Tables 2 and 3 for the number of exposed and unexposed children in the sample by ASD status). Among diagnoses meeting those thresholds in the primary analyses, further exclusions were made in the analyses stratified by offspring sex and ID status due to insufficient numbers of certain exposures in these strata.

After controlling for the study covariates, 37 of the 168 nonchronic maternal diagnoses were nominally significantly associated with autism ($P < 0.05$), 20 of which remained significant after adjusting for multiple testing ($q < 0.05$) (Fig. 1 and Supplementary Table 2). Similarly, 32 of the 68 chronic maternal diagnoses were nominally significantly associated with offspring autism, and 22 remained significant after adjusting for multiple testing ($q < 0.05$) (Fig. 1 and Supplementary Table 3). In total, 42 maternal diagnoses were associated with autism after controlling for covariates and multiple testing adjustments. These included psychiatric conditions (for example, major depressive disorder (hazard ratio (HR) 2.06, 95% confidence interval (CI) 1.78–2.39)), obstetric disorders (for example, premature rupture of membranes (HR 1.13, 95% CI 1.05–1.21), false labor (HR 1.23, 95% CI 1.15–1.31)) and cardiometabolic conditions (for example, diabetes mellitus in pregnancy (HR 1.23, 95% CI 1.12–1.36), primary hypertension (HR 1.34, 95% CI 1.08–1.67)), as well as other diagnoses spanning most other diagnostic categories (fracture of the skull and facial bones (HR 1.98, 95% CI 1.28–3.05), epilepsy and recurrent seizures (HR 1.38, 95% CI 1.16–1.64), breast cancer (HR 1.98, 95% CI 1.25–3.13)). The full overview of all 236 associations following iterative adjustment for potential confounders is presented in Supplementary Tables 2 (nonchronic) and 3 (chronic).

In the multidiagnosis model adjusting for covariates and all 42 maternal diagnoses that were statistically significant in the fully adjusted single-diagnosis models, 30 diagnoses remained statistically significantly associated with autism (15 chronic and 15 nonchronic conditions; Fig. 2 and Supplementary Table 4). Notably, of the 12 diagnoses that were no longer significant after this comorbidity adjustment, half (6) were in the ICD-10 F (psychiatric) category. The full overview of all 42 associations and the coefficients of the covariates are presented in Supplementary Table 4.

Results stratified by the child's sex are presented in Extended Data Fig. 2 (throughout these analyses, we considered biological sex as reported in the registers). In all sex-specific analyses, the CIs around the male- and female-specific estimates overlapped, and we were not able to conclude sex-specific effects for any of the maternal diagnoses. Given the lower prevalence of autism among female individuals (0.8%), we were able to detect an HR of 1.5 with a power of 80% for diagnoses present in 1.1% of pregnancies and an HR of 2.0 for diagnoses present in 0.4% of pregnancies (see the power calculations in Extended Data Fig. 2).

Stratifying the analyses by the child's ID status, we observed that the associations between maternal diagnoses and autism were predominantly driven by associations with autism without ID (Extended Data Fig. 3), although we were not adequately powered to conclude significant differences in the estimates in those strata. The power analyses indicated that we had statistical power (80%) to uncover effects at an HR of 1.5 and 2.0 for diagnoses present in 1.6% and 0.6%, respectively, as indicated in Extended Data Fig. 3. The only maternal diagnosis significantly associated with autism both with and without ID was injury to an unspecified body region.

https://doi.org/10.1038/s41591-024-03479-5

Table 1 | Demographic characteristics of the analytical sample (mother–child dyads)

Variables	Diagnosed with autism (*n*=18,374)	No autism diagnosis (*n*=1,113,525)	Total (*n*=1,131,899)
Maternal age (years) at delivery, mean (s.d.)	30.3 (5.1)	30.6 (4.9)	30.6 (4.9)
Child's sex, *n* (%)			
Female	4,509 (24.5%)	546,916 (49.1%)	551,425 (48.7%)
Male	13,865 (75.5%)	566,609 (50.9%)	580,474 (51.3%)
Child's year of birth, *n* (%)			
1998	1,750 (9.5%)	64,381 (5.8%)	66,131 (5.8%)
1999	1,875 (10.2%)	64,365 (5.8%)	66,240 (5.9%)
2000	1,936 (10.5%)	65,175 (5.9%)	67,111 (5.9%)
2001	1,754 (9.5%)	63,695 (5.7%)	65,449 (5.8%)
2002	1,631 (8.9%)	62,484 (5.6%)	64,115 (5.7%)
2003	1,526 (8.3%)	63,133 (5.7%)	64,659 (5.7%)
2004	1,333 (7.3%)	63,358 (5.7%)	64,691 (5.7%)
2005	1,237 (6.7%)	63,123 (5.7%)	64,360 (5.7%)
2006	1,147 (6.2%)	63,891 (5.7%)	65,038 (5.7%)
2007	1,010 (5.5%)	63,200 (5.7%)	64,210 (5.7%)
2008	838 (4.6%)	64,248 (5.8%)	65,086 (5.8%)
2009	732 (4.0%)	62,203 (5.6%)	62,935 (5.6%)
2010	648 (3.5%)	62,879 (5.6%)	63,527 (5.6%)
2011	460 (2.5%)	58,662 (5.3%)	59,122 (5.2%)
2012	289 (1.6%)	57,762 (5.2%)	58,051 (5.1%)
2013	142 (0.8%)	55,814 (5.0%)	55,956 (4.9%)
2014	53 (0.3%)	56,874 (5.1%)	56,927 (5.0%)
2015	13 (0.1%)	58,278 (5.2%)	58,291 (5.1%)
Maternal income in DKK the year before delivery, mean (s.d.)	203,452 (110,107)	235,762 (214,650)	235,237 (213,400)
Missing, *n* (%)	103 (0.6%)	7,844 (0.7%)	7,947 (0.7%)
Paternal income in DKK the year before delivery, mean (s.d.)	276,857 (196,233)	316,626 (350731)	315,980 (348,806)
Missing, *n* (%)	341 (1.9%)	20,864 (1.9%)	21,205 (1.9%)
Maternal education the year before delivery, *n* (%)			
Primary or lower secondary level	4,828 (26.3%)	211,468 (19.0%)	216,296 (19.1%)
Upper secondary level and secondary vocational education	7,844 (42.7%)	447,248 (40.2%)	455,092 (40.2%)
Short-cycle tertiary, bachelor or equivalent	3,865 (21.0%)	292,433 (26.3%)	296,298 (26.2%)
Master, doctoral or equivalent	1,217 (6.6%)	106,567 (9.6%)	107,784 (9.5%)
Missing, *n* (%)	620 (3.4%)	55,809 (5.0%)	56,429 (5.0%)
Paternal education the year before delivery, *n* (%)			
Primary or lower secondary level	4,828 (26.3%)	211,468 (19.0%)	216,296 (19.1%)
Upper secondary level and secondary vocational education	7,844 (42.7%)	447,248 (40.2%)	455,092 (40.2%)
Short-cycle tertiary, bachelor or equivalent	3,865 (21.0%)	292,433 (26.3%)	296,298 (26.2%)
Master, doctoral or equivalent	1,217 (6.6%)	106,567 (9.6%)	107,784 (9.5%)
Missing, *n* (%)	620 (3.4%)	55,809 (5.0%)	56,429 (5.0%)
Number of days of maternal healthcare visits the year leading up to delivery, *n* (%)			
0	270 (1.5%)	14,837 (1.3%)	15,107 (1.3%)
1–3	13,189 (71.8%)	844,594 (75.8%)	857,783 (75.8%)
4–9	4,639 (25.2%)	241,813 (21.7%)	246,452 (21.8%)
10+	276 (1.5%)	12,281 (1.1%)	12,557 (1.1%)
Child's parity, *n* (%)			
1	9,477 (51.6%)	507,963 (45.6%)	517,440 (45.7%)
2–3	8,217 (44.7%)	558,945 (50.2%)	567,162 (50.1%)
4–5	634 (3.5%)	42,871 (3.9%)	43,505 (3.8%)

https://doi.org/10.1038/s41591-024-03479-5

Table 1 (continued) | Demographic characteristics of the analytical sample (mother–child dyads)

Variables	Diagnosed with autism (*n*=18,374)	No autism diagnosis (*n*=1,113,525)	Total (*n*=1,131,899)
6+	46 (0.3%)	3,746 (0.3%)	3,792 (0.3%)
Child has ID, *n* (%)			
No	15,578 (84.8%)	1,108,855 (99.6%)	1,124,433 (99.3%)
Yes	2,796 (15.2%)	4,670 (0.4%)	7,466 (0.7%)
Mother has at least two children in the cohort, *n* (%)			
No	5,229 (28.5%)	275,100 (24.7%)	280,329 (24.8%)
Yes	13,145 (71.5%)	838,425 (75.3%)	851,570 (75.2%)

Biological sex, as documented in the registers, is reported. DKK, Danish krone.

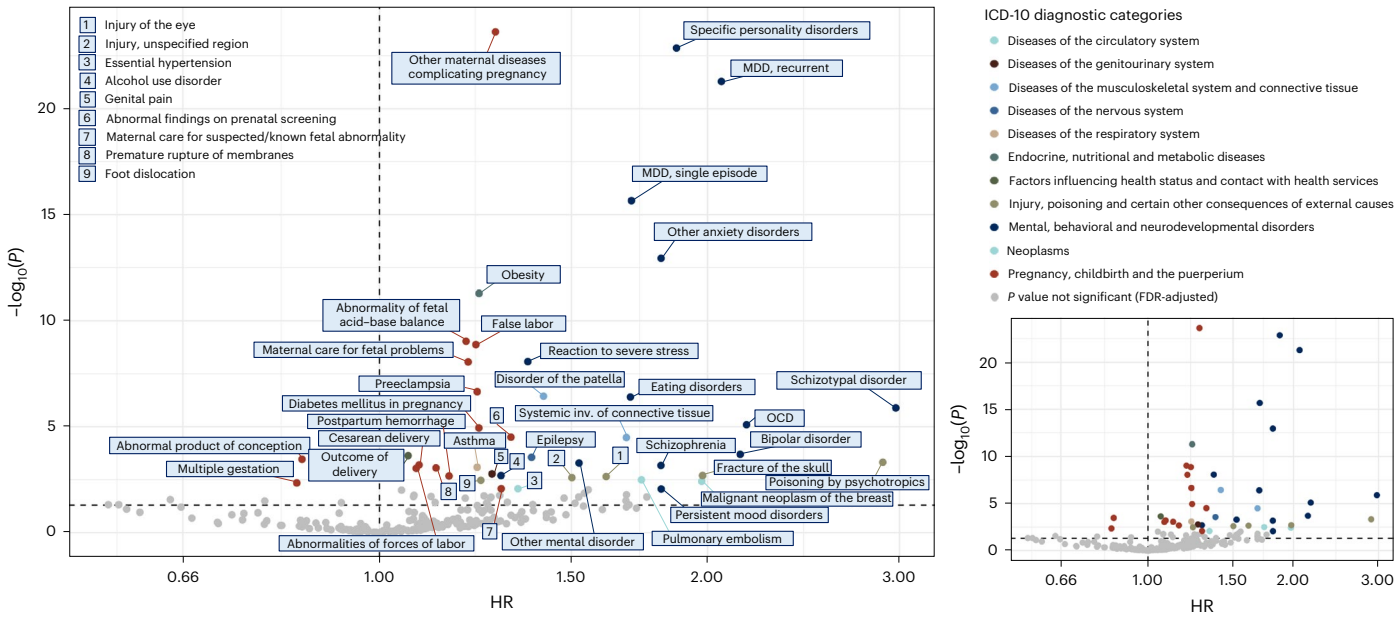


Fig. 1 | Associations between ICD-10 level 3 maternal diagnoses and offspring autism in fully adjusted single-diagnosis models. Point estimates of each association derived from the two-sided Cox proportional hazard model for each diagnosis are illustrated on the *x* axis, with their *P* value ($-\log_{10}(P)$) on the *y* axis. Dots representing each statistically significant association are colored according to the ICD-10 category of the respective diagnosis; nonsignificant associations after correction for multiple testing are shown in gray. The horizontal dashed line represents the *P*-value cutoff for nominal significance ($P = 0.05$). The nonannotated plot on the right presents the same data for a clearer visualization of the coefficient distribution. MDD, major depressive disorder; inv., involvement; OCD, obsessive–compulsive disorder.

The sensitivity analyses demonstrated that our results were robust to our analytical decisions, including the definition of healthcare utilization (Supplementary Tables 5 and 6: models 1–3); restricting the sample to children born by the end of 2009 (that is, with at least 8 years of follow-up data; Supplementary Tables 5 and 6: model 4); varying the length of the exposure period (Supplementary Table 7: models 5 and 6); requiring the presence of two diagnoses to ascertain exposure (Supplementary Table 7: model 7); adjusting for paternal income (Supplementary Table 8: model S8), parity and multiplicity (Supplementary Table 9: models S9–S11); and excluding mothers with an autism diagnosis (Supplementary Table 9: model S12). The key differences between the main and sensitivity analyses included the impact of not adjusting for healthcare utilization, which resulted in many more significant associations (Supplementary Tables 5 and 6: model 3); diagnoses that were significant when not accounting for this variable were enriched for nonspecific conditions ('other', 'unspecified' and 'not elsewhere classified' disorders) or mild conditions (for example, purpura, cystitis, gastroesophageal reflux disease), as well as symptoms (ICD R codes). Moreover, despite the overall close alignment between regression coefficients in models with different lengths of exposure windows, four diagnoses were significant only in models with a longer (72 months) exposure period (migraine, other headache syndromes, osteoarthritis of the knee, paroxysmal tachycardia; Supplementary Table 7: model 6S).

Evidence for familial confounding in the associations between maternal health and autism

Sibling analysis. The sample in the sibling analysis consisted of 851,570 children (the total number of cohort children with a sibling in the cohort). Of the total of 18,374 children with autism in the sample, 12,138 (66%) were from families with siblings discordant for autism status (see Supplementary Table 10 for the covariate distribution in this subset of families). Among those, different numbers of sibling pairs were discordant for both maternal diagnosis and autism status, ranging from 12 to 6,319 pairs for different diagnoses (mean of 749 discordant pairs; Supplementary Table 11).

In the sibling models, we observed a widespread attenuation of the point estimates associated with maternal diagnosis and widening of the CIs, resulting in the loss of statistical significance in most associations (Fig. 3 and Supplementary Table 11). The median change in the point estimates in the sibling models for nonchronic and chronic diagnoses was, respectively, 38% and 71% overall and 51% and 71% when we considered only the associations for which the effects moved closer to the null in the sibling model (Supplementary Table 11).

https://doi.org/10.1038/s41591-024-03479-5

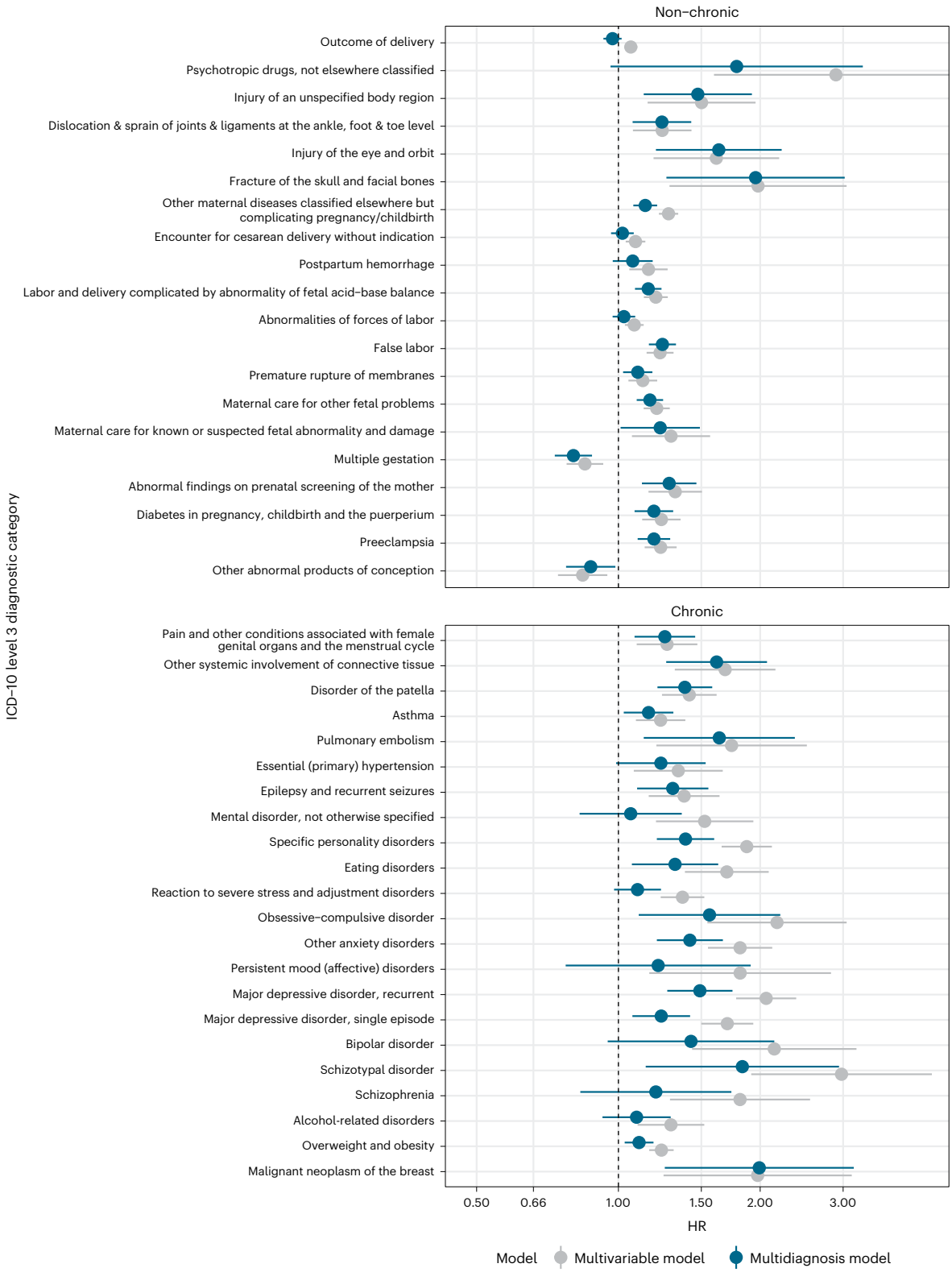


Fig. 2 | Associations between ICD-10 level 3 maternal diagnoses and offspring autism in fully adjusted single-diagnosis models and the multidiagnosis model. Point estimates are HRs adjusted for maternal age at childbirth, child's sex and year of birth, maternal income and education, and maternal healthcare utilization in the 12 months preceding childbirth. Estimates from the multidiagnosis model, in addition to the covariates above, are concurrently adjusted for all significant diagnoses (nonchronic and chronic) in fully adjusted single-diagnosis models (presented in this figure). The error bars represent 95% CIs calculated using point estimates and robust standard errors from the respective regression model.

Results from sibling analysis restricting the sample to male children only were consistent with the results of the main sibling analysis (Supplementary Table 11). However, despite the large sample size, we had low power for some of the diagnoses in the sibling analyses, which resulted in wide CIs, limiting the certainty of the conclusions for some of the diagnoses.

https://doi.org/10.1038/s41591-024-03479-5

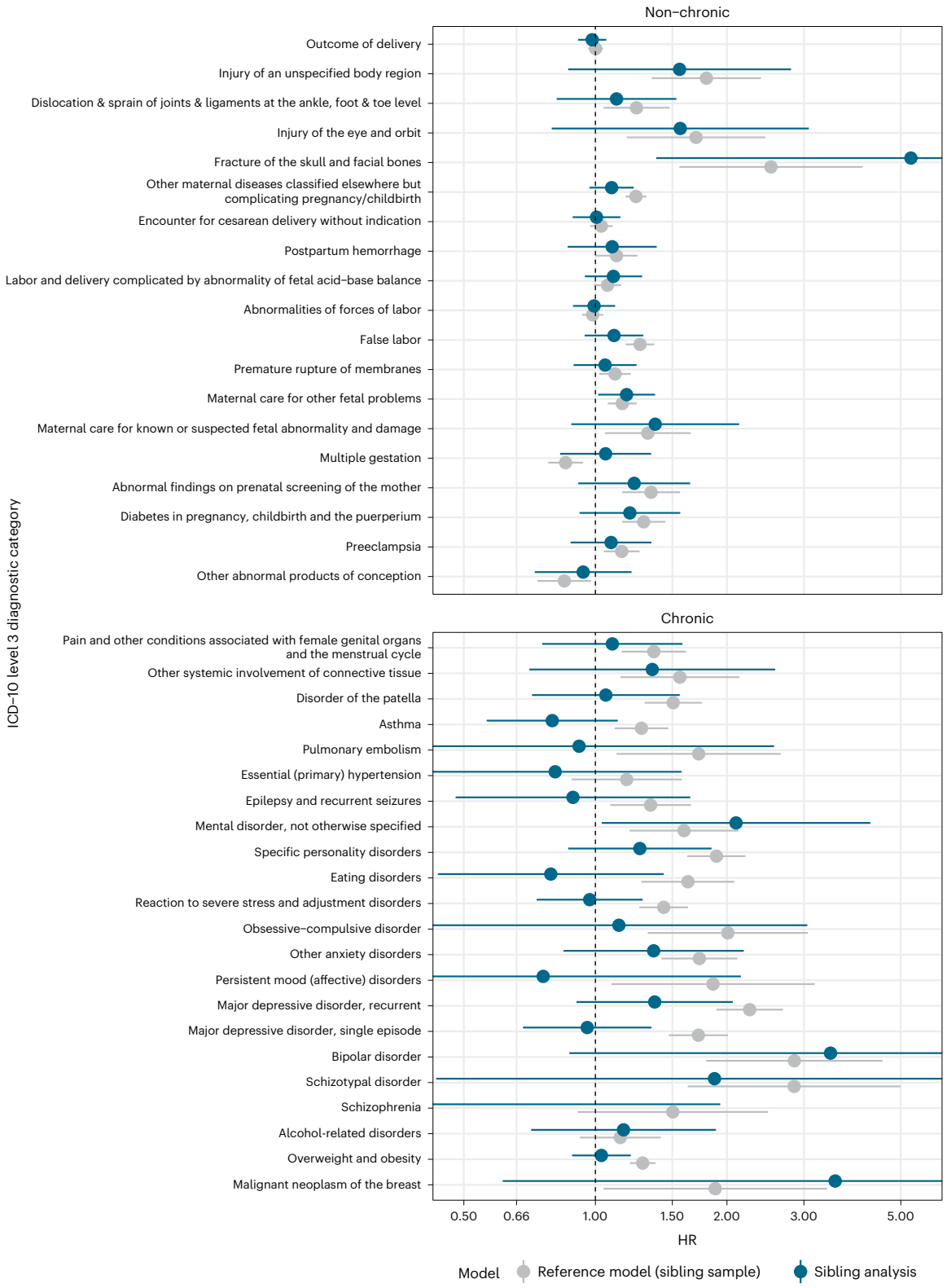


Fig. 3 | Associations between ICD-10 level 3 maternal diagnoses and offspring autism from reference models (single-diagnosis models in the sibling sample) and sibling analysis (stratified by family identification number). Point estimates for each diagnosis are HRs from a model adjusted for maternal age at childbirth, child's sex and year of birth, maternal income and education, and maternal healthcare utilization in the 12 months preceding childbirth. All analyses presented in this figure were restricted to individuals with at least one sibling in the sample (851,570 of the 1,131,899 mother–child dyads in the full birth cohort). The potential differences between the results from the single-diagnosis analysis in the subsample of siblings only and the full sample are likely attributable to the potential differences in sample composition and sample size. Due to the extremely low number of sibling pairs discordant for maternal schizophrenia status and autism (19 pairs), the point estimate from the sibling analysis is out of bounds (HR 0.34, 95% CI 0.06–1.93). The error bars represent 95% CIs calculated using point estimates and robust standard errors from the respective regression model.

https://doi.org/10.1038/s41591-024-03479-5

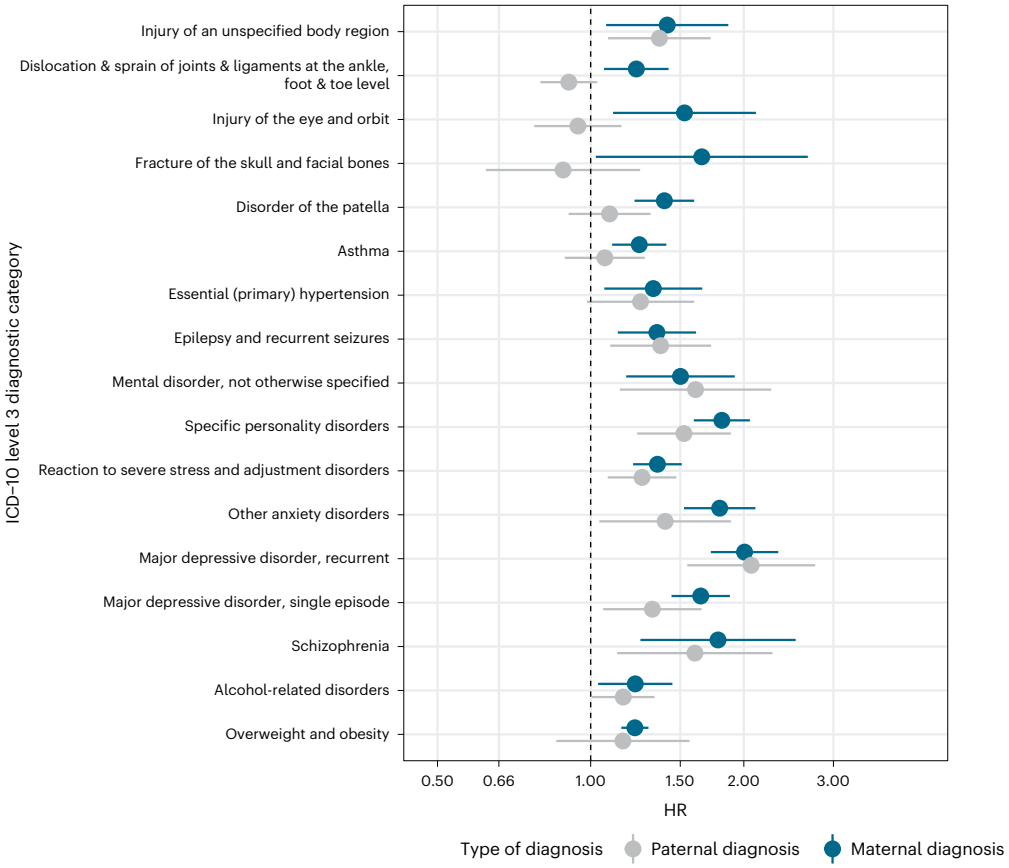


Fig. 4 | Associations between ICD-10 level 3 paternal and maternal diagnoses and offspring autism from single-diagnosis analyses. Point estimates for each diagnosis are HRs from a model adjusted for maternal and paternal ages at childbirth, child's sex and year of birth, maternal and paternal income and education, and maternal and paternal healthcare utilization in the 12 months preceding childbirth, as well as the same diagnosis in the other parent. The error bars represent 95% CIs calculated using point estimates and robust standard errors from the respective regression model.

Paternal analysis. Most of the associations between paternal diagnoses and offspring autism were similar to the associations observed between the corresponding maternal diagnoses and offspring autism (Supplementary Table 12), except for injuries, joint conditions and asthma (Fig. 4). Adjustment for paternal diagnosis resulted in subtle lowering of the point estimates associated with maternal exposure (Supplementary Table 11); however, these differences were not statistically significant.

Figure 5 summarizes the results of comorbidity, sibling and paternal analyses and provides a higher-level summary of the impact of comorbidity adjustment in the multidiagnosis model and familial confounding in estimated effects.

Discussion

In this population-based study of more than 1.1 million pregnancies, we evaluated the associations between a range of maternal diagnoses and autism in the offspring. Using national registry data from Denmark, we have shown that among 236 maternal diagnoses affecting all body systems, 30 were associated with autism in the offspring after adjusting for comorbidity and socioeconomic and demographic covariates and accounting for disorder chronicity and familial correlations. These results remained robust in a series of sensitivity analyses, with no evidence that any of the associations differed by the child's sex. Extensive investigation of familial confounding revealed that most of the observational associations are attributable to family-level factors, challenging the idea that those maternal diagnoses have direct causal effects on autism.

Results from family-based analyses in our study suggest substantial familial confounding in the observational associations between maternal health and child autism. With a few exceptions (discussed below), the diagnoses analyzed in the sibling comparison were not associated with autism within families, and the point estimates in those analyses were frequently attenuated compared to the model in the sibling sample that did not account for familial relationships. For diagnoses for which we observed such attenuation in the sibling models, the median change was 51% for maternal nonchronic conditions and 71% for chronic conditions. Therefore, these results suggest confounding by family-level shared factors that were uncontrolled in the cohort analyses not accounting for familial relationships. For instance, a reduction in the point estimates by ~50%, observed for, for example, premature rupture of membranes or false labor, could indicate confounding of the association between maternal diagnosis and autism by shared genetic factors—which, on average, overlap by 50% between siblings. The presence of such genetic confounding is supported by numerous molecular studies indicating that the genetic variants contributing to autism are also implicated in other traits and disorders[10,20–25] and by family-based approaches demonstrating that autism and these other conditions cluster in families[26,27]. Diagnoses associated with a larger extent of point estimate attenuation in sibling analyses included, for example, multiple gestation and obesity—and these effects are consistent with confounding by shared familial factors that fully overlap between siblings reared together (for example, area pollution, in utero environment)[28,29]. Nevertheless, as discussed by Frisell[30], the reduction in point estimates could also be caused by random measurement errors, which disproportionately affect sibling versus cohort analyses due to the former's reliance on discordance for exposure and outcome. While the excellent validity of diagnostic data in the Danish registers reduces such error—that is, the likelihood

https://doi.org/10.1038/s41591-024-03479-5

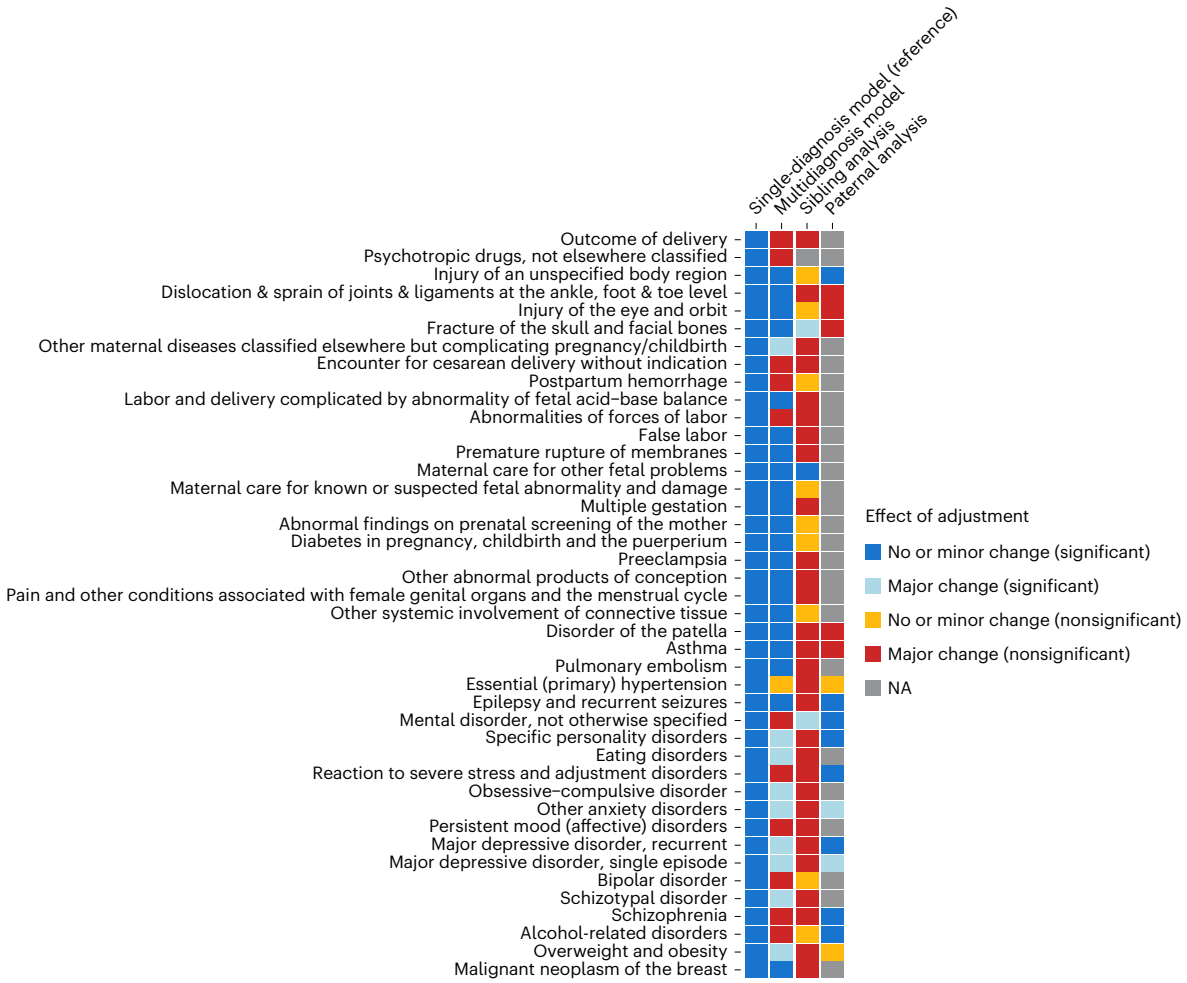


Fig. 5 | Overview of the study results. Tile colors capture the statistical significance of the association between the given maternal diagnosis and offspring autism and the change in the point estimate relative to the reference model ('minor' change refers to an HR change of ≤40%, and 'major' change refers to an HR change of >40%; the numerical parameters underlying this categorization are presented in Supplementary Tables 2–4, 11 and 12). Associations not estimated due to insufficient frequency of the diagnosis in the analytical subsample, or other methodological considerations, are presented as 'NA'. The interpretation of the tile color changes by analysis type or column. In the multidiagnosis model, red and light blue tiles indicate diagnoses whose association with autism was reduced after concurrent adjustment for comorbid conditions, suggesting diagnoses whose original association with autism is most likely to have arisen due to comorbidity with other diagnoses. In sibling analysis, red tiles indicate diagnoses whose point estimate shows a relative change of >40% when compared to the point estimate from the reference model, suggesting familial confounding (that is, $(HR_{sib} - 1)/(HR_{ref} - 1) < 0.6$ or $(HR_{sib} - 1)/(HR_{ref} - 1) > 1.4$); yellow tiles indicate diagnoses that may be associated with autism within families, but our analyses were underpowered to detect such effects (the loss of significance was due to widening of the CIs, with little change in the point estimate; for example, injuries of the eye and an unspecified body region, diabetes in pregnancy). In paternal analysis, red tiles indicate diagnoses associated with autism in the mother, but not in the father, suggesting a lack of evidence for familial confounding (consistent with either direct effects of those conditions on the fetus or indirect genetic effects; for example, injury codes, disorders of the patella and asthma); dark and light blue tiles suggest evidence for familial confounding (without and with evidence for additional contributions of a maternal effect, respectively; for example, direct effects on the fetus and/or indirect genetic effects). Reference models for sibling analysis and the paternal model are provided in Supplementary Tables 11 (model 5) and 12 (model 5), respectively.

that pairs labeled as discordant are, in fact, concordant—it could still affect the comparison between cohort and sibling analyses. This would have the largest potential impact on the chronic conditions, in which misclassification is most likely due to difficulties in the ascertainment of disease onset and remission using registry data and for which we observed the fewest pairs of siblings discordant on both exposure and outcome; therefore, we urge caution in interpreting those effects for the chronic conditions in our and other studies (however, we also note that for those diagnoses, an orthogonal paternal analysis provides supporting evidence for the presence of familial confounding).

For a few associations, the loss of significance in sibling analyses arose due to the widening of the CIs rather than changes in the point estimates attributable to the additional control for familial factors; for those diagnoses (for example, diabetes in pregnancy), the evidence of familial confounding is less clear. For certain other diagnoses, the point estimates also moved further away rather than closer to the null in the sibling models (for example, labor and delivery complicated by abnormality of fetal acid–base balance). Many of the diagnoses associated with the least change in the point estimates in the sibling models and those associated with a change away from the null included conditions diagnosable in the mother but pertaining largely to the fetus (labor and delivery complicated by abnormality of fetal acid–base balance, maternal care for other fetal problems, maternal care for fetal abnormality and damage). While these conditions could be early manifestations of autism, rather than its etiological factors, such possibility—along with potential etiologic contributions of, for example, gestational diabetes—remains to be investigated.

Consistent with attenuation of the point estimates in the sibling analyses, most of the diagnoses that can be made in cis fathers were associated with similar risk estimates irrespective of the parent of origin; as paternal diagnoses have limited direct effects on the fetus during pregnancy, this pattern of results is consistent with confounding by familial factors, in line with the observations in the sibling comparison models. While the inclusion of paternal diagnoses led to a subtle attenuation of the point estimates associated with maternal exposure, these decreases were small and their interpretation is complicated by diverse sources of familial confounding and the potential impact of assortative mating[31,32]. Such a decrease would be expected in situations in which parental diagnoses represent the same familial risk factor (for example, area pollution)[33,34]; instead, our results are consistent with these diagnoses being proxies for genetic liabilities—mostly independent at the parental level but shared at the child level. Nevertheless, we cannot discount the possibility of different sources of confounding for maternal and paternal conditions or the presence of a parentally shared confounder whose effects are masked by statistical fluctuation. As a parental negative control approach relies on the assumption that the sources of confounding between paternal exposures and outcome overlap with those on the maternal side (that is, there are no paternal-specific sources of confounding), the utility of this approach is limited by the extent to which this assumption holds[35]. Additionally, the inclusion of paternal exposure as a negative control in transgenerational studies remains a topic of scientific discussion[36,37] due to the potential effects of paternal exposures on the sperm and the knock-on effect of these germline changes on offspring development[38,39]. However, while epigenetic changes in the sperm have been observed in association with certain paternal conditions occurring before conception[40,41], there is currently no evidence that such effects can persist through two rounds of epigenetic reprogramming in humans occurring after fertilization[42]; such evidence (or validation of the opposite) will be critical for future reconsideration of the paternal negative control approach.

The few maternal diagnoses with no evidence for paternal effects included injury, joint conditions (patella disorder) and asthma, providing preliminary evidence that those diagnoses could be associated with autism either due to their direct effects on the fetus (including, for example, effects of medications taken for these conditions) or indirect genetic influences with (pleiotropic) effects on both maternal diagnosis and the in utero environment[43,44]. Further research incorporating broader family pedigrees and genotype data will be needed to distinguish between these mechanisms.

As in previous studies, complications of pregnancy[45] represented one of the strongest perinatal factors for autism in our sample. As discussed above, most of these diagnoses did not remain significantly associated with autism when comparing sibling pairs discordant for exposure (that is, siblings born to the same mother, after pregnancies with and without those complications), and the point estimates were substantially attenuated in sibling analyses—suggesting the lack of direct causal effects of those pregnancy complications on autism. The obstetric complications for which we did not observe such attenuation include gestational diabetes and postpartum hemorrhage; as we noted above, for those, the presence of familial confounding is less clear.

Likewise, psychiatric and neurological diagnoses, linked to offspring autism in previous studies[1,46], were associated with an increased likelihood of autism when recorded in either the mother or father, and their estimates were substantially attenuated within families—again, indicating familial confounding. The estimates associated with some of the psychiatric diagnoses (for example, personality disorder, anxiety disorders) were higher when the diagnosis was recorded in the mother rather than in the father; although those differences were not statistically significant, we cannot fully discount a potential additional contribution of indirect genetic and/or direct causal effects to the observed associations (including through the effects of medications). Importantly, these familial and nonfamilial mechanisms are not mutually exclusive, highlighting the complex architecture of the risk of neurodevelopmental conditions.

Discerning the extent to which the observational associations between perinatal factors and neurodevelopment are driven by confounding familial factors is critical from the perspective of multiple stakeholders, including pregnant women and their healthcare providers, who should receive reliable information regarding the safety of different exposures in pregnancy; families and caregivers of individuals with autism, who deserve clarity on the role of modifiable perinatal factors in autism etiology; and the research community. While our results diminish the possibility of direct causal contributions of many of these maternal diagnoses to autism and implicate the contributions of underlying familial factors, these findings still offer important etiological insights. 'Familial confounding' should not be conflated with 'irrelevant'. Instead, a more granular understanding of the nature of this confounding—including what familial factors, genetic and nongenetic, are confounding the observational associations and how they influence neurodevelopment—will be critical for a fuller understanding of autism etiology.

Our study offers a comprehensive assessment of the associations between autism and a wide range of maternal diagnoses around pregnancy. The systematic nature of our analysis and ascertainment of all maternal diagnoses with sufficient prevalence for robust effect estimation (~0.1% of pregnant women) enabled a direct comparison of the strength of the estimated associations and allowed us to account for the co-occurrence of these conditions. Additional strengths of this study include the classification of diagnoses into the chronic and nonchronic categories, the accounting for familial confounding in the sibling and paternal analyses, and the availability of extensive demographic, socioeconomic and medical data, which enabled us to implement careful covariate control and extensive sensitivity analyses.

This study also has several limitations. First, we did not account for potential confounding by maternal medication use, which could occur if a diagnosis arises due to medication (for example, as its side effect) and the same medication is also associated with autism in the offspring. Second, we did not have information about several sociodemographic factors, which limited our ability to control for the potential effects of these characteristics (for example, ethnicity, immigration status). Next, our ascertainment of chronic maternal diagnoses was limited to 48 months before childbirth, which, in a birth cohort starting in 1998, is the longest period for which maternal diagnostic codes in Denmark were available fully in the ICD-10 system. While extending this exposure period in truncated datasets had minimal impact on our findings, and the effect sizes associated with a number of conditions were closely aligned in our and other studies—including, for example, depression[47–49], obesity[50,51] and gestational diabetes[52,53]—the associations between four maternal diagnoses and autism were observable only for extended exposure periods. As the appropriate length of the exposure period for the ascertainment of chronic conditions is not clear, we cannot discount the possibility of a type II error and missing other associations due to our use of 48 months preceding birth as the exposure period. Next, our measure of healthcare utilization (total number of days of healthcare encounters) is likely a complex composite of health needs, health awareness, healthcare access, health anxiety, socioeconomic status and urbanicity, among others. Given the profound effects of this measure on the effect sizes in this and other[1,54] autism studies, further research is warranted to dissect these factors and their contribution to the associations reported in our study. While we observed that most of the associations that were no longer significant after including this covariate involved mild and nonspecific diagnoses, symptoms, and 'superficial' injuries, such adjustment could represent an overadjustment (although, in our analytical framework, the measure of healthcare utilization captured factors that could act

as confounders, these are analytically indistinguishable from mediators). Next, while there is currently no evidence for transgenerational epigenetic inheritance in humans, such potential effects cannot be discounted, which could affect the degree to which paternal exposure can be used as a negative control. However, even in contexts in which paternal exposure constitutes an imperfect negative control, its utility has been demonstrated, at least when the negative control exposure exhibits a null association[55]. Next, for chronic conditions, the ascertainment of disease onset and remission using registry data is complicated; for those diagnoses, sibling concordance status was less reliable, potentially increasing random measurement errors. While we highlight this problem and urge caution in interpreting the results from the sibling design for chronic conditions, we present them for comparison to the existing literature (for example, refs. 56–58) and note that the results were consistent with those derived from the paternal analysis. Relatedly, the effect estimates derived from the sibling analyses have an inherently different interpretation from those in nonsibling models, warranting caution when comparing them directly; specifically, the parameters obtained from the stratified Cox regression model describe the effect on the full sibling group, not the individual-level causal effect[59]. Next, for chronic conditions, we could verify through sensitivity analyses that our results were not driven by administrative errors or misdiagnoses that could arise from ascertaining exposure through one record of maternal diagnosis. However, as nonchronic conditions could plausibly be diagnosed at a single encounter, we did not perform such sensitivity analyses for those; therefore, the possibility of misclassification of nonchronic conditions remains. While observing the same results in the main and sensitivity analyses of chronic conditions and the excellent validity of diagnostic data in the Danish registers render a major impact of such misclassification unlikely, it remains a potential limitation. Next, due to multiple testing, the power analyses performed for the stratified analyses (by sex and ID status) might have underestimated the minimal prevalence of maternal diagnosis required; therefore, we urge caution in interpreting the results from those analyses even when supported by our power calculations. Relatedly, the prevalence of ID in individuals with autism in our sample (15.2%) is lower than that reported elsewhere[17], with a potential impact on the ID-stratified analyses. Lastly, the lack of primary care data outside of the pregnancy period limited the ascertainment of milder exposures not necessitating a specialist encounter.

In conclusion, the current study harnessed multiple analytical approaches to provide insights into the associations between maternal diagnoses during pregnancy and the risk of autism in the offspring. We identified new associations between maternal diagnoses and autism, replicated multiple previously established ones and found evidence supporting the role of familial confounding in many of these associations. Our findings draw attention to the importance of maternal health around pregnancy and reinforce the notion that many of the observational associations between perinatal factors and offspring neurodevelopment are likely noncausal in nature. There is a need to validate these conclusions in external datasets and investigate the role of different familial confounders—including both genetic and nongenetic factors—in driving these associations.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-03479-5.

References

1. Kodesh, A. et al. Maternal health around pregnancy and autism risk: a population-based study. *Psychol. Med.* **52**, 4076–4084 (2022).
2. Viktorin, A., Uher, R., Reichenberg, A., Levine, S. Z. & Sandin, S. Autism risk following antidepressant medication during pregnancy. *Psychol. Med.* **47**, 2787–2796 (2017).
3. Xiang, A. H. et al. Association of maternal diabetes with autism in offspring. *JAMA* **313**, 1425–1434 (2015).
4. Croen, L. A., Grether, J. K., Yoshida, C. K., Odouli, R. & Van de Water, J. Maternal autoimmune diseases, asthma and allergies, and childhood autism spectrum disorders: a case–control study. *Arch. Pediatr. Adolesc. Med.* **159**, 151–157 (2005).
5. Atladóttir, H. O. et al. Maternal infection requiring hospitalization during pregnancy and autism spectrum disorders. *J. Autism Dev. Disord.* **40**, 1423–1430 (2010).
6. Khachadourian, V. et al. Maternal psychiatric diagnosis and comorbidities around pregnancy. *Eur. Psychiatry* **66**, e15 (2023).
7. Zaks, N. et al. Association between mental health and reproductive system disorders in women: a systematic review and meta-analysis. *JAMA Netw. Open.* **6**, e238685 (2023).
8. Chucri, T. M. et al. A review of immune transfer by the placenta. *J. Reprod. Immunol.* **87**, 14–20 (2010).
9. Engelhardt, B. Development of the blood–brain barrier. *Cell Tissue Res.* **314**, 119–129 (2003).
10. Grove, J. et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
11. Ejlskov, L. et al. Prediction of autism risk from family medical history data using machine learning: a national cohort study from Denmark. *Biol. Psychiatry Glob. Open Sci.* **1**, 156–164 (2021).
12. Svensson, A. C. et al. Maternal effects for preterm birth: a genetic epidemiologic study of 630,000 families. *Am. J. Epidemiol.* **170**, 1365–1372 (2009).
13. Susser, E., Eide, M. G. & Begg, M. Invited commentary: the use of sibship studies to detect familial confounding. *Am. J. Epidemiol.* **172**, 537–539 (2010).
14. Lipsitch, M., Tchetgen Tchetgen, E. & Cohen, T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* **21**, 383–388 (2010).
15. Shattuck, P. T. et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. *J. Am. Acad. Child Adolesc. Psychiatry* **48**, 474–483 (2009).
16. Parner, E. T., Schendel, D. E. & Thorsen, P. Autism prevalence trends over time in Denmark: changes in prevalence and age at diagnosis. *Arch. Pediatr. Adolesc. Med.* **162**, 1150–1156 (2008).
17. Maenner, M. J. et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR Surveill. Summ.* **72**, 1–14 (2023).
18. Leonard, H. et al. Unpacking the complex nature of the autism epidemic. *Res. Autism Spectr. Disord.* **4**, 548–554 (2010).
19. Mandell, D. S., Novak, M. M. & Zubritsky, C. D. Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics* **116**, 1480–1486 (2005).
20. Thomas, T. R. et al. Clinical autism subscales have common genetic liabilities that are heritable, pleiotropic, and generalizable to the general population. *Transl. Psychiatry* **12**, 247 (2022).
21. Lee, P. H., Feng, Y.-C. A. & Smoller, J. W. Pleiotropy and cross-disorder genetics among psychiatric disorders. *Biol. Psychiatry* **89**, 20–31 (2021).
22. Nudel, R. et al. Maternal pregnancy-related infections and autism spectrum disorder—the genetic perspective. *Transl. Psychiatry* **12**, 334 (2022).
23. Karadag, N. et al. Identification of novel genomic risk loci shared between common epilepsies and psychiatric disorders. *Brain* **146**, 3392–3403 (2023).

24. Smeland, O. B. et al. The shared genetic risk architecture of neurological and psychiatric disorders: a genome-wide analysis. Preprint at *medRxiv* https://doi.org/10.1101/2023.07.21.23292993 (2023).
25. Plummer, J. T., Gordon, A. J. & Levitt, P. The genetic intersection of neurodevelopmental disorders and shared medical comorbidities—relations that translate from bench to bedside. *Front. Psychiatry* **7**, 142 (2016).
26. Gong, T. et al. Understanding the relationship between asthma and autism spectrum disorder: a population-based family and twin study. *Psychol. Med.* **53**, 3096–3104 (2023).
27. Atladóttir, H. O. et al. Association of family history of autoimmune diseases and autism spectrum disorders. *Pediatrics* **124**, 687–694 (2009).
28. Koeppen-Schomerus, G., Spinath, F. M. & Plomin, R. Twins and non-twin siblings: different estimates of shared environmental influence in early childhood. *Twin Res.* **6**, 97–105 (2003).
29. Kendler, K. S., Ohlsson, H., Lichtenstein, P., Sundquist, J. & Sundquist, K. The nature of the shared environment. *Behav. Genet.* **49**, 1–10 (2019).
30. Frisell, T. Invited commentary: sibling-comparison designs, are they worth the effort? *Am. J. Epidemiol.* **190**, 738–741 (2021).
31. Sunde, H. F. et al. Genetic similarity between relatives provides evidence on the presence and history of assortative mating. *Nat. Commun.* **15**, 2641 (2024).
32. Nordsletten, A. E. et al. Patterns of nonrandom mating within and across 11 major psychiatric disorders. *JAMA Psychiatry* **73**, 354–361 (2016).
33. Greenland, S. The effect of misclassification in the presence of covariates. *Am. J. Epidemiol.* **112**, 564–569 (1980).
34. Ogburn, E. L. & Vanderweele, T. J. Bias attenuation results for nondifferentially mismeasured ordinal and coarsened confounders. *Biometrika* **100**, 241–248 (2013).
35. Brew, B. K. & Gong, T. Modelling paternal exposure as a negative control. *Int. J. Epidemiol.* **49**, 1053–1054 (2020).
36. Taylor, K. et al. Effect of maternal prepregnancy/early-pregnancy body mass index and pregnancy smoking and alcohol on congenital heart diseases: a parental negative control study. *J. Am. Heart Assoc.* **10**, e020051 (2021).
37. Keyes, K. M., Davey Smith, G. & Susser, E. Commentary: smoking in pregnancy and offspring health: early insights into family-based and 'negative control' studies? *Int. J. Epidemiol.* **43**, 1381–1388 (2014).
38. Siklenka, K. et al. Disruption of histone methylation in developing sperm impairs offspring health transgenerationally. *Science* **350**, aab2006 (2015).
39. Rodgers, A. B., Morgan, C. P., Bronson, S. L., Revello, S. & Bale, T. L. Paternal stress exposure alters sperm microRNA content and reprograms offspring HPA stress axis regulation. *J. Neurosci.* **33**, 9003–9012 (2013).
40. Chen, X. et al. Whole genome bisulfite sequencing of human spermatozoa reveals differentially methylated patterns from type 2 diabetic patients. *J. Diabetes Investig.* **11**, 856–864 (2020).
41. Soubry, A. et al. Obesity-related DNA methylation at imprinted genes in human sperm: results from the TIEGER study. *Clin. Epigenetics* **8**, 51 (2016).
42. Sasaki, H. & Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* **9**, 129–140 (2008).
43. Wolf, J. B. & Wade, M. J. What are maternal effects (and what are they not)? *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1107–1115 (2009).
44. Huang, Q. Q. et al. Examining the role of common variants in rare neurodevelopment disorders. *Nature* **636**, 404–411 (2024).
45. Gardener, H., Spiegelman, D. & Buka, S. L. Prenatal risk factors for autism: comprehensive meta-analysis. *Br. J. Psychiatry* **195**, 7–14 (2009).
46. Larsson, H. J. et al. Risk factors for autism: perinatal factors, parental psychiatric history, and socioeconomic status. *Am. J. Epidemiol.* **161**, 916–925 (2005).
47. Yin, W. et al. Association between parental psychiatric disorders and risk of offspring autism spectrum disorder: a Swedish and Finnish population-based cohort study. *Lancet Reg. Health Eur.* **40**, 100902 (2024).
48. Hagberg, K. W., Robijn, A. L. & Jick, S. Maternal depression and antidepressant use during pregnancy and the risk of autism spectrum disorder in offspring. *Clin. Epidemiol.* **10**, 1599–1612 (2018).
49. Rai, D. et al. Parental depression, maternal antidepressant use during pregnancy, and risk of autism spectrum disorders: population based case–control study. *BMJ* **346**, f2059 (2013).
50. Li, Y.-M. et al. Association between maternal obesity and autism spectrum disorder in offspring: a meta-analysis. *J. Autism Dev. Disord.* **46**, 95–102 (2016).
51. Sanchez, C. E. et al. Maternal pre-pregnancy obesity and child neurodevelopmental outcomes: a meta-analysis. *Obes. Rev.* **19**, 464–484 (2018).
52. Guo, D., Ju, R., Dai, J. & Zhang, H. Autism spectrum disorders in offspring exposed to maternal gestational diabetes: a meta-analysis and systematic review. *Rev. J. Autism Dev. Disord.* https://doi.org/10.1007/s40489-023-00380-8 (2023).
53. Chen, S., Zhao, S., Dalman, C., Karlsson, H. & Gardner, R. Association of maternal diabetes with neurodevelopmental disorders: autism spectrum disorders, attention-deficit/hyperactivity disorder and intellectual disability. *Int. J. Epidemiol.* **50**, 459–474 (2021).
54. Janecka, M. et al. Association of autism spectrum disorder with prenatal exposure to medication affecting neurotransmitter systems. *JAMA Psychiatry* **75**, 1217–1224 (2018).
55. Weisskopf, M. G., Tchetgen Tchetgen, E. J. & Raz, R. Commentary: on the use of imperfect negative control exposures in epidemiologic studies. *Epidemiology* **27**, 365–367 (2016).
56. Gjerde, L. C. et al. Maternal perinatal and concurrent depressive symptoms and child behavior problems: a sibling comparison study. *J. Child Psychol. Psychiatry* **58**, 779–786 (2017).
57. Gjerde, L. C. et al. Maternal perinatal and concurrent anxiety and mental health problems in early childhood: a sibling-comparison study. *Child Dev.* **91**, 456–470 (2020).
58. Bekkhus, M., Lee, Y., Brandlistuen, R. E., Samuelsen, S. O. & Magnus, P. Maternal anxiety and infants birthweight and length of gestation. A sibling design. *BMC Psychiatry* **21**, 609 (2021).
59. Petersen, A. H. & Lange, T. What is the causal interpretation of sibling comparison designs? *Epidemiology* **31**, 75–81 (2020).

[1]Department of Child and Adolescent Psychiatry, NYU Grossman School of Medicine, New York, NY, USA. [2]Department of Biomedicine, Aarhus University, Aarhus, Denmark. [3]Department of Genetic and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [4]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [5]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [6]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [7]Division of Research, Kaiser Permanente Northern California, Pleasanton, CA, USA. [8]A.J. Drexel Autism Institute, Drexel University, Philadelphia, PA, USA. [9]The Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH), Aarhus, Denmark. [10]National Centre for Register-Based Research, Aarhus BSS, Aarhus University, Aarhus, Denmark. [11]Department of Public Health, Aarhus University, Aarhus, Denmark. [12]Department of Population Health, NYU Grossman School of Medicine, New York, NY, USA. ✉e-mail: magdalena.janecka@nyulangone.org

Methods

Study design and population

The source population for this cohort study comprised all children born in Denmark between January 1, 1998, and December 31, 2015, and their parents identified in the Danish Medical Birth Registry, when possible, or otherwise in the Danish Central Population Register[60–62]. All individuals were followed up through December 2016. Access to the data was approved by the Danish Scientific Ethical Committee system, as well as all relevant register authorities, including the Danish Data Protection Agency, Statistics Denmark and the Danish Health Data Authority.

Outcome

Individuals with autism were ascertained based on the diagnosis of ASD (ICD-10: F84.0, F84.1, F84.5, F84.8 or F84.9; in the absence of F84.2 and F84.3), obtained through linkage with the Danish Psychiatric Central Research Register[61] and Danish National Patient Register[60]. In Denmark, general practitioners or school psychologists refer children with suspected ASD to a child and adolescent psychiatric department, where they undergo a multidisciplinary evaluation. All ASD diagnoses reported to the national registries are determined by child and adolescent psychiatrists and have been reported according to the ICD-10 system since 1994; diagnoses from both inpatient and outpatient contacts, as well as the emergency department, have been reported to the Danish Psychiatric Central Research Register since 1995. The validity of the autism disorder diagnosis (ICD-10 F84.0, childhood autism) in the Danish registers has been confirmed[63]. All children in the sample were followed from birth until the first diagnosis of ASD, emigration, death or the end of follow-up (December 31, 2016), whichever occurred first[63].

Exposure

All reported maternal diagnoses occurring during the exposure detection window (see below) and coded using the ICD-10 served as the exposures. Maternal diagnoses were ascertained through linkage with the Danish Psychiatric Central Research Register (diagnoses of mental disorders) and Danish National Patient Register (diagnoses of nonmental conditions), which have relied on the ICD-10 classification since 1994. The Danish national health registers comprise diagnoses reported by specialists following outpatient or inpatient contacts in public clinics and hospitals; diagnoses made by a general practitioner are not reported to the registries. During pregnancy, midwives or medical doctors report information from prenatal visits. Diagnoses reported to the registry include codes relevant to the current appointment, including both new and preexisting diagnoses but not past diagnoses not pertinent to the current encounter. The validity of the registry-based diagnoses for research has been confirmed for a host of conditions[64,65], and all diagnoses used to assess the Charlson comorbidity index were found to have an excellent (>98%) positive predictive value and low levels of ICD coding errors compared to the discharge notes[66].

The hierarchical organization of diagnostic codes in ICD-10 has four levels, presenting information from the least (level 1) to the most (level 4) specific. The current analyses were based on level 3 diagnostic codes (hereafter referred to as 'diagnoses' for brevity; for example, F33 for major depressive disorder). We used the Chronic Condition Indicator developed by the Agency for Healthcare Research and Quality[67] to assign the diagnoses into the chronic and nonchronic categories. Any level 3 diagnostic category that included both chronic and nonchronic subdiagnoses at level 4 was assessed and classified on a case-by-case basis by clinical experts in the team. All exposure variables were binary indicators of the presence or absence of a given diagnosis in the exposure window.

Exposure window and ascertainment. For nonchronic diagnoses, the exposure window comprised the 12 months preceding childbirth. Assuming a full-term pregnancy, this period captured the pregnancy duration and the 3 months preceding conception. A detection period before pregnancy was set to capture diagnoses occurring shortly before conception, whose residual effects may still influence fetal development. Diagnoses occurring in fewer than ten mothers of children with and/or without autism were excluded from the analyses to minimize the risk of sparse data bias[68].

For chronic diagnoses, the exposure window was defined as the 48 months preceding childbirth, allowing consistent ascertainment of ICD-10 diagnoses for the 1998–2015 birth cohort (the transition from ICD-8 to ICD-10 in Denmark occurred in 1993) while maximizing the exposure period. This definition of exposure window was dictated by the assumption that chronic diagnoses are permanent following their onset, the timing of which cannot be precisely determined using registry-based data. Therefore, a wider detection window increases the sensitivity in capturing chronic maternal diagnoses that may not have been entered into the registry around pregnancy (for example, well-managed conditions that do not require frequent medical attention) but may still affect the fetus. Diagnoses occurring in <20 mothers of children with and/or without autism were excluded from the analyses (note that, due to the broader exposure window, the threshold is higher compared to nonchronic diagnoses).

Covariates

The study covariates included the child's sex and year of birth, maternal age at childbirth, the total number of days of healthcare encounters during the 12 months preceding childbirth, and maternal education and income a year before delivery. Birth year was included to account for the varying incidence of autism and maternal diagnoses over time[69]; maternal age was included to account for its association with offspring autism[70,71] and various maternal medical conditions[72]; the total number of days of recorded healthcare encounters during the 12 months preceding childbirth (outpatient or inpatient contacts in public clinics and hospitals) was a proxy composite measure for health-seeking behaviors, healthcare utilization, healthcare access and maternal morbidity—all of which could influence the likelihood of both maternal (exposure) and child (outcome) diagnoses. Maternal education and income were included in the models to account for the potential effects of socioeconomic status on the likelihood of both maternal diagnoses and offspring autism diagnosis[73,74]. Covariate information was obtained from the Danish Central Population Register, Danish Psychiatric Central Research Register[61] and Danish National Patient Register[60]. Offspring ID (ICD-10 F70–F79) diagnosis and maternal ASD (ICD-8: 299.00, 299.01, 299.03 or ICD-10: F84.0, F84.1, F84.5, F84.8 or F84.9; no ICD-10 F84.2 or F84.3) diagnosis were used for stratified or sensitivity analyses.

Statistical analysis

The analyses comprised two phases. The first phase systematically tested the associations between each maternal diagnosis and autism, including covariate adjustments and several stratified and sensitivity analyses. In the second phase, we evaluated the impact of unmeasured familial factors on these associations.

Phase 1: associations between maternal diagnoses and offspring autism. We evaluated the associations between each maternal diagnosis and autism in the offspring using Cox proportional hazard models implemented in the R software 'survival' package[75], adjusting for the study covariates. Clustering sandwich estimators were used to account for within-family correlations due to the presence of siblings in the dataset.

First, each maternal diagnosis was tested separately for its association with autism, adjusting for covariates. To account for a multitude of tests being performed (equal to the number of distinct maternal diagnoses), we applied false discovery rate (FDR) correction[76] for multiple testing on empirical P values. The permissible FDR (q value) was set at 0.05. Second, we adjusted for possible comorbidity between

different maternal diagnoses as follows: all chronic and nonchronic diagnoses with a statistically significant association with autism in the fully adjusted single-diagnosis models (FDR $q < 0.05$) were concurrently included in a multidiagnosis model, in addition to the full set of covariates.

We ran additional models, stratifying by the child's sex and a co-occurring diagnosis of ID. As this resulted in reduced sample sizes for analytical subgroups (for example, female individuals with autism, individuals with co-occurring ID), we conducted appropriate power analyses to establish the minimal prevalence of maternal diagnosis to conclude the presence or absence of stratum-specific effects.

Due to the minimal missingness of administrative data in the Danish registers, cases with missing covariate information were removed from the analyses.

Sensitivity analyses. We conducted a set of sensitivity analyses to evaluate the impact of our analytical choices. To interrogate the robustness of the results to different definitions of healthcare utilization, we reanalyzed the data using different definitions of this covariate, including (1) the number of distinct diagnoses in the 12 months before childbirth and (2) the number of days of encounters with the healthcare system in the 48 months before childbirth, as well as (3) by omitting the covariate from the analyses. To inspect the sensitivity of the results to the definition of the exposure and follow-up periods, we (4) reanalyzed the data in a subsample of children born by the end of 2009, ensuring >8 years of follow-up for all children; next, we restricted the sample to children with maternal health information in the ICD-10 system for at least 72 months before childbirth (as ICD-10 was introduced in Denmark in 1993, this included children born after 1999) and retested the associations for both (5) shorter (12 months) and (6) longer (72 months) exposure periods to inspect the impact of varying exposure periods for the ascertainment of chronic conditions; (7) to examine the impact of exposure definition, we reanalyzed the association between maternal chronic disorders and autism, ascertaining only those conditions diagnosed on at least two distinct occasions within the exposure period (note that mothers with a single diagnosis and their children were excluded from this analysis). In the multidiagnosis model, to evaluate a broader measure of familial socioeconomic status, we (8) adjusted for paternal education and income. Furthermore, to account for potential parity and birth-order effects, we (9) adjusted for parity and (10) restricted the sample to first-born children and (11) to singletons only. Finally, to ensure that the observed associations did not arise because of the co-occurrence of maternal autism and other maternal conditions, we (12) limited the sample to children from mothers without an ASD diagnosis.

Phase 2: impact of unmeasured familial confounding on associations between maternal diagnoses and autism. We used two orthogonal approaches to elucidate familial confounding. First, we implemented a sibling design to contrast autism likelihood in siblings discordant for maternal diagnosis status in the exposure period. Second, we used the negative control of paternal exposure design to compare the associations between offspring autism and both maternal and paternal diagnoses.

Discordant sibling analysis. To assess the effect of time-invariant shared familial confounders on the observed associations, we repeated the fully adjusted analysis treating maternal unique identification numbers as a stratification variable in the Cox regression models and restricting the sample to children with at least one maternal sibling in the cohort. To allow for a direct comparison of the sibling and main (phase 1) results despite the sample restriction, we repeated the phase 1 analyses in the subsample of children with at least one sibling. Additionally, in a sensitivity analysis, we conducted sibling analyses restricting the sample to male children only to evaluate the potential impact of different recurrence risks[77] based on the sex of the sibling with and without autism (due to the lower prevalence of autism in female individuals and the relatively few autism-discordant female-only sibling pairs, analogous analyses in female individuals are underpowered). As the sibling approach relies on accurate ascertainment of the timing of the exposure to determine exposed and unexposed sibling status, it may not provide a reliable inference for chronic diagnoses due to the difficulty in ascertaining the disease onset—a limitation in all sibling analyses examining chronic maternal conditions.

Negative control of paternal exposure. We repeated the fully adjusted analyses (phase 1), additionally including the same ICD-10 paternal diagnosis received during the same exposure period in the regression models. This enabled a direct comparison of the effects of maternal and paternal diagnoses on the likelihood of autism, providing potential insights regarding familial confounding. Concurrent adjustment for maternal diagnoses in paternal analyses minimizes the potential inflation of paternal coefficients by factors related to maternal condition; such inflation could otherwise arise due to a correlation between parents, induced by, for example, assortative mating[78]. We analyzed the effects of both chronic and nonchronic diagnoses with statistically significant effects in the phase 1 analysis after controlling for multiple testing but excluded those not applicable to cis men.

All analyses were implemented in R software (version 4.1.3)[79].

Inclusion and ethics

The research included local (Danish) researchers at all stages of the research process and is locally relevant. The roles and responsibilities in this project were agreed on ahead of the research. The study was approved by a local ethics review committee.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The Danish Scientific Ethical Committee system, as well as all relevant register authorities, including the Danish Data Protection Agency, Statistics Denmark and the Danish Health Data Authority, approved access to the data under strict conditions regarding access and data export. Under these conditions, there are no provisions for exporting individual-level data, all or in part, to another institution in or outside of Denmark. The minimum datasets, including all summary statistics of the measures of associations used to draw the study conclusions, are presented in the supplementary material. The corresponding author will respond to any potential additional queries within 2 weeks of receiving the query.

Code availability

All analytical code is available via GitHub at https://github.com/v-k-lab/mat.dx.asd.dk.

References

60. Lynge, E., Sandegaard, J. L. & Rebolj, M. The Danish National Patient Register. *Scand. J. Public Health* **39**, 30–33 (2011).
61. Mors, O., Perto, G. P. & Mortensen, P. B. The Danish Psychiatric Central Research Register. *Scand. J. Public Health* **39**, 54–57 (2011).
62. Pedersen, C. B. The Danish Civil Registration System. *Scand. J. Public Health* **39**, 22–25 (2011).
63. Lauritsen, M. B. et al. Validity of childhood autism in the Danish Psychiatric Central Register: findings from a cohort sample born 1990–1999. *J. Autism Dev. Disord.* **40**, 139–148 (2010).
64. Sundbøll, J. et al. Positive predictive value of cardiovascular diagnoses in the Danish National Patient Registry: a validation study. *BMJ Open* **6**, e012832 (2016).

65. Albaek Jacobsen, H., Jess, T. & Larsen, L. Validity of inflammatory bowel disease diagnoses in the Danish National Patient Registry: a population-based study from the north Denmark region. *Clin. Epidemiol.* **14**, 1099–1109 (2022).
66. Thygesen, S. K., Christiansen, C. F., Christensen, S., Lash, T. L. & Sørensen, H. T. The predictive value of ICD-10 diagnostic coding used to assess Charlson comorbidity index conditions in the population-based Danish National Registry of Patients. *BMC Med. Res. Methodol.* **11**, 83 (2011).
67. *Chronic Condition Indicator (CCI) for ICD-10-CM (Beta Version)* (Agency for Healthcare Research and Quality, 2018).
68. Greenland, S., Mansournia, M. A. & Altman, D. G. Sparse data bias: a problem hiding in plain sight. *BMJ* **352**, i1981 (2016).
69. Schendel, D. E. & Thorsteinsson, E. Cumulative incidence of autism into adulthood for birth cohorts in Denmark, 1980–2012. *JAMA* **320**, 1811–1813 (2018).
70. Sandin, S. et al. Advancing maternal age is associated with increasing risk for autism: a review and meta-analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **51**, 477–486 (2012).
71. Janecka, M. et al. Parental age and differential estimates of risk for neuropsychiatric disorders: findings from the Danish Birth Cohort. *J. Am. Acad. Child Adolesc. Psychiatry* **58**, 618–627 (2019).
72. Biro, M. A., Davey, M.-A., Carolan, M. & Kealy, M. Advanced maternal age and obstetric morbidity for women giving birth in Victoria, Australia: a population-based study. *Aust. N. Z. J. Obstet. Gynaecol.* **52**, 229–234 (2012).
73. Daalderop, L. A., Lagendijk, J., Steegers, E. A. P., El Marroun, H. & Posthumus, A. G. Psychological distress during pregnancy and adverse maternal and perinatal health outcomes: the role of socioeconomic status. *Int. J. Gynecol. Obstet.* **163**, 920–930 (2023).
74. Delobel-Ayoub, M. et al. Socioeconomic disparities and prevalence of autism spectrum disorders and intellectual disability. *PLoS ONE* **10**, e0141964 (2015).
75. Therneau, T. M. & Li, H. Computing the Cox model for case cohort designs. *Lifetime Data Anal.* https://doi.org/10.1023/A:1009691327335 (1999).
76. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* https://doi.org/10.1111/j.2517-6161.1995.tb02031.x (1995).
77. Palmer, N. et al. Association of sex with recurrence of autism spectrum disorder among siblings. *JAMA Pediatr.* **171**, 1107–1112 (2017).
78. Madley-Dowd, P., Rai, D., Zammit, S. & Heron, J. Simulations and directed acyclic graphs explained why assortative mating biases the prenatal negative control design. *J. Clin. Epidemiol.* **118**, 9–17 (2020).
79. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2016).

Acknowledgements

This work was supported by grants from the National Institute of Mental Health (M.J., A.R., S.S. (grant no. MH124817 to M.J.) and V.K. (award no. T32-MH122394 to V.K.)) and the Lundbeck Foundation (iPSYCH, grant nos. R102-A9118 and R155-2014-1724). This work was also supported, in part, by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (A.R., M.J., S.S., J.D.B., D.S. and J.G. (grant no. HD098883 to A.R.)). We would like to acknowledge the generous support of the Seaver Foundation to J.D.B., A.R., S.S. and M.J. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies or the authors' employers.

Author contributions

M.J. conceived the study concept and design. V.K. and M.J. conceived the study design and statistical methods. V.K. and E.S.A. carried out all statistical analyses. V.K. prepared all the figures and tables. V.K., A.R. and M.J. obtained funding for the project. M.J. performed project supervision. V.K. and M.J. drafted the paper. All authors provided feedback on the study design and revised and approved the manuscript.

Competing interests

V.K. is currently employed by Takeda Pharmaceutical Company outside of the submitted work. The other authors declare no competing interests.

Additional information

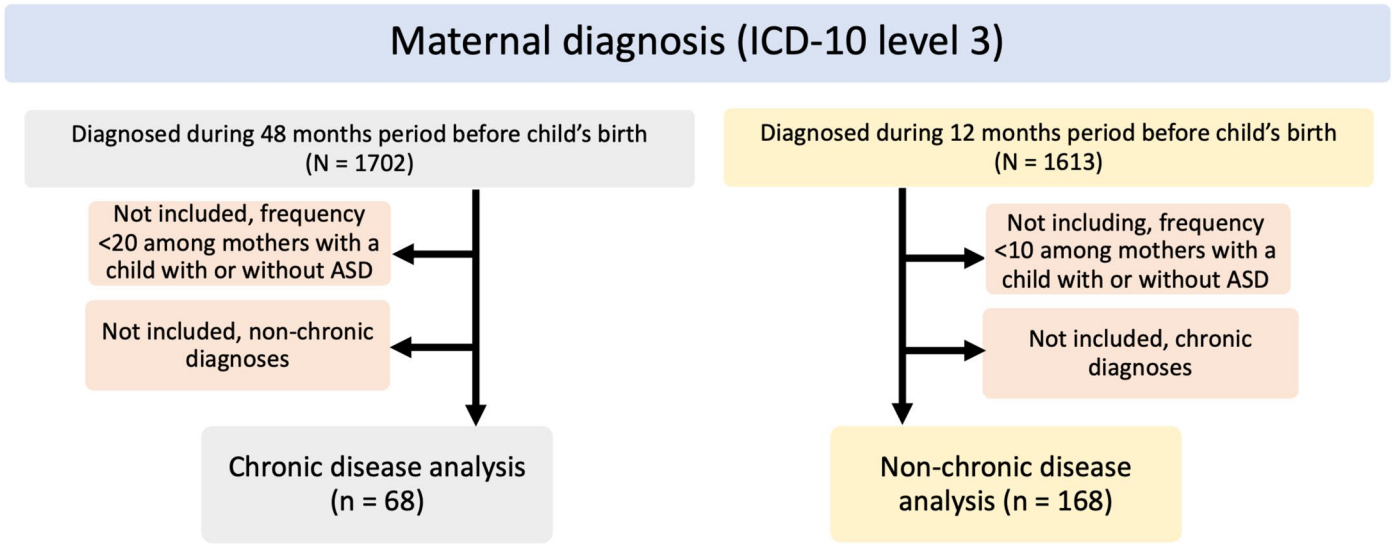
Extended data is available for this paper at https://doi.org/10.1038/s41591-024-03479-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-024-03479-5.

Correspondence and requests for materials should be addressed to Magdalena Janecka.

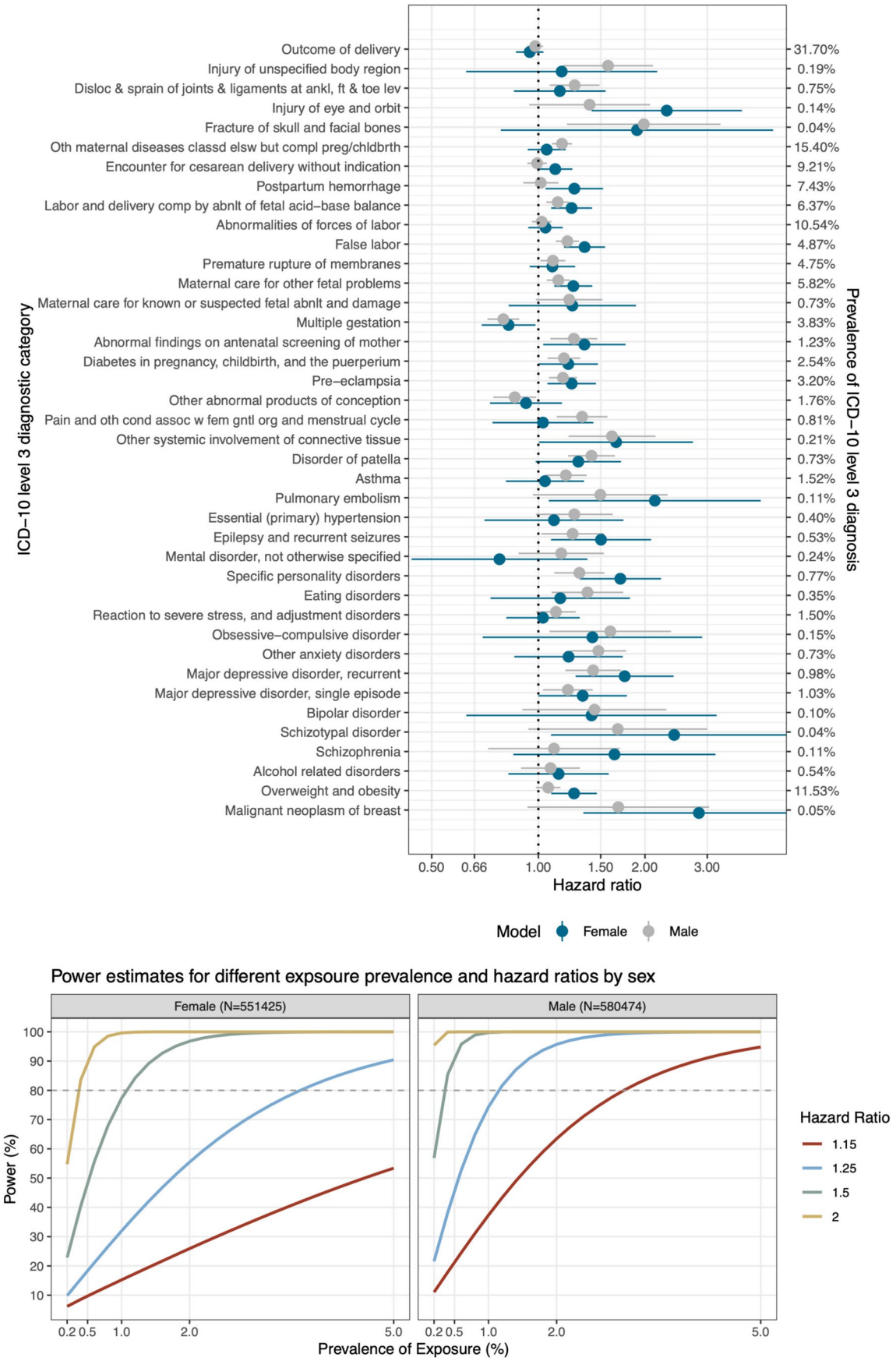
Peer review information *Nature Medicine* thanks Eric Fombonne and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Sonia Muliyil, in collaboration with the *Nature Medicine* team.

https://doi.org/10.1038/s41591-024-03479-5



Extended Data Fig. 1 | Flowchart of inclusion of maternal diagnosis around pregnancy in analyses. Flowchart of the number of maternal diagnoses included in the study.

https://doi.org/10.1038/s41591-024-03479-5

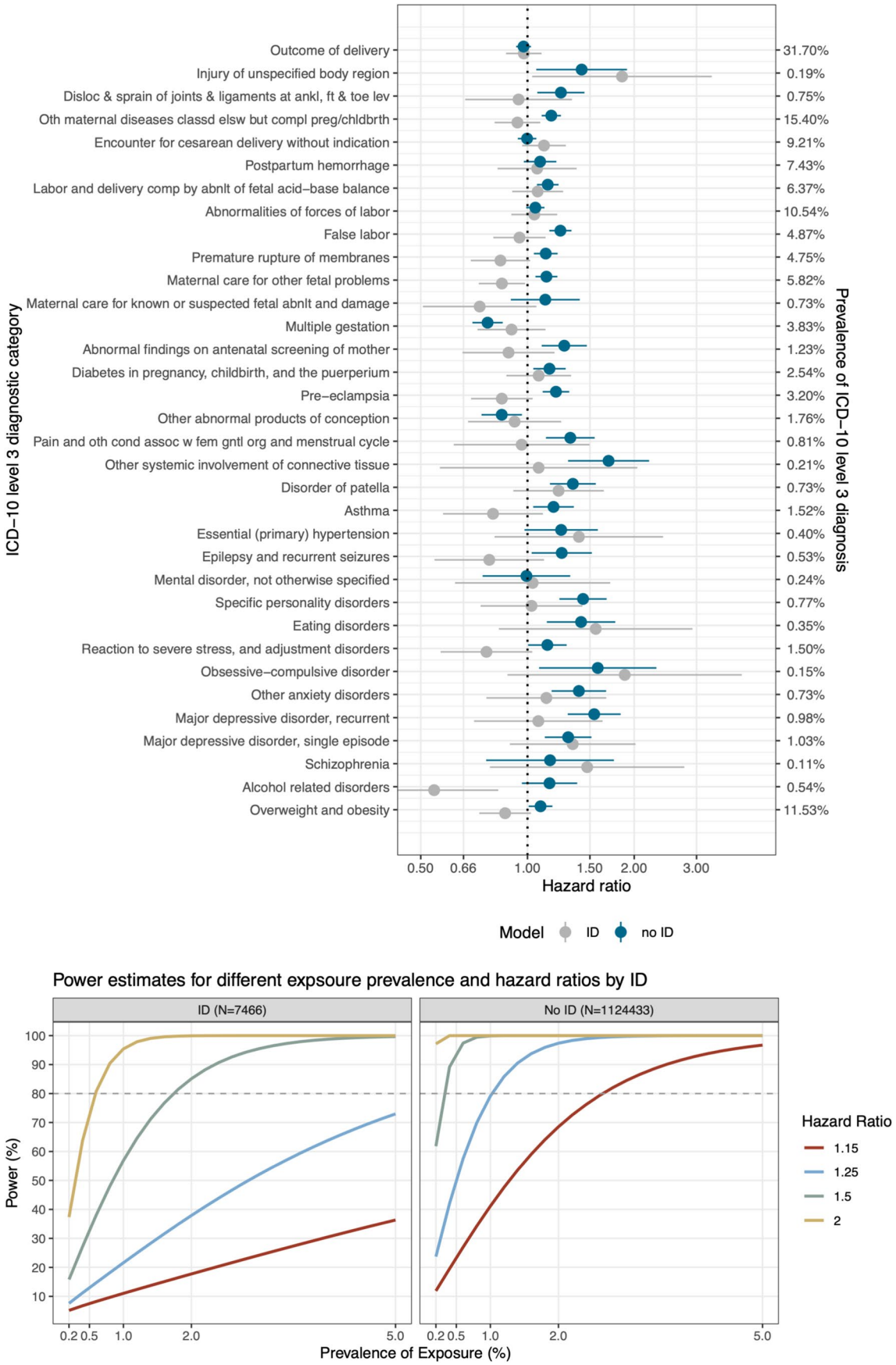


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41591-024-03479-5

Extended Data Fig. 2 | Associations between ICD-10 level 3 maternal diagnoses and offspring autism for non-chronic and chronic diagnoses stratified by sex of the child. Point estimates are hazard ratios (HRs) adjusted for maternal age at birth, child's year of birth, maternal income and education, and maternal health care encounter in the 12-month period preceding childbirth, and concurrently included all the diagnoses (non-chronic and chronic) in this figure. The error bars represent 95% confidence intervals, calculated using point estimates and robust standard errors from respective regression model. Prevalence of ICD-10 diagnoses were estimated using the entire sample, without stratification by child sex.

https://doi.org/10.1038/s41591-024-03479-5



Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41591-024-03479-5

Extended Data Fig. 3 | Associations between ICD-10 level 3 maternal diagnoses and offspring autism for non-chronic and chronic diagnoses stratified by intellectual disability (ID) in the child. Point estimates are hazard ratios adjusted for maternal age at birth, child's sex and year of birth, maternal income and education, and maternal health care encounter in the 12-month period preceding childbirth, and concurrently included all the diagnoses (non-chronic and chronic) in this figure. The error bars represent 95% confidence intervals, calculated using point estimates and robust standard errors from respective regression model. Prevalence of ICD-10 diagnoses were estimated using the entire sample, without stratification by child sex.

nature portfolio

Corresponding author(s): Magdalena Janecka

Last updated by author(s): 08/29/2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	All analyses were performed using R software (version 4.1.3), including following packages: survival

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The Danish Scientific Ethical Committee system, as well as all relevant register authorities, including the Danish Data Protection Agency, Statistics Denmark, and the Danish Health Data Authority, approved access to these data under strict conditions regarding access and data export. Under these conditions there are no provisions for exporting individual level data, all or in part, to another institution, in or outside of Denmark. All summary statistics of the measures of associations used to draw the study conclusions are presented in the supplemental material. The corresponding author will respond to any potential additional queries.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Throughout the manuscript we refer to data on biological sex available in Danish national registers. No corresponding gender identify information was available.
Reporting on race, ethnicity, or other socially relevant groupings	The data comes from the Danish national resource. No groupings based on participants ancestry, race or ethnicity has been performed.
Population characteristics	The dataset is drawn from a birth cohort including children born in Denmark 1998-2015, followed up until 2017. In all analyses, we adjusted for year of child's birth, maternal age at child's birth, and duration of follow-up. Additionally, we adjusted for socioeconomic maternal characteristics (education and income). In additional sensitivity we further adjusted for paternal age, paternal education and income. The sample was 48.7% female and 51.3% male, with median length of follow-up of 9.7 years (IQR: 5.3-14.3).
Recruitment	We included all children live born in Denmark 1998-2015 and their parents, with no additional ascertainment criteria. In the final analytical sample we excluded children lacking data on covariates.
Ethics oversight	Access to the Danish registry data has been approved by The Danish Scientific Ethical Committee system. The study was additionally reviewed by the institutional review board at the NYU Grossman School of Medicine and considered exempt.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We selected all children born in Denmark within the cohort years. Prior to the analyses we have performed power analysis showing that we'd be extremely well powered even for very rare outcomes.
Data exclusions	The only exclusion criterion was lack of appropriate covariate data. All exclusion criteria were pre-determined.
Replication	The findings have yet to be replicated in additional national birth cohorts. However, the association findings between maternal diagnoses of obstetric, metabolic and psychiatric conditions, and the risk of autism spectrum disorder in children, replicate several already existing studies. As the manuscript presents analyses of retrospective data, no experimental replication was performed as part of the study design.
Randomization	It is an observational study with no experimenter-introduced randomization.
Blinding	No blinding was performed as the analyses were carried out on retrospective data, where both the exposure and outcome status were known at the point of analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA