

ARTICLE

OPEN



Comorbidities in autism spectrum disorder and their etiologies

Vahe Khachadourian ^{1,2,3}, Behrang Mahjani ^{1,2}, Sven Sandin^{1,2}, Alexander Kolevzon^{1,2}, Joseph D. Buxbaum ^{1,2,3,4}, Abraham Reichenberg^{1,2,3,5} and Magdalena Janecka ^{1,2,3,4,6}✉

© The Author(s) 2023

Individuals with autism spectrum disorder (ASD), in addition to the core features of the disease, experience a higher burden of co-occurring medical conditions. This study sought to describe the frequency and distribution of comorbid conditions in individuals with ASD, and systematically evaluate the possibility that pre- and postnatal exposures (e.g., preterm birth, hypoxia at birth, traumatic brain injury, and fetal alcohol syndrome) associated with ASD may also be linked with distinct comorbidities. We used the SPARK study database, launched by the Simons Foundation Autism Research Initiative (SFARI). Comorbidities considered in the study included neurological, cognitive, psychiatric, and physical conditions. The study sample consisted of 42,569 individuals with ASD and their 11,389 non-ASD siblings (full and half siblings). Majority (74%) of individuals with ASD had at least one comorbidity, and had a greater average number of comorbidities than their non-ASD siblings. Preterm birth and hypoxia at birth were the most common peri-natal exposures in the sample. In logistic regression models adjusted for covariates, these exposures were associated with several distinct comorbidities in ASD cases, including attention and behavior problems, psychiatric and neurological disorders, and growth conditions. A similar pattern of association was also observed in non-ASD siblings. Our findings underscore that individuals with ASD experience a greater burden of comorbidities, which could be partly attributable to the higher rates of perinatal exposures compared to their non-ASD siblings. Study findings, if replicated in other samples, can inform the etiology of comorbidity in ASD.

Translational Psychiatry (2023)13:71; <https://doi.org/10.1038/s41398-023-02374-w>

INTRODUCTION

Over the past 20 years, the reported prevalence of autism spectrum disorder (ASD) in the US has tripled, with current estimates indicating a prevalence of 1.85% among children [1]. The core symptoms of ASD include social communication issues, restricted interests, and repetitive behaviors. In addition to these core diagnostic features, affected individuals experience a higher burden of co-occurring conditions [2–9] impacting a wide range of body systems. For instance, people with an ASD diagnosis are more likely to be overweight or obese [2, 7, 9], or have a co-occurring accompanying psychiatric condition, such as anxiety and mood disorders [10], compared to those without ASD. Prevalence estimates of several other conditions, including epilepsy [5], sleep disorders [4], and gastrointestinal disease [6] are also much higher in ASD than those observed in the general population.

The comorbidities in ASD might have atypical manifestations and symptomatology, often making their recognition more difficult [11]. Despite these potential difficulties in the diagnosis, the reported comorbidity burden in ASD is still substantial - in a population-based study in Sweden, Lundström et al. [12] found that over 50% of individuals with ASD had four or more comorbid conditions.

The observations of higher rates of certain health conditions in ASD [2–9] suggest the possibility of a shared etiology between these different comorbidities and ASD. This possibility is

substantiated by the observations that some ASD-associated risk genes often have pleiotropic effects on a range of other disorders and phenotypes (e.g., DDX3X and gait disturbance [13]). Despite this evidence, in non-syndromic forms of ASD, potential etiological overlaps between ASD and its comorbidities have not been systematically explored. This is despite the fact that investigation of comorbid conditions [14] associated with ASD has the potential to offer insights into the etiology and risk architecture of the disorder.

The current study attempted therefore to address this knowledge gap by (1) describing the frequency and distribution of comorbidities in individuals with ASD, (2) examining if pre- and postnatal exposures previously associated with ASD may be linked to distinct comorbidities in cases, thereby highlighting the potentially 'pleiotropic' effects of these exposures, and (3) conducting a comparative analysis in non-ASD siblings to examine if the associations between these environmental exposures and comorbidities might occur independently of the ASD diagnosis.

METHODS

Study design and population

We used the SPARK study database, launched by the Simons Foundation Autism Research Initiative (SFARI). The SPARK study recruited families with one or more children with an ASD diagnosis from 21 clinical sites throughout the US [15]. The SPARK database contains data collected from

¹Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ²Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ³Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁴Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁵Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁶Department of Genetic and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ✉email: magdalena.janecka@mssm.edu

Received: 21 September 2022 Revised: 9 February 2023 Accepted: 17 February 2023

Published online: 25 February 2023

parents, including demographic characteristics, birth and developmental history, and medical diagnoses. All phenotypic data and biospecimens are collected remotely to allow participants to complete the study protocol at their convenience. The scientific community can access SPARK data free of charge (<https://www.sfari.org/resource/spark/>).

We used all available SPARK sample with phenotypic information, excluding the observations marked for possible unreliable ASD diagnosis or medical data. To minimize recall bias and increase the reliability of the data on early-life exposures, we restricted the sample to individuals born between 1999 and 2019 who were less than 18 years of age at the time of registration into the SPARK study. Finally, we excluded the individuals with maternal age outside the range of 13–55 years, or paternal age of less than 13 years at child's birth, due to suspected administrative errors in the records. A study flow diagram is presented in Fig. S1.

The study protocol was reviewed and approved by the institutional review board of the Icahn School of Medicine at Mount Sinai.

Exposure and outcome

Pregnancy and birth-related conditions served as the main exposure variables, including preterm birth (gestational age of <37 weeks), fetal alcohol syndrome, serious prenatal infection (e.g., German measles), hypoxia at birth, and bleeding into brain during delivery. Hypoxia at birth was a binary variable that captured those children with hypoxia at birth who were admitted to a neonatal intensive care unit (NICU). Postnatal exposures included lead poisoning, brain infections such as bacterial meningitis, and encephalitis, and traumatic brain injury requiring hospitalization.

Medical history data in SPARK is based on parental report of professional diagnosis and is divided into several domains. A positive reply for each domain is followed by more specific and detailed questions regarding the medical conditions within that domain. In this study, as outcomes we considered medical conditions ascertained in the domains of (a) birth or pregnancy complications, (b) attention or behavior disorders, (c) speech and language, intellectual disability/cognitive Impairment, learning disability (LD), or other developmental delay or developmental disability, (d) growth conditions, (e) neurological conditions, (f) vision or hearing conditions, (g) mood, depression, anxiety or obsessive-compulsive disorder (OCD), and (h) sleep, feeding/eating or toileting problems.

For brevity and ease of reading, we refer to these outcomes as comorbidities, even when discussing the results from the non-ASD siblings, where unlike in children with ASD, there is no index condition.

Covariates

Family- and parental-level covariates included the annual household income, and father's and mother's highest education levels, which served as a proxy for socio-economic status (SES), shown to be associated with ASD risk [16], ASD diagnosis [17–19] as well as health disparities [19, 20] which could impact the probability of receiving a comorbid diagnosis. Paternal and maternal ages at the time of delivery were adjusted for to account for the possibility of higher rates of certain pre- and postnatal exposures and medical outcomes among the children of very young / older parents (e.g., [21, 22]). Child's year of birth, age of the child at evaluation, and survey version, were additional covariates considered in this study, allowing us to control for the differences in the outcome/exposure rates that could be induced by the timing or the method of evaluation. Finally, we adjusted for child's sex as well as the parent-reported race and ethnicity of the child to control for potential differences in the prevalence of the pre- and post-natal exposures [20, 23] and ASD diagnosis by gender, race and ethnicity [17–19, 24].

Analysis

All analyses were performed using R programming language (version 4.0.0). The analysis of the non-ASD siblings included both full- and half-siblings. Characteristics of the analytical sample by ASD status were summarized using means and frequencies. Prevalence of the exposures and comorbidities were estimated by ASD status. These estimates among the non-ASD siblings were standardized to the year of evaluation, age, and gender distribution observed among children with ASD in order to ensure that possible differences between these groups are not driven by demographic or recruitment characteristics. The standardization was performed by weighting the non-ASD siblings to mimic the year of evaluation, age, and gender distributions observed among children with ASD [25]. We repeated the analysis estimating the prevalence of exposures

and comorbidities in sibling pairs (two siblings from each family) who were discordant for ASD diagnosis.

The binary logistic regression model was used to assess the associations between the exposures and outcomes of interest among children with ASD. Each outcome was evaluated in a separate model, including the exposures of interest and covariates. Considering potential clustering effects due to multiple deliveries among mothers in the study, we used clustered sandwich estimator, implemented in the clusterSEs package (v2.6.2). To avoid reporting unreliable associations due to the sparse data bias [26], we did not analyze the exposure-outcome pairs with a recorded outcome frequency of less than 10 among exposed / unexposed children. To account for multiple testing, we applied false discovery rate correction [27] to the empirical *p*-values.

Next, in each exposure-outcome model, we also adjusted for all other comorbidities grouped with the outcome (as per comorbidity groups identified by SPARK) in order to obtain estimates that are independent of the correlation patterns between these closely related comorbidities. Finally, we estimated associations between the exposure-outcome pairs among the non-ASD siblings and contrasted the estimates with those obtained in the sample of children with ASD.

Missing data for the covariates were imputed using mice R package, applying multivariate imputation by chained equation [28]. The sample was restricted to the individuals with medical history data; hence for the exposure and outcome variables, no missing data imputation was required. We performed 5 imputations using information from all outcome and exposure variables as well as covariates (annual household income, and father's and mother's highest education levels, paternal age and maternal age, child' gender and year of birth, age of the child at evaluation, survey version, race, and ethnicity).

RESULTS

There was a total of 42,564 individuals with ASD, and 11,390 of their non-ASD siblings (born 1999–2019) with complete medical data. Forty-seven individuals with maternal age beyond the range of 13–55 or paternal age of less than 13, and 1936 individuals who were 18 and older at the time of registration into SPARK, were excluded from the analysis. The final analytic sample included 51,971 participants (40,582 with ASD and 11,389 siblings without ASD) born to 34,929 mothers.

Seventy-nine percent of children with ASD, and 49% non-ASD siblings were male. Mean maternal age at delivery was 29.5 among children with ASD and 29.1 among children without ASD diagnosis. Mean paternal age was also similar among cases and controls, 31.8 and 31.5 years respectively. Table 1 and Table S1 provide further details about sociodemographic characteristics of the analytical sample by ASD status.

Prevalence of pre- and postnatal exposures

Individuals with ASD had a higher standardized prevalence of all of the pre-and postnatal exposures. For example, the prevalence of intraventricular hemorrhage among children with ASD was 0.9% compared to 0.3% among non-ASD sibling controls (Fig. 1). Similarly, prevalence of brain infection (0.3% vs <0.1%), fetal alcohol syndrome (1.2% vs 0.2%), infection in pregnancy (0.3% vs 0.1%), lead poisoning (0.3%, 0.1%), and traumatic brain injury (0.5%, 0.2%) were more than two-folds higher among children with ASD compared to non-ASD sibling controls. Relative to non-ASD siblings, children with ASD diagnosis also had a higher prevalence of hypoxia at birth (6.9% vs 4.6%, *p*-value < 0.05) and were more likely to be born preterm (13.2% vs 10.0%, *p*-value < 0.05). The same pattern was observed when we restricted the sample to pairs of siblings who were discordant for ASD diagnosis (Fig. S2).

Prevalence and co-occurrence of comorbidities

Children with ASD had a substantially higher standardized prevalence of all comorbidities analyzed in this study compared to their siblings without an ASD diagnosis (all *p*-values < 0.05). Attention deficit hyperactivity disorder (ADHD) was the most common comorbidity, affecting more 1 in every 3 children with

ASD (35.3%), much higher than 1 in 6 (16.8%) among non-ASD siblings. Learning disability (23.5%) and intellectual disability (21.7%) were the next most-common comorbid conditions among children with ASD. Figure 2 presents prevalence of the analyzed comorbid conditions in the SPARK sample by ASD status and Fig. S3 illustrates this in pairs of siblings who were discordant for ASD diagnosis.

The proportion of individuals with any comorbidity diagnosis was higher in individuals with ASD compared to their non-ASD siblings (Fig. 3), mirroring the patterns observed for the individual diagnoses. Stratifying the sample based on the presence of at least one comorbidity, the burden of comorbidities remained higher among individuals with ASD than their non-ASD siblings. The correlations between pairs of the 28 comorbidities that served as outcomes in our analyses were predominantly positive (Fig. S4). The correlations in the non-ASD siblings (Fig. S5) were mostly similar to the ones observed in the ASD sample. However, the strength of the correlations in the non-ASD siblings was often lower than those observed in individuals with ASD.

Comparing the prevalence of comorbidities between children with no pre-and postnatal exposure to those with at least one pre-and postnatal exposure we found a relatively higher comorbidity prevalence among both ASD children as well as their non-ASD siblings (Fig. S6).

Association of pre-and postnatal exposures with comorbid conditions among children with ASD

There was a total of 224 exposure-comorbidity pairs, reflecting all combinations between 8 exposures and 28 outcomes (Fig. 4a).

Table 1. Demographic characteristics of the analytical sample by ASD status ($N = 51,971$).

	ASD ($N = 40,582$)	Control ($N = 11,389$)
Child's year of birth		
1999–2004, n (%)	5579 (13.7%)	1638 (14.4%)
2005–2009, n (%)	11,823 (29.1%)	3481 (30.6%)
2010–2014, n (%)	15,883 (39.1%)	3985 (35.0%)
2015–2019, n (%)	7297 (18.0%)	2285 (20.1%)
Gender		
Female, n (%)	8430 (20.8%)	5805 (51.0%)
Male, n (%)	32,152 (79.2%)	5584 (49.0%)
Maternal age at delivery, mean (SD)	29.5 (6.0)	29.1 (5.5)
Paternal age at delivery, mean (SD)	31.8 (6.6)	31.5 (6.2)

Due to the rarity of some exposures, the association between 34 of these pairs could not be estimated among children with ASD. Among the remaining associations, we observed that exposure was associated with predominantly higher, rather than lower rates of comorbidity. The exposures associated with the highest number of comorbidities included preterm birth, hypoxia at birth, and fetal alcohol syndrome.

Adjusting the associations for other comorbidities within the same domain (related comorbidities), the magnitude of the odds ratios for some of the associations attenuated (Fig. 4b). Although the number of significant associations reduced from 100 to 55, the overall pattern remained similar as in the previous analyses, with positive associations accounting for a large portion of the total significant associations. Tables 2S–9S provide frequency and associations of the exposure-outcome pairs. Although the figures present the point estimates for the assessed associations and their statistical significance, the uncertainty around these estimates is captured by the corresponding confidence intervals presented in Tables 2S–9S.

Association of pre-and postnatal exposures with comorbidity among non-ASD siblings

To provide further context for interpretation of the results, we estimated the associations between exposures and ASD-comorbid conditions among the non-ASD siblings. Preterm birth and hypoxia at birth were the most common exposures in the non-ASD children, allowing for evaluation of their associations with comorbidity within this smaller sample. Figure S7 presents these associations alongside the estimates obtained in the sample of individuals with ASD. Similarly as in individuals with ASD, short stature, difficulty gaining weight, and motor delay had the strongest association with preterm birth, while motor delay, disruptive mood dysregulation disorder, and social communication disorder were the three comorbidities with the strongest association with hypoxia at birth. Most of the exposure-comorbidity estimates were suggestive of potential positive associations ($OR > 1$), however, for several of them, the width of the confidence intervals prevented from drawing clear conclusions.

DISCUSSION

We estimated the prevalence of pre- and postnatal exposures, as well as comorbidity patterns by ASD status, replicating the widely reported increased burden of co-occurring health conditions among individuals with ASD. Our analyses show that the burden of comorbidity in individuals with ASD remains elevated even in comparison to their non-ASD siblings, suggesting that shared familial factors are unlikely to fully account for these observations.

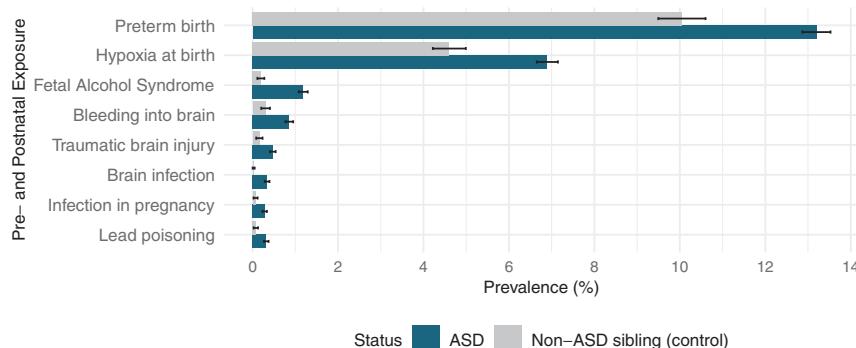


Fig. 1 Barplots illustrating the prevalence (%) of pre- and postnatal exposures by ASD status. Prevalence estimates in unaffected siblings (grey bars) were standardized to year of evaluation, age and sex distribution in ASD cases (blue bars).

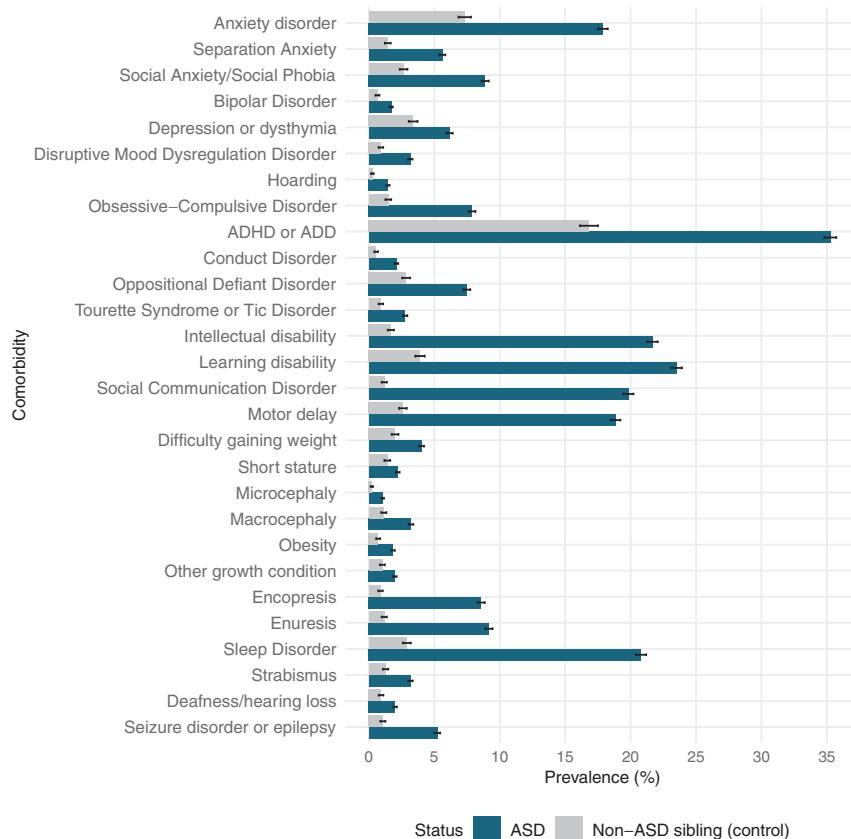


Fig. 2 Barplots illustrating the prevalence of comorbid conditions by ASD status. Prevalence estimates in unaffected siblings (grey bars) were standardized to year of evaluation, age and sex distribution in ASD cases (blue bars).

Our results suggest that the pre- and postnatal exposures considered in this study are both (i) more common in children with ASD, compared to their non-ASD siblings, and (ii) associated with an increased burden of specific comorbidities—potentially indicating the ‘pleiotropic’ effects of these exposures (or their genetic underpinnings), on both ASD and its comorbidities (Fig. 5). Furthermore, (iii) observing these exposure-comorbidity associations among the non-ASD siblings indicates that they may occur independently of the ASD diagnosis. These results suggest that the higher rates of certain comorbidities in ASD may be partly attributable to the higher rates of the underlying risk factors (environmental exposures, or the underlying genetic variation) among the affected individuals, rather than to downstream effects of ASD itself.

In the initial analyses adjusting for covariates, we evaluated associations between 190 exposure-comorbidity pairs among individuals with ASD, observing predominantly positive associations. Further adjustment for other comorbid conditions potentially correlated with the outcome in each model decreased the number of significant associations, suggesting that some of the associations observed in the initial analyses were attributable to the clustering (positive correlation) of comorbid conditions. Nevertheless, even after this adjustment, majority of the significant associations persisted, with hypoxia at birth, preterm birth, and fetal alcohol syndrome accounting for most of the significant exposure-comorbidity associations.

Notably, many of these associations are in line with those reported in samples not ascertained for ASD diagnosis, including the association between strabismus and both pre-term birth [29] and fetal alcohol syndrome [30], and exposures directly affecting the brain (bleeding into brain, hypoxia at birth, brain infection,

and traumatic brain injury) and epilepsy [31, 32]. This suggests that the etiology of conditions that tend to be comorbid in ASD may be not distinctly different from the etiology of the same conditions in the general population. This is also in line with the associations observed in non-ASD siblings in our study. While the rates of comorbidities were lower in siblings, many of the exposure-outcome associations remained significant in individuals without an ASD diagnosis, corroborating studies in the general population [29–32] and indicating that the recorded associations were not attributable to the ASD diagnosis alone.

This pattern of results, whereby exposures are seemingly independently associated with both ASD status and its comorbidities, could suggest their independent effects on these different outcomes. Such ‘pleiotropic’ effects (underpinned, or not, by the genetic variation) could help explain the high frequency of comorbid diagnoses in ASD. Given the controls in our study were siblings of cases, it remains possible that the liability correlated with the exposure exerts its effects on both ASD and medical comorbidities—rendering the effects observed in individuals with ASD and their siblings quantitatively, but *not qualitatively* different.

The study findings have several clinical and research implications. The higher rates of co-occurring medical conditions [9], along the health disparities [33] among individuals with ASD, highlights the need for closer monitoring and better screening of these individuals for additional medical conditions. The positive correlations observed between certain exposures and comorbidities can potentially help with the early identification and treatment of these comorbidities. The exposure distribution and comorbidity pattern in individuals with ASD may also help explain the heterogeneity of ASD, thus potentially

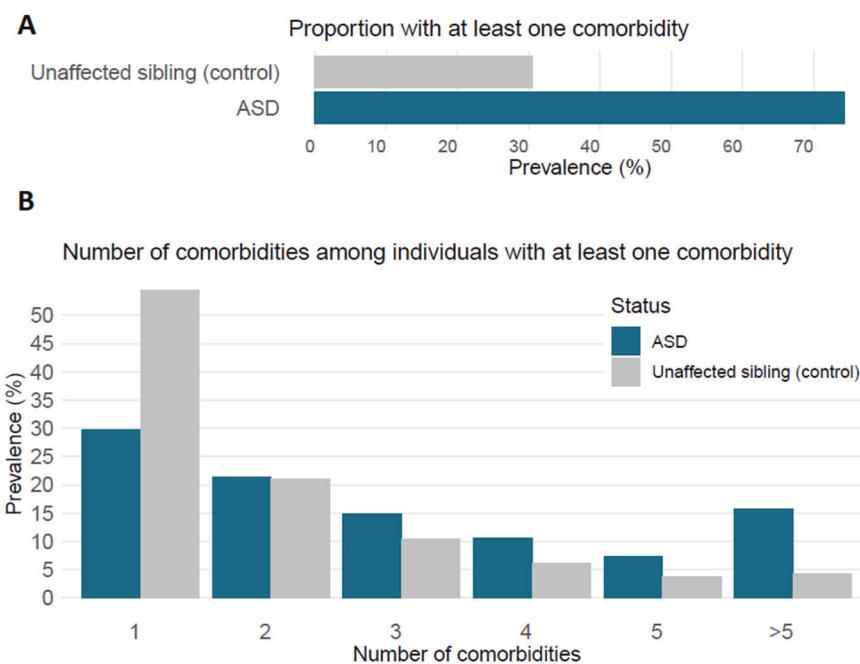


Fig. 3 Burden of comorbid conditions in ASD cases and their unaffected siblings. **A** Prevalence of individuals with >1 comorbidity by ASD status. **B** Distribution of the number of comorbidities among those with at least one comorbidity by ASD status.

offering a scope for stratification of these individuals for diagnosis and/or research purposes. Finally, these associations and their comparisons with effects in unaffected individuals can provide further insights into the complex etiology of ASD and its accompanying comorbidities.

The study has several limitations. First, the study population is self-selected and, despite the efforts to increase representation of diverse groups [34], given the non-probabilistic sampling and recruitment of study participants, we cannot exclude the risk of selection bias. Nevertheless, the prevalence of comorbid conditions observed in the SPARK sample approximate those reported elsewhere, increasing the generalizability of our results, and suggesting limited impact of self-selection on study findings. For example, the reported prevalence of epilepsy was 5% in SPARK, cf. 7% in a recent meta-analysis [35]; similarly, ADHD was reported in 35% of children in the SPARK study, and 28.2% (95% CI: 13.3–43.0) in a sample derived from population-based cohort [36]. Additionally, while all participants in our study were born in or before 2019, medical evaluation of a small fraction of them could have occurred during the COVID-19 pandemic. While the effects of the pandemic on our findings would be limited by the fact that medical evaluation reflected previously reported, rather than current conditions, we were not able to fully evaluate the magnitude of COVID-19's impact on our results due to inability to ascertain the exact individuals evaluated during the pandemic. Finally, our results represent the associations observed in broadly defined ASD, including both its familial and sporadic forms, with likely distinct etiologies. It remains to be elucidated how polygenic load, rare high-impact variants, and other non-genetic factors interact with the pre- and post-natal exposures interrogated in our study to result in an array of medical comorbidities.

Although the associations in our study are adjusted for a pool of covariates known to be correlated with pre- and postnatal exposures and/or comorbidities, no conclusions can be drawn about causality, especially given potential measurement error and uncontrolled confounding [37] (e.g., access to care and genetic factors). Of note, even though adjusting for additional comorbidities yields exposure-comorbidity associations that are independent of other correlated comorbidities, such adjustments do not

always result in estimates that are closer to the causal parameters. We also cannot rule out the possibility of measurement error (e.g., due to self-reporting, recall bias, or difficulty of diagnosing comorbidity in children with ASD) affecting the estimated associations. Nevertheless, observing similar patterns as reported in other general population samples and the existence of several null and a few negative associations are suggestive that potential measurement error is unlikely to be a major concern. The potential for relative overdiagnosis of comorbidity in children with ASD, as well as possible polypharmacy among them, could have contributed to the higher burden of comorbidity in ASD children compared to non-ASD siblings. Nevertheless, although we did not have data to explore potential impacts of these factors, it is unlikely that the observed differences in prevalence of comorbidities between these populations could be primarily attributable to these factors.

A few of the covariates had high rates of missing values, nevertheless, we conducted multiple imputations, which, contingent on untestable assumptions regarding the missingness patterns, can eliminate missing data bias. The sensitivity analysis restricting the analytical sample to those with complete data yielded results that were consistent with those obtained from the analysis of the dataset with imputed missing covariates. Additionally, coding of some variables in SPARK does not distinguish between missing and negative values, therefore, there is a chance that some of the observations coded as "null" or "no" were in fact missing, and thus potential false negatives. Given that the observed associations corroborate those observed in other, non-ASD samples, potential bias is minimal. Lack of data on the date of diagnosis of comorbidities limited our analytical approach and did not allow 'time to event' analysis in order to account for the fact that some of the individuals not yet diagnosed with ASD/comorbidity may receive a diagnosis later in life. Although SPARK is one of the largest ASD samples, the low prevalence of certain exposures and outcomes limited our statistical power in evaluating their associations, especially in the smaller sample of the non-ASD siblings.

In conclusion, this study followed a systematic approach, evaluating pre- and postnatal exposure and comorbidity

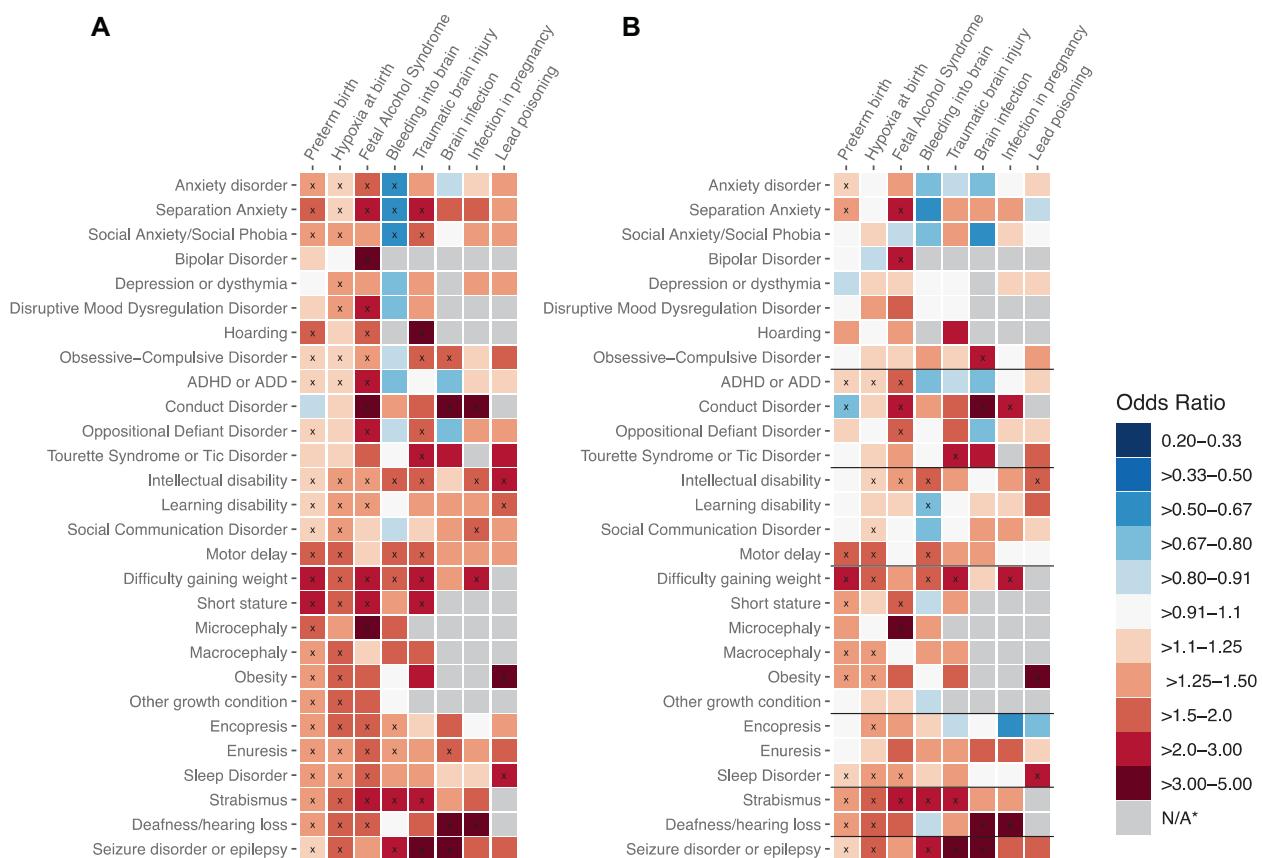


Fig. 4 Associations between pre- and postnatal exposures and comorbidity in individuals with ASD. **A** The model for each comorbidity was adjusted for maternal and paternal education and age at delivery, child's gender and year of birth, age of the child at evaluation, survey version, race, and annual household income. **B** In addition to the covariates above, the model for each comorbidity was adjusted for other comorbidities within the same, SPARK-defined disorder domain (domains are demarcated with solid black lines). Blue tiles indicate lower and red tiles higher probability of the diagnosis of the given comorbid condition (y-axis) in individuals with certain environmental exposure (x-axis). Grey tiles indicate the associations that could not be estimated due to insufficient number of exposed individuals with the comorbidity to allow reliable coefficient estimate. * p value < 0.05.

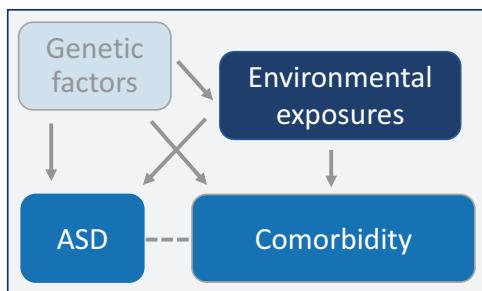


Fig. 5 Schematic illustrating potential relationships between genetic vectors, environmental exposures, ASD and its comorbid conditions.

associations in ASD. We demonstrated that the common comorbidities in individuals with ASD are often associated with pre- and postnatal exposures also linked to ASD [38]. The higher prevalence of comorbidity in ASD, even compared to non-ASD siblings, draw attention to the importance of timely diagnosis and management of comorbidity in ASD, which in turn could offer improved prognosis and quality of life among these individuals. Finally, our preliminary evidence from associations between pre- and postnatal exposures and later comorbidities in the non-ASD siblings suggests that these associations are not limited to ASD and could partly be due to common underlying cause/s.

REFERENCES

- Baio J, Wiggins L, Christensen DL, Maenner MJ, Daniels J, Warren Z, et al. Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2016. MMWR Surveill Summ. 2020;69:1–12.
- Hill AP, Zuckerman KE, Fombonne E. Obesity and autism. Pediatrics. 2015;136:1051–61. <https://doi.org/10.1542/peds.2015-1437>.
- Buck TR, Viskochil J, Farley M, Coon H, McMahon WM, Morgan J, et al. Psychiatric comorbidity and medication use in adults with autism spectrum disorder. J Autism Dev Disord. 2014;44:3063–71. <https://doi.org/10.1007/s10803-014-2170>.
- Richdale AL, Schreck KA. Sleep problems in autism spectrum disorders: prevalence, nature, & possible biopsychosocial aetiologies. Sleep Med Rev. 2009;13:403–11. <https://doi.org/10.1016/j.smrv.2009.02.003>.
- Mouridsen SE, Rich B, Isager T. Epilepsy in disintegrative psychosis and infantile autism: a long-term validation study. Dev Med Child Neurol. 1999;41:S0012162299000213. <https://doi.org/10.1017/S0012162299000213>.
- Horvath K, Perman JA. Autistic disorder and gastrointestinal disease. Curr Opin Pediatr. 2002;14:583–7. <https://doi.org/10.1097/00008480-200210000-00004>.
- Egan AM, Dreyer ML, Odar CC, Beckwith M, Garrison CB. Obesity in young children with autism spectrum disorders: prevalence and associated factors. Child Obes. 2013;9:125–31. <https://doi.org/10.1089/chi.2012.0028>.
- Vohra R, Madhavan S, Sambamoothi U. Comorbidity prevalence, healthcare utilization, and expenditures of Medicaid enrolled adults with autism spectrum disorders. Autism. 2017;21:995–1009. <https://doi.org/10.1177/1362361316665222>.
- Croen LA, Zerbo O, Qian Y, Massolo ML, Rich S, Sidney S, et al. The health status of adults on the autism spectrum. Autism. 2015;19:814–23. <https://doi.org/10.1177/1362361315577517>.
- Kirsch AC, Huebner ARS, Mehta SQ, Howie FR, Weaver AL, Myers SM, et al. Association of comorbid mood and anxiety disorders with autism spectrum disorder. JAMA Pediatr. 2020;174:63. <https://doi.org/10.1001/jamapediatrics.2019.4368>.

11. Belardinelli C, Raza M, Taneli T. Comorbid behavioral problems and psychiatric disorders in autism spectrum disorders. *J Child Dev Disord*. 2016;02. <https://doi.org/10.4172/2472-1786.100019>
12. Lundström S, Reichenberg A, Melke J, Råstam M, Kerekes N, Lichtenstein P, et al. Autism spectrum disorders and coexisting disorders in a nationwide Swedish twin study. *J Child Psychol Psychiatry*. 2015;56:702–10. <https://doi.org/10.1111/jcpp.12329>.
13. Tang L, Levy J, Guillory S, Halpern D, Zweifach J, Giserman-Kiss I, et al. Prospective and detailed behavioral phenotyping in DDX3X syndrome. *Mol Autism*. 2021;12:36. <https://doi.org/10.1186/s13229-021-00431-z>.
14. Miles JH, Takahashi TN, Bagby S, Sahota PK, Vaslow DF, Wang CH, et al. Essential versus complex autism: definition of fundamental prognostic subtypes. *Am J Med Genet Part A*. 2005;135A:171–80. <https://doi.org/10.1002/ajmg.a.30590>.
15. Feliciano P, Daniels AM, Snyder LG, Beaumont A, Camba A, Esler A, et al. SPARK: A US Cohort of 50,000 families to accelerate autism research. *Neuron*. 2018;97:488–93. <https://doi.org/10.1016/j.neuron.2018.01.015>.
16. Rai D, Lewis G, Lundberg M, Araya R, Svensson A, Dalman C, et al. Parental socioeconomic status and risk of offspring autism spectrum disorders in a Swedish population-based study. *J Am Acad Child Adolesc Psychiatry*. 2012;51:467–76.e6. <https://doi.org/10.1016/j.jaac.2012.02.012>.
17. Thomas P, Zahorodny W, Peng B, Kim S, Jani N, Halperin W, et al. The association of autism diagnosis with socioeconomic status. *Autism*. 2012;16:201–13. <https://doi.org/10.1177/1362361311413397>.
18. Durkin MS, Maenner MJ, Meaney FJ, Levy SE, DiGuiseppi C, Nicholas JS, et al. Socioeconomic inequality in the prevalence of autism spectrum disorder: evidence from a U.S. cross-sectional study. Myer L, ed. *PLoS One*. 2010;5:e11551. <https://doi.org/10.1371/journal.pone.0011551>
19. Mandell DS, Wiggins LD, Carpenter LA, Daniels J, DiGuiseppi C, Durkin MS, et al. Racial/ethnic disparities in the identification of children with autism spectrum disorders. *Am J Public Health*. 2009;99:493–8. <https://doi.org/10.2105/AJPH.2007.131243>.
20. Chen E, Martin AD, Matthews KA. Understanding health disparities: the role of race and socioeconomic status in children's health. *Am J Public Health*. 2006;96:702–8. <https://doi.org/10.2105/AJPH.2004.048124>.
21. Ekéus C, Otterblad Olausson P, Hjern A. Psychiatric morbidity is related to parental age: a national cohort study. *Psychol Med*. 2006;36:269–76. <https://doi.org/10.1017/S0033291705006549>.
22. Mayo JA, Lu Y, Stevenson DK, Shaw GM, Eisenberg ML. Parental age and preterm birth: a population-based cohort of nearly 3 million California livebirths from 2007 to 2012. *J Perinatol*. Published online December 8, 2020. <https://doi.org/10.1038/s41372-020-00894-7>
23. Di Renzo GC, Rosati A, Sarti RD, Cruciani L, Cutuli AM. Does fetal sex affect pregnancy outcome? *Gend Med*. 2007;4:19–30. [https://doi.org/10.1016/S1550-8579\(07\)80004-0](https://doi.org/10.1016/S1550-8579(07)80004-0).
24. Werling DM, Geschwind DH. Sex differences in autism spectrum disorders. *Curr Opin Neurol*. 2013;26:146–53. <https://doi.org/10.1097/WCO.0b013e32835ee548>.
25. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Third. Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
26. Greenland S, Mansournia MA, Altman DG. Sparse data bias: a problem hiding in plain sight. *BMJ*. 2016;352:i1981. <https://doi.org/10.1136/bmj.i1981>.
27. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B*. 1995;57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
28. Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med*. 2016;4:30. <https://doi.org/10.3978/j.issn.2305-5839.2015.12.63>.
29. Gulati S, Andrews CA, Apkarian AO, Musch DC, Lee PP, Stein JD. Effect of gestational age and birth weight on the risk of strabismus among premature infants. *JAMA Pediatr*. 2014;168:850. <https://doi.org/10.1001/jamapediatrics.2014.946>.
30. Strömlund K. Ocular abnormalities in the fetal alcohol syndrome. *Acta Ophthalmol Suppl (Oxf)*. 1985;171:1–50.
31. Annegers JF, Coan SP. The risks of epilepsy after traumatic brain injury. *Seizure*. 2000;9:453–7. <https://doi.org/10.1053/seiz.2000.0458>.
32. Sung CY, Chu NS. Epileptic seizures in intracerebral haemorrhage. *J Neurol Neurosurg Psychiatry*. 1989;52:1273–6. <https://doi.org/10.1136/jnnp.52.11.1273>.
33. Bishop-Fitzpatrick L, Kind AJH. A scoping review of health disparities in autism spectrum disorder. *J Autism Dev Disord*. 2017;47:3380–91. <https://doi.org/10.1007/s10803-017-3251-9>.
34. SPARK: Simons Foundation Powering Autism Research for Knowledge. The Importance of Racial and Ethnic Diversity in Autism Research. Published 2023. Accessed January 14, 2023. https://sparkforautism.org/discover_article/diversity-in-autism-research/

ACKNOWLEDGEMENTS

We are grateful to all of the families in SPARK, the SPARK clinical sites and SPARK staff. We would like to acknowledge the generous support of the Beatrice and Samuel A. Seaver Foundation.

AUTHOR CONTRIBUTIONS

VK and MJ conceived the study idea; VK and MJ planned the analysis. VK and MJ obtained the funding. VK carried out the analysis and prepared the first draft of the manuscript. MJ provided supervision. All authors contributed to the design of the study, interpretation of the findings, and critical revisions. All authors approved the version to be published and are accountable for the accuracy and integrity of their areas of involvement in the study.

FUNDING

VK was supported by National Institute of Mental Health Award T32 MH122394. MJ, AR and SS were supported by MH123417. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the authors' employers.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41398-023-02374-w>.

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

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023