DNA Methylation, Environmental Adversities and Attention Problems in Children Born Very Preterm

Samiha Kazi, BS[1]; Barry M. Lester, PhD[2,3,4]; Todd M. Everson, PhD[5]; Carmen J. Marsit, PhD[5]; Brian S. Carter, MD[6], Jennifer Check, MD[7], Jennifer Helderman, MD[7], Julie A. Hofheimer, PhD[8], Elisabeth C. McGowan, MD[3], Charles R. Neal, MD[9], Steven L. Pastyrnak, PhD[10], Lynne M. Smith, MD[11], Thomas Michael O'Shea, MD[8], Marie Camerota, PhD*[3,4]

[1]Department of Biology, Brown University, Providence, RI
[2]Department of Pediatrics, Alpert Medical School of Brown University and Women and Infants Hospital, Providence, RI
[3]Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Providence, RI
[4]Brown Center for the Study of Children at Risk, Alpert Medical School of Brown University and Women and Infants Hospital, Providence, RI
[5]Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA
[6]Department of Pediatrics-Neonatology, Children's Mercy Hospital, Kansas City, MO
[7]Department of Pediatrics, Wake Forest School of Medicine, Winston-Salem, NC
[8]Department of Pediatrics, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC
[9]Department of Pediatrics, University of Hawaii John A. Burns School of Medicine, Honolulu, HI
[10]Department of Pediatrics, Spectrum Health-Helen DeVos Hospital, Grand Rapids, MI
[11]Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA

***Corresponding Author:** Marie Camerota, PhD, Brown Center for the Study of Children at Risk, Women and Infants Hospital of Rhode Island, 101 Dudley Street, Providence, RI, 02905, USA; Phone: 401-274-1122, ext. 48905; Email: marie_camerota@brown.edu

Abstract

Children born preterm are at higher risk of attention problems compared to children born at term, yet not all preterm children are affected. The purpose of this study was to identify epigenetic predictors of early childhood attention problems among children born very preterm and to examine the joint role of epigenetic and environmental factors in predicting attention problems in this population. We studied 242 participants from a multi-site study of infants born < 30 weeks gestational age. Neonatal buccal swabs were assayed for DNA methylation levels at over 450,000 CpG sites and age acceleration metrics were calculated using existing epigenetic clocks. A composite of postnatal environmental adversity was calculated using maternal reported risk factors. Attention problems were assessed in early childhood (mean age 6.58 years) using the Conner's Kiddie Continuous Performance Test 2nd Edition. After adjustment for multiple testing, we found that DNA methylation of 9 CpG sites was associated with childhood attention problems. Several CpGs were located in genes previously linked to neurodevelopmental traits and inflammation in prior epigenome-wide and genome-wide association studies. Greater environmental adversity was also associated with increased attention problems. When tested together, DNA methylation and environmental adversity independently predicted attention problems. This study is the first to show associations between DNA methylation, environmental adversity, and objectively measured attention problems in school-age children born very preterm. These results could shed light on the etiology of attention problems in this population and may help us identify at birth preterm children at highest risk for later ADHD diagnosis.

DNA Methylation, Environmental Adversities and Attention Problems in Children Born Very Preterm

Children who are born preterm (< 37 weeks gestational age [GA]) face an increased risk of developing behavioral and neurodevelopmental problems (e.g., Aarnoudse-Moens et al., 2009). Attention problems including Attention Deficit Hyperactivity Disorder (ADHD) are one of the most common psychiatric conditions faced by preterm infants (Franz et al., 2018). Children born preterm are at 2 to 4 times higher risk for developing ADHD compared to term-born peers, with the highest risk among children born very preterm (< 32 weeks GA; Johnson & Marlow, 2011). Studying the development of attention problems in children born very preterm could help identify children at highest risk for long-term impairments who might benefit from treatment or early intervention (Pingault et al., 2014).

Although infants born very preterm are at higher risk of developing attention problems, there is great variability in outcomes among this population (e.g., Camerota et al., 2022). Different environmental and biological mechanisms likely contribute to these discrepancies (Thapar et al., 2013). Early-life environmental adversities, such as low socioeconomic status and limited access to healthcare, have been shown to have detrimental effects on individuals' physical and mental health, and are associated with ADHD diagnoses and symptom severity (Spencer et al., 2022). Prior studies have primarily focused on investigating risk factors for attention problems among full-term children with relatively few studies specifically investigating risk for attention problems among children born very preterm. Some the environmental risk factors associated with attention problems in very preterm populations are similar to those investigated in full-term populations, such as maternal antenatal smoking and low maternal educational attainment (Downey et al., 2015).

Beyond environmental factors, there is also a strong genetic component underlying the risk for attention problems (Thapar et al., 2013). In addition to differences in genotype, differential regulation of gene expression is emerging as an important predictor of attention problems in children (Cecil & Nigg, 2022). Epigenetics refers to the molecular processes that can modulate gene expression without directly changing DNA sequences themselves. Most human studies focus on DNA methylation as an epigenetic mechanism (Hamza et al., 2019). DNA methylation refers to the addition of a methyl group in an area of the DNA sequence where a cytosine nucleotide

is followed by a guanine nucleotide (termed a CpG site). DNA methylation of a portion of DNA can change the likelihood of gene expression, with consequences for child behavior and development (Lester et al., 2016).

Epigenome-wide association studies (EWAS) are a way to link DNA methylation of specific CpG sites to disease risk (Wei et al., 2021). Several prior EWAS have investigated the association between DNA methylation and ADHD (Chen et al., 2018; Goodman et al., 2020; Mooney et al., 2020; Neumann et al., 2020; Walton et al., 2017; Wang et al., 2020; Wilmot et al., 2016). These studies have found significant associations between methylation of various CpGs and child attention problems (Neumann et al., 2020). DNA methylation at birth seems particularly influential: a prior study found that nine CpGs were identified at birth that predicted symptoms, while no CpGs at school age were associated with concurrent ADHD symptoms (Neumann et al., 2020). One EWAS was previously conducted in the same sample of very preterm infants studied here, but with a primary focus on parent-reported attention problems at age 2 (Camerota, Lester, Castellanos, et al., 2024). This prior study provides evidence that neonatal DNA methylation at numerous CpG sites is associated with early-emerging attention problems in children born very preterm.

Besides EWAS, another approach to analyzing epigenetic data is to construct polyepigenetic scores. Epigenetic clocks are one type of polyepigenetic score that have been created to study biological aging. Depending on the population, biological aging can refer to developmental maturity (in pediatric populations) or bodily decline (adult or geriatric populations). Epigenetic clocks were originally developed in adults, and research using these clocks showed that age acceleration (i.e., an individual having greater epigenetic age in relation to their chronological age) predicted greater risk for chronic diseases and mortality (Fransquet et al., 2019). Recent clocks such as PedBE (McEwen et al., 2020) and NEOage clocks (Graw et al., 2021) have been developed for use with children. Our group specifically developed the NEOage clocks to estimate gestational and chronological age acceleration among very preterm infants, with the goal of improving prediction of long-term outcomes in this high risk group. Studying preterm infants from an epigenetic age lens is interesting since this population of infants is born before they should be, making it unclear whether age acceleration (as a potential marker for biological maturity) predicts better or worse outcomes. We have previously shown that the NEOage clock is associated with neonatal neurobehavioral characteristics

in the current sample (Paniagua et al., 2023), yet it remains unknown whether the NEOage clock predicts longer-term outcomes for this population, such as attention problems. A recent study found that epigenetic estimates of GA were inversely associated with ADHD symptoms in early childhood but that epigenetic age acceleration did not uniquely predict ADHD risk (Salontaji et al., 2024). However this prior study was primarily comprised of term-born children. Thus, this is the first study to investigate epigenetic age acceleration and attention problems in very preterm children.

Studies have yet to compare results from EWAS and epigenetic clock analyses to understand epigenetic risk factors for attention problems among children born very preterm. Additionally, the existing epigenetic studies of ADHD have yet to investigate the joint role of environmental and epigenetic factors as predictors of attention problems, though our prior work shows that both are important predictors of cognitive outcomes in children born very preterm (Camerota et al., 2024). In the current study, we aimed to (1) conduct an EWAS to examine epigenetic predictors of attention problems in early childhood (ages 5-7) and pinpoint individual CpG sites associated with children's inattention, (2) test associations between neonatal epigenetic age acceleration and attention problems, and (3) test the joint contribution of environmental and epigenetic factors in predicting attention problems among children born very preterm. We hypothesized that epigenetic and environmental risk factors would both predict risk for attention problems in this population.

Methods

Participants and Procedures

Participants were drawn from the Neonatal Neurobehavior and Outcomes in Very Preterm
Infants (NOVI) Study, a multi-site study of infants born < 30 weeks GA. Participants were
recruited from nine university affiliated NICUs across six research sites from April 2014 to June
2016. Inclusion criteria were: (1) birth < 30 weeks GA, (2) parental ability to speak English or
Spanish, (3) residence within 3 hours of the NICU and follow-up clinic. Exclusion criteria

included major congenital anomalies, maternal age < 18 years, cognitive impairment, and death. Parents of eligible infants were approached when infants were 31-32 weeks GA or when survival to discharge was deemed likely by the attending neonatologist. Researchers at each site obtained informed consent in line with each institution's review board.

Medical and demographic information of participants were collected through medical record abstraction and maternal report at the time of enrollment. Buccal swabs were collected at the time of NICU discharge. Child neurobehavioral assessments were conducted at yearly follow-up visits from ages 2 to 7 years. Participants were included in this analysis if they had neonatal DNA methylation data and at least one assessment of attention problems from ages 5 to 7 years.

Measures

Neonatal DNA Methylation

Genomic DNA was extracted from buccal swab samples, collected near term-equivalent age,
using the Isohelix Buccal Swab system (Boca Scientific), quantified using the Quibit
Fluorometer (Thermo Fisher, Waltham, MA, USA) and aliquoted into a standardized
concentration for subsequent analyses. DNA samples were randomly plated across 96-well plates
and given to the Emory University Integrated Genomics Core for bisulfite modification using
the EZ DNA Methylation Kit (Zymo Research, Irvine, CA), and subsequent assessment of
genome-wide DNAm using the Illumina MethylationEPIC Beadarray (Illumina, San Diego, CA)
following standardized methods based on the manufacturer's protocol.

Pre-processing of data followed a previously described workflow (Everson et al., 2020). Array data underwent Noob normalization (Aryee et al., 2014; Liu & Siegmund, 2016). Samples with poor detection p-values or sex-mismatch were excluded. We excluded probes with median detection p-values < 0.05, those on the X or Y chromosome, those with single nucleotide polymorphisms (SNP) within the binding region, and those that could cross-

hybridize to other regions of the genome (Pidsley et al., 2016). Array data were standardized across Type-I and Type-II probe designs with beta-mixture quantile normalization (Pidsley et al., 2013; Teschendorff et al., 2013).

We then took measures to diminish multiple testing burden and increase our ability to identify meaningful associations. First, we administered the CoMeBack pipeline to identify co-methylated regions (CMRs) which are clusters of highly-correlated, proximal CpG sites (Gatev et al., 2020). Principal components analysis is conducted for each CMR and the first principal component is allocated to each cluster as a summary of DNAm levels at that CMR. The CoMeBack pipeline identified 73,746 CMRs representing the DNAm of 206,195 CpG sites; 500,128 CpG sites were not included in CMRs and were preserved as individual CpG sites. Next, we kept out CpGs or CMRs with low variability (SD < 0.02); sites with low variability are more prone to measurement error and are less likely to result in duplicable findings (Logue et al., 2017). To decrease the possibility of invalid or non-reproducible findings, we investigated each CpG and CMR for outliers and documented values that fell 3 interquartile ranges (IQR) below the 25th percentile or 3 IQR above the 75th percentile to missing.

After exclusions and data reduction, 452,453 loci (60,917 CMRs and 391,536 CpGs) were obtainable from 542 samples for this study (83% of 651 with buccal swab consent; 77% of entire NOVI cohort). For simplicity in the results, we refer to each loci as a CpG but note where significant results were located in a CMR. The DNAm data used in this study are attainable through NCBI Gene Expression Omnibus (GEO) via accession series GSE128821.

Conners Kiddie Continuous Performance Test (K-CPT2)

The Conners Kiddie Continuous Performance Test 2nd Edition (K-CPT 2) is an assessment of attention deficits for children aged 4-7 years that is widely used across clinical settings (Shaked et al., 2020). The K-CPT 2 has been shown to differentiate between preschool children with and without ADHD (I.-C. Chen et al., 2021). The K-CPT2 consists of five blocks (sets of trials)

with two sub-blocks made up of 20 trials each. In each trial, participants are shown an image of an object on the computer screen. They must respond (press a keyboard button) when they see certain objects (“targets”) but refrain from responding when they see one specific object (“non-targets”). Each block consists of two sub-blocks with a short (1.5 second) and long (3 second) inter-stimulus interval (ISI). Trials with short ISIs track fast-rate tasks and assess participants’ rapid information processing while trials with long ISIs track slow-rate tasks and assess participants’ alertness. Participants complete a total of 200 trials over 7.5 minutes. In this analysis, participants were considered to have valid data if they completed at least 75% of trials. This standardized, computerized assessment produces a number of scores related to four dimensions of attention: Inattentiveness, Impulsivity, Sustained Attention and Vigilance. Omission errors (i.e., number of missed targets) are a commonly-used metric of inattention. Age-adjusted omission error T-scores were the primary outcome in this analysis.

Participants in NOVI completed the K-CPT at either the 5, 6, and/or 7 year follow-up visits. For participants who completed the K-CPT at multiple visits, only the first administration was used.

Age Acceleration

Age acceleration is defined as the difference between an individual’s chronological age and their epigenetic age. In line with prior studies in this sample (Graw et al., 2021), two chronological age metrics were examined. Postnatal age (PNA) in the NOVI study was defined as the elapsed time from an infant’s birth to their buccal swab collection at NICU discharge. Post-menstrual age (PMA) was defined as the elapsed time from an infant’s conception to the time of buccal collection at NICU discharge. Epigenetic age estimates of PMA and PNA were determined using the NEOage epigenetic clocks compatible with the Illumina EPIC array derived from infants in the NOVI cohort (Graw et al., 2021). To prevent overfitting in the derivation of epigenetic age, we implemented a leave-one-out strategy, where each participant’s predicted epigenetic age was based on a training data set that excluded them and their siblings. For both PMA and PNA, age acceleration was defined as the residuals when epigenetic age is regressed on chronological age in an unadjusted linear model.

Environmental Adversity (Discharge to 48 months)

We constructed a cumulative environmental adversity index for children based on data from 3 timepoints in the study (24 months; 36 months; 48 months). At each timepoint, six risk factors were assessed: (a) limited socioeconomic resources, defined as household income < 150% of the poverty line and/or Hollingshead level V; (b) limited childcare assistance, defined as no reported secondary caregiver; (c) caregiver psychiatric disorder, defined as primary caregiver scoring above the clinical cutoff for depression and/or anxiety (T-score ≥ 63) on the Brief Symptom Inventory (Derogatis & Melisaratos, 1983), in the mild to severe depression range (total score ≥ 6) on the Quick Inventory of Depressive Symptomology, or self-report of any diagnosis, treatment, or counseling for depression and/or anxiety; (d) family disruption/separation, defined as any reported Child Protective Services involvement; (e) high stressful life events, defined as the sum of stressful life events >1SD above the sample mean; and (f) crowded household, defined as four or more children reported to be living in the household (**Table 1**). The presence of each risk factor was scored as one point. We computed a cumulative adversity index at each of the 3 timepoints calculated as the proportion of encountered risk factors. For this analysis, we took the average score across all 3 timepoints as an approximation of overall postnatal environmental adversity (up to 48 months).

Covariates

Since DNAm levels tend to vary within different types of cells, it is important to account for this heterogeneity when using mixed cell samples, such as buccal tissue in this study. The proportions of epithelial, fibroblast and immune cells in the buccal tissue samples for our study were calculated using existing reference methylomes (Zheng et al., 2018). Due to an inverse relationship between the immune and epithelial cell proportions in our collected data, we adjusted all epigenetic analyses for epithelial cell proportion in order to account for cellular heterogeneity. We also accounted for batch effects in all epigenetic analyses by adjusting for sample plate.

In addition to these technical covariates, all epigenetic models were also adjusted for infant sex, study site, age at buccal swab collection, gestational age, neonatal medical morbidities, and maternal prenatal smoking (obtained from her medical record). Neonatal medical morbidities were defined as the sum of four possible morbidities: bronchopulmonary dysplasia, serious brain

injury, retinopathy of prematurity, and necrotizing enterocolitis/sepsis (Bassler et al., 2009; McGowan et al., 2022).

Statistical Analysis

First, we conducted an EWAS to investigate the association of DNAm at each of 452,453 CpG sites and child attention problems at age 5-7 years. We used generalized estimating equation (GEE) models with robust standard errors to regress K-CPT 2 omission error T-scores (dependent variable) on DNAm at each CpG site, considering the nesting of children within families and covariates (study site, infant sex, infant GA at birth, infant GA at buccal swab, maternal prenatal smoking, neonatal medical morbidities, cell type composition [proportion of epithelial cells], and sample plate). P-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) (Benjamini & Hochberg, 1995). CpG sites associated with attention problems within a 5% FDR cutoff were deemed significant. To ensure meaningful interpretation, we rescaled DNAm at each CpG site by dividing the raw data by the CpG-specific interquartile range (IQR) such that the coefficients obtained from the GEE models represent the expected change in attention problem T-scores associated with a change in DNAm from the 25th to the 75th percentile of observed data.

Next, we estimated two GEE models to examine the association between epigenetic age acceleration (PMA and PNA) and attention problem T-scores. In these models, age acceleration was the independent variable and attention problem T-scores were the dependent variable. The same covariates listed above were used in these models.

Finally, we tested the association of both environmental adversity and DNA methylation with attention problems. Initially, we examined environmental adversity as a primary focal predictor, by utilizing the cumulative environmental adversity index as the independent variable and attention problems as the dependent variable. Then, we estimated joint models that included both environmental adversity and any significant epigenetic variables (either DNA methylation of specific CpGs from the EWAS or either PNA or PMA acceleration). These joint models allowed us to compare coefficients to understand how the relationship between epigenetics and attention problems changed when additionally accounting for environmental adversity, and vice versa.

Results

Of the 704 infants that were enrolled in the NOVI Study, 242 were included in this analysis (Figure 1). Children included in this analysis had fewer neonatal medical morbidities compared to those excluded (0.71 vs 0.97, $p < .001$), specifically lower rates of bronchopulmonary dysplasia (44% vs. 55%, $p < .01$) and serious brain injury (6.2% vs. 17%, $p < .001$). Included children also had younger GA (39.6 vs. 41.1 weeks, $p < .001$) and lower weight (2907 vs. 3070 grams, $p < .02$) at NICU discharge, probably due to shorter NICU length of stay (87 vs. 98 days, $p < .001$). All comparisons of included and excluded children are shown in **Supplemental Table 1**.

Participants in the analysis sample were 51% male and the average GA at birth was 27.13 weeks ($SD = 1.87$). Maternal race in this sample was distributed as follows: Asian (4.7%), Black/African American (18%), White (46%), multiracial (18%), and unknown/not reported (12%; **Table 2**). The majority of mothers in the sample (88%) had a high school degree. The average environmental adversity score was 0.20 ($SD = 0.16$, range = 0 to 0.72). The average omission error T-score in the sample was 58.6 ($SD = 12.3$, range = 41 to 90).

EWAS Findings

Based on our analyses, DNA methylation at 9 CpG sites were associated with attention problems among very preterm infants (**Table 3**). None of these CpG sites were identified as co-methylated regions (CMRs). Of these 9 CpG sites, there were 2 negative associations (lower DNAm levels at CpG were associated with more attention problems) and 7 positive associations (higher DNAm levels at CpG were associated with more attention problems). The magnitude of the coefficients ranged from 3.34 to 4.88, indicating that a difference in methylation from the 25th to the 75th percentile is associated with a roughly 3 to 4 point difference in attention problem T-scores. Of the 9 significant results, 6 of the CpG sites are located in specific genes while 3 of the CpG sites are intergenic (**Table 4**).

We found several relevant traits associated with the significant CpGs and annotated genes from the conducted EWAS (**Table 4**). Five CpGs (cg16623157, cg08333974, cg08483160, cg20863501, cg13709858) were associated with previous entries in the EWAS catalog, in relation to DNAm levels and diseases (e.g. COPD, type 2 diabetes) or protein levels (e.g. c-

reactive protein [CRP], pentatricopeptide repeat proteins [PGR3]). Three CpGs (cg25972938, cg08333974, cg13709858) were located in genes that have been associated with neurodevelopmental traits such as reaction time, educational attainment and mathematical ability (*SLC12A4*, *CACNA2D4*, *OSBP2*) in prior GWAS.

Epigenetic Age Acceleration

In our covariate-adjusted models, we found no significant association between either PMA (b = -0.48, SE = 0.80, p = .55) or PNA acceleration (b = -0.40, SE = 0.57, p = .49) and attention problems.

Joint Environmental and Epigenetic Models

We then estimated a series of models to examine the joint association between epigenetics, environmental adversity, and child attention problems. To do so, we first estimated an epigenetics-only model that included all 9 significant CpG sites found to be individually significant from the EWAS (**Table 5**, Model 1). Compared to the EWAS models that tested each CpG individually, when modeled together, 5 of the 9 remained significant (cg13709858, cg08483160, cg08333974, cg25972938, cg16623157) though the magnitude of the coefficients was slightly reduced (from 3.43-4.88 to 2.11-2.66; **Table 5**, Model 2). We then estimated a model that included all 9 significant CpG sites from the EWAS along with environmental adversity (**Table 5**, Model 3). In this model, 4 of the 5 CpGs from Model 2 remained significant (cg13709858, cg08483160, cg08333974, cg25972938). In addition, there was a significant association between environmental adversity and child attention problems (b =9.55, SE = 4.63, $p < .05$) indicating that children exposed to higher levels of adversity tended to have higher attention problem scores. Thus, in our final joint model, DNAm and environmental adversity were uniquely associated with child attention problems.

Discussion

The purpose of this study was to identify epigenetic predictors of early childhood attention problems among very preterm children and to examine the joint role of epigenetic and environmental factors in predicting attention problems in this population. Using EWAS methods, we found 9 CpGs that were significantly associated with attention problems from ages 5-7 among

very preterm children, 7 of which were positively associated and 2 of which were negatively associated. Effect sizes for these associations ranged from 0.27 to 0.40 indicating small to medium sized effects. Conversely, we found no significant association between age acceleration (either PMA or PNA) and early childhood attention problems. Finally, we found a significant association between environmental adversity and early childhood attention problems, showing that very preterm infants who encountered more adversity within their home environment also tended to have more attention problems in early childhood. Examining the cumulative effects of epigenetic predictors and environmental adversity revealed that both factors uniquely contributed to risk for attention problems in this population.

From our EWAS, 3 of the 9 significant CpGs (*cg25972938, cg08333974, cg13709858*) are located in genes (*SLC12A4, CACNA2D4, OSBP2*, respectively) that have been shown in prior GWAS to have associations with neurodevelopmental traits such as mathematical ability, intelligence, educational attainment, and reaction time (Davies et al., 2018; Lee et al., 2018; Okbay et al., 2022; Savage et al., 2018). Attention problems have been shown to be predictive of later educational attainment and mathematical ability, showing that epigenetic regulation of these particular genes may provide information about long-term neurodevelopmental consequences for children born preterm (Loe & Feldman, 2007). The CpGs we identified in our EWAS have also been shown in prior EWAS to be associated with certain diseases (e.g. COPD, type 2 diabetes) and protein levels (e.g. CRP; Gadd et al., 2022; Hillary et al., 2023). Interestingly, CRP levels are thought to be a marker for systemic inflammation (Saccaro et al., 2021) and inflammation and inflammatory diseases have been increasingly researched due to their co-occurrence with ADHD (Leffa et al., 2018). Given that *cg13709858* was identified as a significant CpG in our EWAS and has also been shown to correlate with CRP, it suggests the importance of studying inflammation as a precursor to neurodevelopmental outcomes in preterm infants (e.g., Kuban et al., 2017; O'Shea et al., 2014). Moreover, although prior EWAS have been conducted within the NOVI cohort investigating DNA methylation and child neurobehavior, none of the significant CpGs from this analysis overlapped with any of the prior studies (Camerota, Lester, McGowan, et al., 2024; Everson et al., 2019) indicating that objectively assessed attentional problems may have unique biological underpinnings compared to other neurobehavioral measures. It is particularly noteworthy that we found no overlap in the results of our EWAS versus those in a recently conducted EWAS with parent-reported attention

problems in NOVI (Camerota, Lester, Castellanos, et al., 2024). These differences may be due to these distinct measures used to determine outcomes of attention problems (parent report versus objective assessment) or the different timepoints of assessment (age 2 years versus 5-7 years).

In contrast to findings from our EWAS, there were no significant associations found between any of our epigenetic age acceleration measures and early childhood attention problems. These null findings are consistent with a recent study looking at epigenetic age acceleration and ADHD symptoms among term-born children (Salontaji et al., 2024). Thus, it remains unclear if epigenetic age acceleration is specifically associated with attention deficits during childhood. On the other hand, the PedBE clock has been studied with a relatively small sample of very preterm infants and results showed that accelerated biological age (PMA) was associated with worse cognitive and language scores at 18 months (Gomaa et al., 2022). Although this prior study included neurodevelopmental outcomes assessed at a much younger age in comparison to our study, these findings still demonstrate that epigenetic aging among this population of very preterm infants may serve as an indicator of neurodevelopmental outcomes. It is possible that age acceleration at birth is informative of earlier emerging, foundational cognitive abilities, or that age acceleration measured closer to the assessment of attention problems would be a stronger predictor as opposed to age acceleration at birth. These remain important avenues for future research.

When analyzing the association between environmental adversity and child attention problems, our findings showed greater environmental adversity predicted worse attention problems among children born preterm. Previous studies show that exposure to specific environmental exposures (e.g., pesticides; air pollutants) as well as socioeconomic conditions are associated with ADHD risk in children (Saez et al., 2018; Spencer et al., 2022). These prior studies have primarily focused on full-term children, while our findings focus on a very preterm cohort, thereby supplementing the research conducted in this field by expanding it to a high-risk population. In addition to assessing the impact of environmental factors alone, we also analyzed the cumulative effects of environmental adversity and epigenetic factors on attention problems. When examined together, our findings show that neither the association of DNAm or environmental adversity with attention problems notably declined when the two factors were modeled together. Typically, epigenetics has been thought of as a

mechanism that might explain the relationship between environmental adversity and phenotypic outcomes, as environmental risk factors can influence patterns of DNAm (Hamza et al., 2019). However, our findings demonstrate the opposite; neonatal epigenetics and environmental adversity appear to work independently of each other in predicting risk for childhood attention problems. Thus, it does not appear that the impact of environmental adversity on attention problems is due to DNAm, nor is the association of DNAm with attention problems confounded by environmental adversity.

Strengths of our study include our use of a well-characterized, prospective cohort of children born very preterm, a vulnerable population at increased risk for attention problems. Despite being at increased risk, very preterm infants remain an understudied population (McCormick & Litt, 2017). The findings of our study help bring awareness to the importance of studying epigenetics and environmental adversity in this particular group, as these two factors may allow us to better predict (and possibly be better prepared) for possible adverse outcomes for these children. Our study also utilizes objective measurement of attention problems. Parent-report and objective assessments have different strengths and weaknesses; our use of K-CPT2 scores means our assessment of attention problems is not prone to biases that may be present in parental reports (e.g., recall bias, desirability bias). It is also necessary to acknowledge our results in the context of the study's limitations. While the NOVI cohort itself is relatively large, the number of children with epigenetic data and attention scores in early childhood was comparatively smaller. Thus, we may have been underpowered to detect smaller effects in this study. There were also some differences noted in the children included versus excluded, indicating that children included in this analysis had fewer neonatal morbidities compared to those excluded. Thus, our results may not generalize to all very preterm infants. Additionally, we utilized a cumulative environmental adversity index to assess the relationship between individuals' attention problems and their environment. Although using a cumulative index is a widely used approach across studies (Fox, 2002), the use of a collective index limits our ability to pinpoint the exact aspects of adverse childhood experiences that are most strongly associated with attention problems. This would be an important area of future study. Moreover, DNAm levels in our study are based on buccal swab data and thus our results do not necessarily convey information about processes occurring in the brain. Lastly, while our study examined attention problems and ADHD-related symptoms, we lack genetic data on this cohort

and therefore cannot account for genetic underpinnings of ADHD among this cohort of very preterm infants. While prior studies have indicated that preterm children are at higher risk for ADHD due to increased immaturity at birth, and that genetic factors do not explain this increased vulnerability, it remains important to consider underlying genetic factors that may contribute to ADHD risk in this population (Lindström et al., 2011).

Conclusion

In summary, we found that DNAm levels at specific CpG sites are associated with attention problems at ages 5 to 7 years among children born very preterm. By examining both environmental and epigenetic factors in a joint model, we concluded that these factors likely operate independently in the prediction of attention problems, rather than as confounders or mediators of one another. Future studies should investigate if these same CpGs continue to predict attention problems during the later elementary school years or if they predict ADHD diagnosis. The information from this study could potentially be used to improve identification of risk for attention problems among the very preterm population, to pinpoint individual children who may benefit from earlier treatment or intervention.

References

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. *Pediatrics*, *124*(2), Article 2. https://doi.org/10.1542/peds.2008-2816

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*(10), 1363–1369. https://doi.org/10.1093/bioinformatics/btu049

Barker, E. D., Jaffee, S. R., Uher, R., & Maughan, B. (2011). The contribution of prenatal and postnatal maternal anxiety and depression to child maladjustment. *Depression and Anxiety*, *28*(8), 696–702. https://doi.org/10.1002/da.20856

Bassler, D., Stoll, B. J., Schmidt, B., Asztalos, E. V., Roberts, R. S., Robertson, C. M. T., Sauve, R. S., & for the Trial of Indomethacin Prophylaxis in Preterms Investigators. (2009). Using a Count of Neonatal Morbidities to Predict Poor Outcome in Extremely Low Birth Weight Infants: Added Role of Neonatal Infection. *Pediatrics*, *123*(1), 313–318. https://doi.org/10.1542/peds.2008-0377

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Camerota, M., Lester, B. M., Castellanos, F. X., Carter, B. S., Check, J., Helderman, J., Hofheimer, J. A., McGowan, E. C., Neal, C. R., Pastyrnak, S. L., Smith, L. M., O'Shea, T. M., Marsit, C. J., & Everson, T. M. (2024). Epigenome-wide association study identifies neonatal DNA methylation associated with two-year attention problems in children born very preterm. *Translational Psychiatry*, *14*(1), 126.

https://doi.org/10.1038/s41398-024-02841-y

Camerota, M., Lester, B. M., McGowan, E. C., Carter, B. S., Check, J., Dansereau, L. M., DellaGrotta, S. A., Helderman, J. B., Hofheimer, J. A., Loncar, C. M., Neal, C. R., O'Shea, T. M., Pastyrnak, S. L., Smith, L. M., Abrishamcar, S., Hüls, A., Marsit, C. J., & Everson, T. M. (2024). Contributions of prenatal risk factors and neonatal epigenetics to cognitive outcome in children born very preterm. *Developmental Psychology*. https://doi.org/10.1037/dev0001709

Campagna, M. P., Xavier, A., Lechner-Scott, J., Maltby, V., Scott, R. J., Butzkueven, H., Jokubaitis, V. G., & Lea, R. A. (2021). Epigenome-wide association studies: Current knowledge, strategies and recommendations. *Clinical Epigenetics*, *13*(1), 214. https://doi.org/10.1186/s13148-021-01200-8

Cecil, C. A. M., & Nigg, J. T. (2022). Epigenetics and ADHD: Reflections on Current Knowledge, Research Priorities and Translational Potential. *Molecular Diagnosis & Therapy*, *26*(6), 581–606. https://doi.org/10.1007/s40291-022-00609-y

Chen, I.-C., Chang, C.-H., Chang, Y., Lin, D.-S., Lin, C.-H., & Ko, L.-W. (2021). Neural Dynamics for Facilitating ADHD Diagnosis in Preschoolers: Central and Parietal Delta Synchronization in the Kiddie Continuous Performance Test. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *29*, 1524–1533. https://doi.org/10.1109/TNSRE.2021.3097551

Chen, Y.-C., Sudre, G., Sharp, W., Donovan, F., Chandrasekharappa, S. C., Hansen, N., Elnitski, L., & Shaw, P. (2018). Neuroanatomic, epigenetic and genetic differences in monozygotic twins discordant for attention deficit hyperactivity disorder. *Molecular Psychiatry*, *23*(3), 683–690. https://doi.org/10.1038/mp.2017.45

Davies, G., Lam, M., Harris, S. E., Trampush, J. W., Luciano, M., Hill, W. D., Hagenaars, S. P.,

Ritchie, S. J., Marioni, R. E., Fawns-Ritchie, C., Liewald, D. C. M., Okely, J. A., Ahola-Olli, A. V., Barnes, C. L. K., Bertram, L., Bis, J. C., Burdick, K. E., Christoforou, A., DeRosse, P., … Deary, I. J. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nature Communications*, *9*(1), 2098. https://doi.org/10.1038/s41467-018-04362-x

Davis, E. G., Humphreys, K. L., McEwen, L. M., Sacchet, M. D., Camacho, M. C., MacIsaac, J.
L., Lin, D. T. S., Kobor, M. S., & Gotlib, I. H. (2017). Accelerated DNA methylation age in adolescent girls: Associations with elevated diurnal cortisol and reduced hippocampal volume. *Translational Psychiatry*, *7*(8), e1223–e1223. https://doi.org/10.1038/tp.2017.188

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory
report. *Psychological Medicine*, *13*(3), 595–605. https://doi.org/10.1017/S0033291700048017

Downey, L. C., O'Shea, T. M., Allred, E. N., Kuban, K., McElrath, T. F., Warner, D. D., ... &
Study, E. L. G. A. N. E. (2015). Antenatal and early postnatal antecedents of parent-reported attention problems at 2 years of age. *The Journal of pediatrics*, *166*(1), 20-25. https://doi.org/10.1016/j.jpeds.2014.08.004

Everson, T. M., Marsit, C. J., Michael O'Shea, T., Burt, A., Hermetz, K., Carter, B. S.,
Helderman, J., Hofheimer, J. A., McGowan, E. C., Neal, C. R., Pastyrnak, S. L., Smith, L. M., Soliman, A., DellaGrotta, S. A., Dansereau, L. M., Padbury, J. F., & Lester, B. M. (2019). Epigenome-wide Analysis Identifies Genes and Pathways Linked to Neurobehavioral Variation in Preterm Infants. *Scientific Reports*, *9*(1), 6322. https://doi.org/10.1038/s41598-019-42654-4

Everson, T. M., O'Shea, T. M., Burt, A., Hermetz, K., Carter, B. S., Helderman, J., Hofheimer, J.
A., McGowan, E. C., Neal, C. R., Pastyrnak, S. L., Smith, L. M., Soliman, A., DellaGrotta, S. A., Dansereau, L. M., Padbury, J. F., Lester, B. M., & Marsit, C. J. (2020). Serious neonatal morbidities are

associated with differences in DNA methylation among very preterm infants. *Clinical Epigenetics*, *12*(1), 151. https://doi.org/10.1186/s13148-020-00942-1

Faraone, S. V., & Mick, E. (2010). Molecular Genetics of Attention Deficit Hyperactivity Disorder. *Psychiatric Clinics of North America*, *33*(1), 159–180. https://doi.org/10.1016/j.psc.2009.12.004

Felt, J. M., Yusupov, N., Harrington, K. D., Fietz, J., Zhang, Z. “Zach,” Sliwinski, M. J., Ram, N., O’Donnell, K. J., Meaney, M. J., Putnam, F. W., Noll, J. G., Binder, E. B., & Shenk, C. E. (2023). Epigenetic age acceleration as a biomarker for impaired cognitive abilities in adulthood following early life adversity and psychiatric disorders. *Neurobiology of Stress*, *27*, 100577. https://doi.org/10.1016/j.ynstr.2023.100577

Fox, M. A. (2002). Evaluating cumulative risk assessment for environmental justice: A community case study. *Environmental Health Perspectives*, *110*(suppl 2), 203–209. https://doi.org/10.1289/ehp.02110s2203

Fransquet, P. D., Wrigglesworth, J., Woods, R. L., Ernst, M. E., & Ryan, J. (2019). The epigenetic clock as a predictor of disease and mortality risk: A systematic review and meta-analysis. *Clinical Epigenetics*, *11*(1), 62. https://doi.org/10.1186/s13148-019-0656-7

Franz, A. P., Bolat, G. U., Bolat, H., Matijasevich, A., Santos, I. S., Silveira, R. C., Procianoy, R. S., Rohde, L. A., & Moreira-Maia, C. R. (2018). Attention-Deficit/Hyperactivity Disorder and Very Preterm/Very Low Birth Weight: A Meta-analysis. *Pediatrics*, *141*(1), e20171645. https://doi.org/10.1542/peds.2017-1645

Gadd, D. A., Hillary, R. F., McCartney, D. L., Zaghlool, S. B., Stevenson, A. J., Cheng, Y., Fawns-Ritchie, C., Nangle, C., Campbell, A., Flaig, R., Harris, S. E., Walker, R. M., Shi, L., Tucker-Drob, E. M., Gieger, C., Peters, A., Waldenberger, M., Graumann, J., McRae, A. F., … Marioni, R. E.

(2022). Epigenetic scores for the circulating proteome as tools for disease prediction. *eLife*, *11*, e71802. https://doi.org/10.7554/eLife.71802

Gatev, E., Gladish, N., Mostafavi, S., & Kobor, M. S. (2020). CoMeBack: DNA methylation array data analysis for co-methylated regions. *Bioinformatics (Oxford, England)*, *36*(9), 2675–2683. https://doi.org/10.1093/bioinformatics/btaa049

Gomaa, N., Konwar, C., Gladish, N., Au-Young, S. H., Guo, T., Sheng, M., Merrill, S. M., Kelly,

E., Chau, V., Branson, H. M., Ly, L. G., Duerden, E. G., Grunau, R. E., Kobor, M. S., & Miller, S. P. (2022). Association of Pediatric Buccal Epigenetic Age Acceleration With Adverse Neonatal Brain Growth and Neurodevelopmental Outcomes Among Children Born Very Preterm With a Neonatal Infection. *JAMA Network Open*, *5*(11), e2239796. https://doi.org/10.1001/jamanetworkopen.2022.39796

Goodman, S. J., Burton, C. L., Butcher, D. T., Siu, M. T., Lemire, M., Chater-Diehl, E., Turinsky,

A. L., Brudno, M., Soreni, N., Rosenberg, D., Fitzgerald, K. D., Hanna, G. L., Anagnostou, E., Arnold, P. D., Crosbie, J., Schachar, R., & Weksberg, R. (2020). Obsessive-compulsive disorder and attention-deficit/hyperactivity disorder: Distinct associations with DNA methylation and genetic variation. *Journal of Neurodevelopmental Disorders*, *12*(1), 23. https://doi.org/10.1186/s11689-020-09324-3

Graw, S., Camerota, M., Carter, B. S., Helderman, J., Hofheimer, J. A., McGowan, E. C., Neal,

C. R., Pastyrnak, S. L., Smith, L. M., DellaGrotta, S. A., Dansereau, L. M., Padbury, J. F., O'Shea, M., Lester, B. M., Marsit, C. J., & Everson, T. M. (2021). NEOage clocks—Epigenetic clocks to estimate post-menstrual and postnatal age in preterm infants. *Aging*, *13*(20), 23527–23544. https://doi.org/10.18632/aging.203637

Hamza, M., Halayem, S., Bourgou, S., Daoud, M., Charfi, F., & Belhadj, A. (2019). Epigenetics and ADHD: Toward an Integrative Approach of the Disorder Pathogenesis. *Journal of Attention Disorders*, *23*(7), 655–664. https://doi.org/10.1177/1087054717696769

Hillary, R. F., McCartney, D. L., Bernabeu, E., Gadd, D. A., Cheng, Y.,

Chybowska, A. D.,
Smith, H. M., Murphy, L., Wrobel, N., Campbell, A., Walker, R. M., Hayward, C., Evans, K. L., McIntosh, A. M., & Marioni, R. E. (2023). *Blood-based epigenome-wide analyses on the prevalence and incidence of nineteen common disease states*. https://doi.org/10.1101/2023.01.10.23284387

Johnson, S., & Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatric*
Research, *69*(5 Pt 2), 11R-8R.
https://doi.org/10.1203/pdr.0b013e318212faa0

Lanata, C. M., Chung, S. A., & Criswell, L. A. (2018). DNA methylation 101: What is important to know about DNA methylation and its role in SLE risk and disease heterogeneity. *Lupus Science & Medicine*, *5*(1), e000285. https://doi.org/10.1136/lupus-2018-000285

Kuban, K. C. K., Joseph, R. M., O'Shea, T. M., Heeren, T., Fichorova, R. N., Douglass, L., Jara, H., Frazier, J. A., Hirtz, D., Rollins, J. V., & Paneth, N. (2017). Circulating Inflammatory-Associated Proteins in the First Month of Life and Cognitive Impairment at Age 10 Years in Children Born Extremely Preterm. *The Journal of Pediatrics*, *180*, 116-123.e1. https://doi.org/10.1016/j.jpeds.2016.09.054

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A.,
Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H., Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Leffa, D. T., Torres, I. L. S., & Rohde, L. A. (2018). A Review on the Role of Inflammation in
Attention-Deficit/Hyperactivity Disorder. *Neuroimmunomodulation*, *25*(5–6), 328–333. https://doi.org/10.1159/000489635

Lindström, K., Lindblad, F., & Hjern, A. (2011). Preterm Birth and Attention-Deficit/Hyperactivity Disorder in Schoolchildren. *Pediatrics*, *127*(5), 858–865. https://doi.org/10.1542/peds.2010-1279

Liu, J., & Siegmund, K. D. (2016). An evaluation of processing methods for HumanMethylation450 BeadChip data. *BMC Genomics*, *17*(1), 469. https://doi.org/10.1186/s12864-016-2819-7

Loe, I. M., & Feldman, H. M. (2007). Academic and Educational Outcomes of Children With ADHD. *Journal of Pediatric Psychology*, *32*(6), 643–654. https://doi.org/10.1093/jpepsy/jsl054

Logue, M. W., Smith, A. K., Wolf, E. J., Maniates, H., Stone, A., Schichman, S. A., McGlinchey, R. E., Milberg, W., & Miller, M. W. (2017). The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples. *Epigenomics*, *9*(11), Article 11. https://doi.org/10.2217/epi-2017-0078

McCormick, M. C., & Litt, J. S. (2017). The Outcomes of Very Preterm Infants: Is It Time to Ask Different Questions? *Pediatrics*, *139*(1), e20161694. https://doi.org/10.1542/peds.2016-1694

McEwen, L. M., O'Donnell, K. J., McGill, M. G., Edgar, R. D., Jones, M. J., MacIsaac, J. L., Lin, D. T. S., Ramadori, K., Morin, A., Gladish, N., Garg, E., Unternaehrer, E., Pokhvisneva, I., Karnani, N., Kee, M. Z. L., Klengel, T., Adler, N. E., Barr, R. G., Letourneau, N., … Kobor, M. S. (2020). The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proceedings of the National Academy of Sciences*, *117*(38), 23329–23335. https://doi.org/10.1073/pnas.1820843116

McGowan, E. C., Hofheimer, J. A., O'Shea, T. M., Kilbride, H., Carter, B. S., Check, J., Helderman, J., Neal, C. R., Pastyrnak, S., Smith, L. M., Camerota, M., Dansereau, L. M., Della Grotta, S. A., & Lester, B. M. (2022). Analysis of Neonatal Neurobehavior and Developmental Outcomes Among Preterm Infants. *JAMA Network Open*, *5*(7), e2222249. https://doi.org/10.1001/jamanetworkopen.2022.22249

Mooney, M. A., Ryabinin, P., Wilmot, B., Bhatt, P., Mill, J., & Nigg, J. T.

(2020). Large
epigenome-wide association study of childhood ADHD identifies peripheral DNA methylation associated with disease and polygenic risk burden. *Translational Psychiatry*, *10*(1), 8. https://doi.org/10.1038/s41398-020-0710-4

Moore, L. D., Le, T., & Fan, G. (2013). DNA Methylation and Its Basic Function.
Neuropsychopharmacology, *38*(1), 23–38. https://doi.org/10.1038/npp.2012.112

Neumann, A., Walton, E., Alemany, S., Cecil, C., González, J. R., Jima, D. D., Lahti, J.,
Tuominen, S. T., Barker, E. D., Binder, E., Caramaschi, D., Carracedo, Á., Czamara, D., Evandt, J., Felix, J. F., Fuemmeler, B. F., Gutzkow, K. B., Hoyo, C., Julvez, J., … Tiemeier, H. (2020). Association between DNA methylation and ADHD symptoms from birth to school age: A prospective meta-analysis. *Translational Psychiatry*, *10*(1), 398. https://doi.org/10.1038/s41398-020-01058-z

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J.,
Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., … Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437–449. https://doi.org/10.1038/s41588-022-01016-z

O'Shea, T. M., Joseph, R. M., Kuban, K. C. K., Allred, E. N., Ware, J., Coster, T., Fichorova, R. N., Dammann, O., & Leviton, A. (2014). Elevated blood levels of inflammation-related proteins are associated with an attention problem at age 24 mo in extremely preterm infants. *Pediatric Research*, *75*(6), 781–787. https://doi.org/10.1038/pr.2014.41

O'Shea, T. M., McGrath, M., Aschner, J. L., Lester, B., Santos, H. P., Marsit, C., Stroustrup, A.,
Emmanuel, C., Hudak, M., McGowan, E., Patel, S., Fry, R. C., Smith, P. B., Newby, K. L., Jacobson, L. P., Parker, C. B., & on behalf of

program collaborators for Environmental influences on Child Health Outcomes. (2022). Environmental influences on child health outcomes: Cohorts of individuals born very preterm. *Pediatric Research*. https://doi.org/10.1038/s41390-022-02230-5

Palladino, V. S., McNeill, R., Reif, A., & Kittel-Schneider, S. (2019). Genetic risk factors and gene–environment interactions in adult and childhood attention-deficit/hyperactivity disorder. *Psychiatric Genetics*, *29*(3), 63–78. https://doi.org/10.1097/YPG.0000000000000220

Paniagua, U., Lester, B. M., Marsit, C. J., Camerota, M., Carter, B. S., Check, J. F., Helderman, J., Hofheimer, J. A., McGowan, E. C., Neal, C. R., Pastyrnak, S. L., Smith, L. M., DellaGrotta, S. A., Dansereau, L. M., O'Shea, T. M., & Everson, T. M. (2023). Epigenetic age acceleration, neonatal morbidities, and neurobehavioral profiles in infants born very preterm. *Epigenetics*, *18*(1), 2280738. https://doi.org/10.1080/15592294.2023.2280738

Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, *14*(1), 293. https://doi.org/10.1186/1471-2164-14-293

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C., & Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, *17*(1), Article 1. https://doi.org/10.1186/s13059-016-1066-1

Pingault, J.-B., Côté, S. M., Vitaro, F., Falissard, B., Genolini, C., & Tremblay, R. E. (2014). The developmental course of childhood inattention symptoms uniquely predicts educational attainment: A 16-year longitudinal study. *Psychiatry Research*, *219*(3), 707–709. https://doi.org/10.1016/j.psychres.2014.06.022

Saccaro, L. F., Schilliger, Z., Perroud, N., & Piguet, C. (2021). Inflammation, Anxiety, and Stress in Attention-Deficit/Hyperactivity Disorder. *Biomedicines*, *9*(10), 1313. https://doi.org/10.3390/biomedicines9101313

Saez, M., Barceló, M. A., Farrerons, M., & López-Casasnovas, G. (2018). The association between exposure to environmental factors and the occurrence of attention-deficit/hyperactivity disorder (ADHD). A population-based retrospective cohort study. *Environmental Research*, *166*, 205–214. https://doi.org/10.1016/j.envres.2018.05.009

Salontaji, K., Haftorn, K. L., Sanders, F., Page, C. M., Walton, E., Felix, J. F., Bekkhus, M., Bohlin, J., Tiemeier, H., & Cecil, C. A. M. (2024). Gestational epigenetic age and ADHD symptoms in childhood: A prospective, multi-cohort study. *Molecular Psychiatry*, s41380-024-02544–2. https://doi.org/10.1038/s41380-024-02544-2

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C. A., Nagel, M., Awasthi, S., Barr, P. B., Coleman, J. R. I., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., … Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, *50*(7), 912–919. https://doi.org/10.1038/s41588-018-0152-6

Shaked, D., Faulkner, L. M. D., Tolle, K., Wendell, C. R., Waldstein, S. R., & Spencer, R. J. (2020). Reliability and validity of the Conners' Continuous Performance Test. *Applied Neuropsychology: Adult*, *27*(5), 478–487. https://doi.org/10.1080/23279095.2019.1570199

Shuster, C. L., Sheinkopf, S. J., McGowan, E. C., Hofheimer, J. A., O'Shea, T. M., Carter, B. S., Helderman, J. B., Check, J., Neal, C. R., Pastyrnak, S. L., Smith, L. M., Loncar, C., Dansereau, L. M., DellaGrotta, S. A., Marsit, C., & Lester, B. M. (2023). Neurobehavioral and Medical Correlates of Autism Screening: 2-Year Outcomes for Infants Born Very Preterm. *The*

Journal of Pediatrics, 113536. https://doi.org/10.1016/j.jpeds.2023.113536

Spencer, N. J., Ludvigsson, J., Bai, G., Gauvin, L., Clifford, S. A., Abu Awad, Y., Goldhaber-Fiebert, J. D., Markham, W., Faresjö, Å., White, P. A., Raat, H., Jansen, P., Nikiema, B., Mensah, F. K., McGrath, J. J., & EPOCH Collaborative Group. (2022). Social gradients in ADHD by household income and maternal education exposure during early childhood: Findings from birth cohort studies across six countries. *PLOS ONE*, *17*(3), e0264709. https://doi.org/10.1371/journal.pone.0264709

Tarca, A. L., Romero, R., & Draghici, S. (2006). Analysis of microarray experiments of gene expression profiling. *American Journal of Obstetrics and Gynecology*, *195*(2), 373–388. https://doi.org/10.1016/j.ajog.2006.07.001

Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., & Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics (Oxford, England)*, *29*(2), Article 2. https://doi.org/10.1093/bioinformatics/bts680

Turecki, G., & Meaney, M. J. (2016). Effects of the Social Environment and Stress on Glucocorticoid Receptor Gene Methylation: A Systematic Review. *Biological Psychiatry*, *79*(2), 87–96. https://doi.org/10.1016/j.biopsych.2014.11.022

Walton, E., Pingault, J.-B., Cecil, C. A. M., Gaunt, T. R., Relton, C. L., Mill, J., & Barker, E. D. (2017). Epigenetic profiling of ADHD symptoms trajectories: A prospective, methylome-wide study. *Molecular Psychiatry*, *22*(2), 250–256. https://doi.org/10.1038/mp.2016.85

Wang, Y., Qian, M., Tang, D., Herbstman, J., Perera, F., & Wang, S. (2020). A powerful and flexible weighted distance-based method incorporating interactions between DNA methylation and environmental factors on health

outcomes. *Bioinformatics*, *36*(3), 653–659. https://doi.org/10.1093/bioinformatics/btz630

Wei, S., Tao, J., Xu, J., Chen, X., Wang, Z., Zhang, N., Zuo, L., Jia, Z., Chen, H., Sun, H., Yan,
Y., Zhang, M., Lv, H., Kong, F., Duan, L., Ma, Y., Liao, M., Xu, L., Feng, R., … Jiang, Y. (2021). Ten Years of EWAS. *Advanced Science*, *8*(20), 2100727. https://doi.org/10.1002/advs.202100727

Wilmot, B., Fry, R., Smeester, L., Musser, E. D., Mill, J., & Nigg, J. T. (2016). Methylomic
analysis of salivary DNA in childhood ADHD identifies altered DNA methylation in *VIPR2*. *Journal of Child Psychology and Psychiatry*, *57*(2), 152–160. https://doi.org/10.1111/jcpp.12457

Zheng, S. C., Webster, A. P., Dong, D., Feber, A., Graham, D. G., Sullivan, R., Jevons, S., Lovat,
L. B., Beck, S., Widschwendter, M., & Teschendorff, A. E. (2018). A novel cell-type deconvolution algorithm reveals substantial contamination by immune cells in saliva, buccal and cervix. *Epigenomics*, *10*(7), Article 7. https://doi.org/10.2217/epi-2018-0037

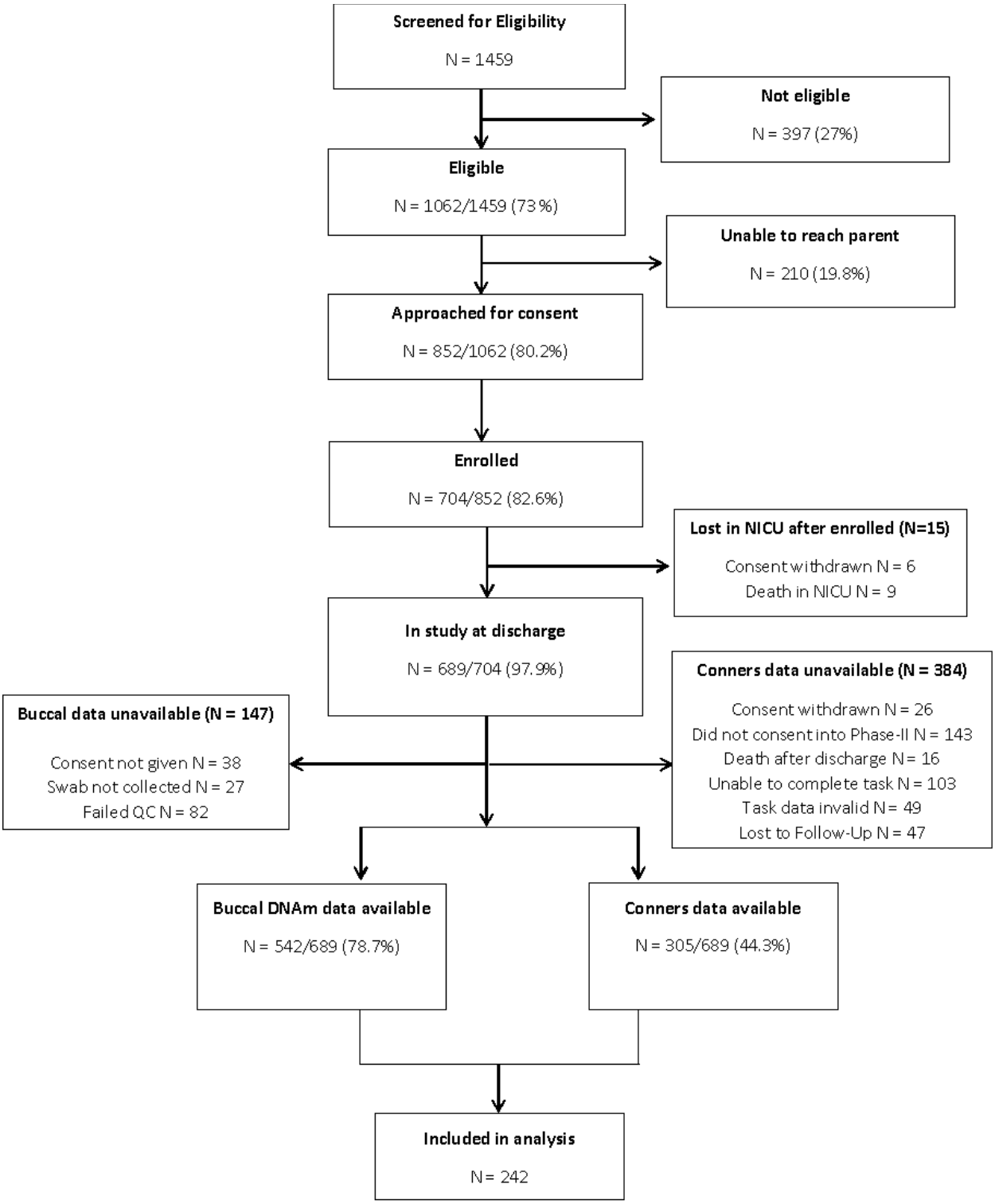

Figure 1. Study flowchart.

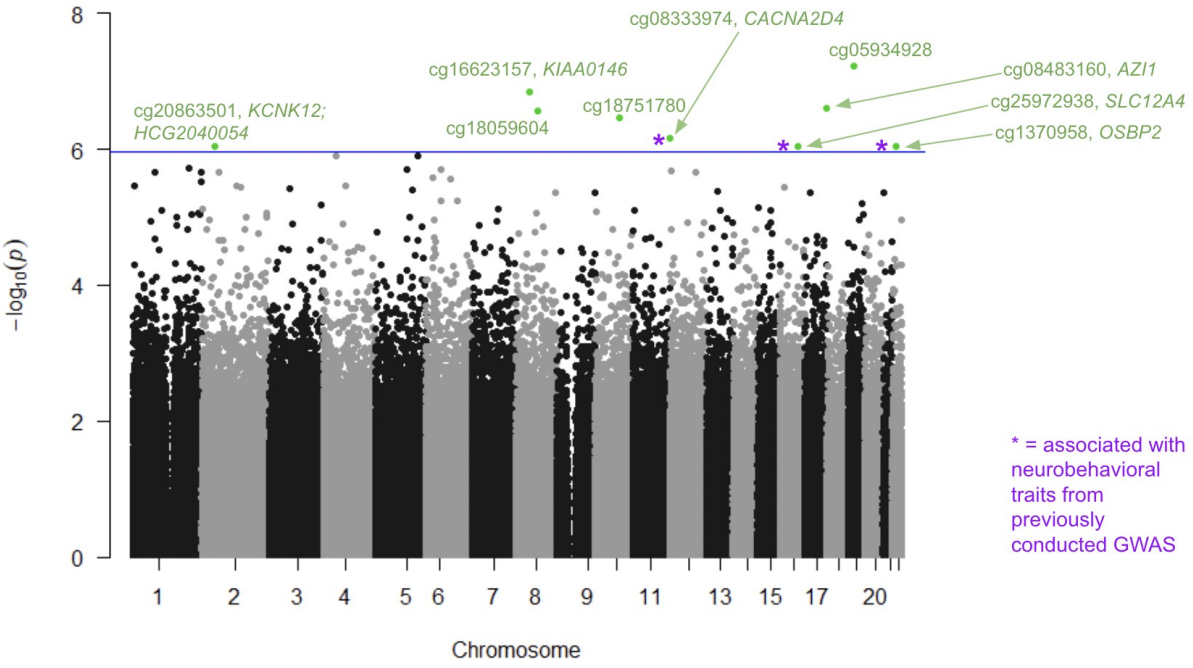

Figure 2. Manhattan plot showing CpG sites significantly associated with child inattention scores (FDR < 5%). The x-axis indicates genomic location and the y-axis indicates the -log10(p values) from the EWAS models examining CpG methylation in relation to Conner's KCPT omission error T-scores, controlling for cellular heterogeneity, sample batch, child sex, recruitment site, gestational age at birth, gestational age at sample collection, neonatal medical morbidities, and maternal prenatal smoking. Gene annotations are added for all CpGs yielding significant associations after FDR adjustment. CpGs without a gene annotation were intergenic. An asterisk is used to denote CpGs located in genes that have previously been associated with neurobehavioral traits (e.g., educational attainment, mathematical ability) in the GWAS database.

Table 1
Prevalence of postnatal risk factors at each timepoint

Risk Factor	**Defining Characteristics**	**24 months**	**36 months**	**48 months**
Limited socioeconomic resources	Household income < 150% of the poverty line and/or Hollingshead level V	41%	40%	39%
Limited childcare assistance	No reported secondary caregiver	10%	15%	9%
Caregiver psychiatric disorder	Primary caregiver scoring above the clinical cutoff for depression and/or anxiety (T-score ≥ 63) on the Brief Symptom Inventory, in the mild to severe depression range (total score ≥ 6) on the Quick Inventory of Depressive Symptomatology, or self-report of any diagnosis, treatment, or counseling for depression and/or anxiety	42%	38%	37%
Family disruption/separation	Any reported Child Protective Services involvement	12%	3%	1%
Stressful life events	The sum of stressful life events >1SD above the sample mean	15%	7%	7%
Crowded household	Four or more children reported to be living in the household	13%	13%	15%

Table 2
Demographic and medical characteristics of the sample

Maternal characteristics (N = 212)	**M (SD) or % (n)**
Minority race or ethnicity	55% (117/211)
American Indian / Alaska Native race	0% (0/212)
Asian race	4.7% (10/212)
Native Hawaiian/Other Pacific Islander race	0.94% (2/212)
Black or African American race	18% (38/212)
White race	46% (98/212)
More than one race	18% (38/212)
Unknown/Race not reported	12% (26/212)
Hispanic/Latinx ethnicity	25% (53/212)
Low SES: Hollingshead level 5	7.6% (16/211)
Maternal education: < HS/GED	12% (25/210)
No partner	21% (44/211)
Pre-pregnancy obesity	40% (82/206)
Prenatal tobacco use	11% (24/211)
Neonatal characteristics (N = 242)	**M (SD) or % (n)**
Multiple gestation	31% (74/241)
Vaginal delivery	26% (63/241)
Neonatal medical morbidities (count)	0.7167 (0.778)
Severe retinopathy of prematurity	4.1% (10/241)
Necrotizing enterocolitis/sepsis	17% (42/241)
Bronchopulmonary dysplasia	44% (106/241)
Serious brain injury	6.2% (15/240)
Sex = Male	51% (123/242)
GA at birth (weeks)	27.13 (1.87)
Head circumference (cm)	24.51 (2.41)
GA at NICU discharge (weeks)	39.55 (3.97)
Length of NICU stay (days)	86.95 (33)
Birth weight (g)	957.7 (268)
Weight at discharge (g)	2907 (781)

Note. GA, gestational age; HS, high school; GED, General Equivalency Diploma; NICU, Neonatal intensive care unit; SES, socioeconomic status. Minority race or ethnicity was defined as any non-White race (e.g., Black,

Asian) or ethnicity (e.g., Hispanic and/or Latinx). Serious brain injury included parenchymal echodensity, periventricular leukomalacia, or ventricular dilation diagnosed via cranial ultrasound.

Table 3
EWAS results for statistically significant CpG sites (FDR <5%)

CpG	**Location**	**Gene Annotation**	**Coefficient**	**Std Error**	***p* value (raw)**	***p* value (FDR)**
cg16623157	chr8:48385636	KIAA0146	4.88	0.93	1.45E-07	0.0313
cg25972938	chr16:67995561	SLC12A4	-3.84	0.78	9.11E-07	0.0465
cg05934928	chr19:19866258	N/A (intergenic)	3.71	0.68	5.88E-08	0.0266
cg08333974	chr12:19563637	CACNA2D4	3.43	0.69	6.76E-07	0.0465
cg08483160	chr17:79183966	AZI1	4.28	0.83	2.47E-07	0.0313
cg20863501	chr2:47754578	HCG2040054; KCNK12	4.17	0.85	9.04E-07	0.0465
cg18751780	chr10:87318404	N/A (intergenic)	3.39	0.67	3.52E-07	0.0318
cg18059604	chr8:79312147	N/A (intergenic)	3.34	0.65	2.77E-07	0.0313
cg13709858	chr22:31226258	OSBP2	-3.75	0.76	9.25E-07	0.0465

Table 4
CpGs associated with child attention problems (FDR < 5%) are linked to genes and outcomes in the GWAS and EWAS Catalogs

CpG	Location	Gene Annotation	EWAS Catalog Hits	EWAS Catalog Entries	GWAS Catalog Hits	GWAS Catalog Entries
cg16623157	chr8:48385636	KIAA0146	3	Tissue, age	--	
cg25972938	chr16:67995561	SLC12A4	0		28	Mathematical ability, BMI
cg05934928	chr19:19866258	N/A (intergenic)	--		--	
cg08333974	chr12:19563337	CACNA2D4	2	Whole blood, incident COPD, incident chronic kidney disease	24	Educational attainment, mathematical ability, reaction time
cg08483160	chr17:79183966	AZI1	4	Age	--	
cg20863501	chr2:47754578	HCG2040054; KCNK12	1	Protein levels, implication of methylation	1	
cg18751780	chr10:87318404	N/A (intergenic)	--		--	
cg18059604	chr8:79312147	N/A (intergenic)	--		--	
cg13709858	chr22:31226258	OSBP2	4	Type 2 diabetes and methylation, CRP protein levels, PRG3 protein	51	Mathematical ability, educational attainment, intelligence

levels

	Model 1: Individual CpGs	**Model 2: All CpGs**	**Model 3: All CpGs + Environment**
DNAm (cg13709858)	-3.75 (0.76)***	-2.49 (0.81)**	-2.61(0.81)***
DNAm (cg08483160)	4.28 (0.83)***	2.32 (0.87)**	2.52 (0.89)**
DNAm (cg08333974)	3.43 (0.69)***	2.57 (0.74)***	2.28 (0.76)**
DNAm (cg05934928)	3.71 (0.68)***	0.21 (1.08)	0.07 (1.10)
DNAm (cg25972938)	-3.84 (0.78)***	-2.11 (0.81)**	-1.86 (0.92)*
DNAm (cg20863501)	4.17 (0.85)***	0.72 (1.04)	0.88 (1.03)
DNAm (cg18059604)	3.34 (0.65)***	0.30 (1.13)	0.70 (1.19)
DNAm (cg16623157)	4.88 (0.93)***	2.66 (1.09)*	1.96 (1.18)
DNAm (cg18751780)	3.39 (0.67)***	1.09 (1.18)	0.94 (1.23)
Environmental Adversity			9.55 (4.63)*

Table 5
Results from joint environmental and epigenetic models

Note. Models were estimated as generalized estimating equation (GEE) models accounting for covariates and nesting of children within families (i.e., multiple births). Model 1 included DNAm at each individual CpG as separate predictors (same as epigenome-wide association study models). Model 2 included all CpGs at the same time with all covariates. Model 3 included all CpGs with covariates as well as cumulative environmental

adversity. The outcome for all models was omission error T-scores, where higher scores indicated more attention problems.
$+p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Supplemental Table 1
Demographic and medical characteristics of included and excluded participants

	Included (N = 212)	**Excluded (N = 405)**	
Maternal characteristics	**M (SD) or % (n)**	**M (SD) or % (n)**	***p* value**
Minority race or ethnicity	55% (117/211)	58% (230/395)	0.57
American Indian / Alaska Native race	0% (0/212)	0.25% (1/405)	1.00
Asian race	4.7% (10/212)	3.2% (13/405)	0.48
Native Hawaiian/Other Pacific Islander race	0.94% (2/212)	1.5% (6/405)	0.85
Black or African American race	18% (38/212)	22% (88/405)	0.31
White race	46% (98/212)	40% (163/405)	0.18
More than one race	18% (38/212)	24% (98/405)	0.09
Unknown/Race not reported	12% (26/212)	8.9% (36/405)	0.24
Hispanic/Latinx ethnicity	25% (53/212)	22% (89/405)	0.46
Low SES: Hollingshead level 5	7.6% (16/211)	11% (42/394)	0.28
Maternal education: < HS/GED	12% (25/210)	13% (53/394)	0.68
No partner	21% (44/211)	27% (108/394)	0.09
Pre-pregnancy obesity	40% (82/206)	31% (120/386)	0.04
Prenatal tobacco use	11% (24/211)	17% (67/396)	0.09
	Included (N = 242)	**Excluded (N = 462)**	
Neonatal characteristics	**M (SD) or % (n)**	**M (SD) or % (n)**	***p* value**
Multiple gestation	31% (74/241)	24% (110/456)	0.07
Vaginal delivery	26% (63/241)	30% (138/455)	0.28
Neonatal medical morbidities (count)	0.7167 (0.778)	0.969 (0.927)	<0.001
Severe retinopathy of prematurity	4.1% (10/241)	6.8% (31/456)	0.21
Necrotizing enterocolitis/sepsis	17% (42/241)	19% (86/456)	0.72
Bronchopulmonary dysplasia	44% (106/241)	55% (251/456)	0.01
Serious brain injury	6.2% (15/240)	17% (77/454)	<0.001
Sex = Male	51%	58%	0.08

	(123/242)	(268/462)	
GA at birth (weeks)	27.13 (1.87)	26.94 (1.94)	0.21
Head circumference (cm)	24.51 (2.41)	24.44 (2.44)	0.73
GA at NICU discharge (weeks)	39.55 (3.97)	41.05 (6.01)	<0.001
Length of NICU stay (days)	86.95 (33)	97.99 (48.6)	<0.001
Birth weight (g)	957.7 (268)	943.3 (287)	0.51
Weight at discharge (g)	2907 (781)	3070 (960)	0.02

Note. GA, gestational age; HS, high school; GED, General Equivalency Diploma; NICU, Neonatal intensive care unit; SES, socioeconomic status. Minority race or ethnicity was defined as any non-White race (e.g., Black, Asian) or ethnicity (e.g., Hispanic and/or Latinx). Serious brain injury included parenchymal echodensity, periventricular leukomalacia, or ventricular dilation diagnosed via cranial ultrasound.