Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Claire A. Rowcliffe Date

In-utero exposure to indoor air pollution or tobacco smoke and cognitive development in a South African birth cohort study

By

Claire A. Rowcliffe

Master of Public Health

Epidemiology

_____ [Chair's signature]

Anke Hüls, PhD

Committee Chair

In-utero exposure to indoor air pollution or tobacco smoke and cognitive development in a South African birth cohort study

By

Claire Rowcliffe

Bachelor of Science

University of Tennessee – Knoxville

2018

Thesis Committee Chair: Anke Hüls, PhD

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2021

Abstract

In-utero exposure to indoor air pollution or tobacco smoke and cognitive development in a South African birth cohort study

By Claire A. Rowcliffe

Background and Aims: There is increasing evidence indicating that air pollution exposure is associated with neuronal damage. Since pregnancy is a critical window of vulnerability, air pollution exposure during this period could have adverse effects on neurodevelopment. This study aims 1) to analyze associations of prenatal exposure to indoor air pollution (particulate matter with diameters ≤10µm, PM_{10}) and tobacco smoke on neurodevelopment and 2) to determine if these associations are mediated by deviations of epigenetic gestational age from chronological gestational age (ΔGA).

Methods: Data of 734 children from the South African Drakenstein Child Health Study were analyzed. Prenatal PM_{10} exposure was measured using devices placed in the families' homes. Maternal smoking during pregnancy was determined by maternal urine cotinine measures. Bayley Scales of Infant Development were used to measure cognition, language, motor, and adaptive behavior development at two years of age. Overall composite scores were calculated as the average of the four sub-scores. Linear regression models adjusted for maternal age, gestational age, sex of child, ancestry, birth weight/length, and socioeconomic status were used to explore associations. A mediation analysis was conducted to analyze if the associations were mediated by ΔGA using DNA methylation measurements from cord blood.

Results: An increase of one interquartile range in PM_{10} (54.40 µg/m^3) was significantly associated with cognition sub-scores (β-estimate [95%-confidence interval]: -0.01 [-0.22, 0.00]). Maternal smoking was significantly associated with lower overall composite scores (-1.84 [-3.52, -0.16]) and lower adaptive behavior sub-scores (-3.39 [-5.63, -1.14]). Other scores were not associated with PM_{10} or smoking. Associations were not significantly mediated by ΔGA (e.g., for PM_{10} and cognition, proportion mediated [95%-confidence interval]: 1% [-12, 27%]).

Conclusion: We found an association of prenatal exposure to indoor air pollution (PM_{10}) and tobacco smoke on neurodevelopment at 2 years of age. Further research is needed to understand underlying biological mediators.

In-utero exposure to indoor air pollution or tobacco smoke and cognitive development in a South African birth cohort study

By

Claire Rowcliffe

Bachelor of Science

University of Tennessee – Knoxville

2018

Thesis Committee Chair: Anke Hüls, PhD

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Epidemiology

2021

Acknowledgements

I would like to give special thanks to Dr. Hüls, my advisor and chairperson on my committee, and the Hüls lab group for all the assistance and feedback on this study.

I would like to thank all the co-authors for this study: Junyu Chen, Aneesa Vanker, Nastassja Koen, Meaghan Jones, Nicole Gladish, Kirsten A. Donald, Catherine J Wedderburn, Michael S Kobor, Heather J Zar, and Dan J Stein.

In addition, I would like to thank all the earlier researchers on whose work I am able to build.

And finally, many thanks to my loving family for all the support in this effort.

Introduction

There is increasing evidence indicating that air pollution exposure is associated with an impairment of the central nervous system (CNS) and neuronal damage. Parts of the CNS develop at different stages due to cellular mechanisms like cell proliferation, differentiation, and migration. If these mechanisms were to be disrupted due to toxic environmental exposures, proper brain formation would not occur, leading to improper cognitive function. Failure in cell proliferation or cell migration because of exposure to toxic insults has profound deleterious effects on the developing brain (Costa et al., 2004). Though neurons are able to make new synapses throughout life, the period of brain development when synaptogenesis occurs is critical for the formation of the basic circuitry of the nervous system (Costa et al., 2019; Rodier, 1995). Since pregnancy is a critical window of developmental vulnerability for the brain, exposure to air pollution during this period could have a severe impact on neurodevelopment.

Currently, there is much research addressing an association between air pollution and neurological impairment. For instance, many studies have reported significant associations between air pollution and the development of dementia and Alzheimer's in older populations (Delgado-Saborit et al., 2021). Studies have reported a significant association between increased exposure to $PM_{2.5}$ and decreased cognitive function scores (Tallon et al., 2017), reduced reasoning ability (Tonne et al., 2014), slower reaction time (Cullen et al., 2018), worsened memory (J. A. Ailshire & Crimmins, 2014; Jennifer A. Ailshire & Clarke, 2015), and overall cognitive decline (Cacciottolo et al., 2017; Weuve et al., 2012).

Despite the knowledge that pregnancy is a critical window of developmental vulnerability for the brain, fewer studies have shown effects of prenatal exposure to air pollution on infants' cognitive development. Some studies report significant associations between air pollution and Autism Spectrum Disorder (ASD) in children. Two studies in Taiwan and in Pennsylvania, respectively, reported an increased risk of ASD associated with PM (Jung et al., 2013; Talbott et al., 2015). A similar study in two cohorts (North Carolina and California) also reported an association between PM exposure and ASD, particularly when exposure occurred in the third trimester of pregnancy (Kalkbrenner et al., 2015).

While there is some research addressing neurodevelopment, most studies have focused on the effects of outdoor air pollution. One cohort determined that prenatal exposure to polycyclic aromatic hydrocarbons (PAH), a chemical produced from burning coal, oil, or gasoline, was associated with a lower developmental index and IQ, and the odds of cognitive developmental delay were significantly greater for children with high prenatal exposure (Perera et al., 2006, 2009). Other studies have shown that exposure to NO_2 and benzene have an inverse association with mental development and result in reduced psychomotor development (Guxens et al., 2012, 2014). Lastly, a Japanese cohort determined that exposure to suspended PM, NO_2 and SO_2 had a negative association with verbal and fine motor development (Yorifuji et al., 2017). Although

these studies are assessing the association between prenatal exposure to air pollution and neurodevelopment, there is a clear gap in literature investigating indoor air pollution. Indoor air pollution is a leading cause of death and disability, worldwide (Lim et al., 2012) and is a particularly pertinent area of research for low-to-middle income countries due to the high percentage of households that still use traditional methods for cooking and heating. Based on these studies, this study is assessing an association currently under-researched.

Additionally, multiple studies have found associations between tobacco smoke and reduced neurological function or impaired neurodevelopment. For instance, a number of papers ascertained that smoking is associated with reduced neurological function in a variety of cognitive domains such as verbal memory, visual search speeds (Richards et al., 2003), and information processing speed (Starr et al., 2007). Another study determined that smoking was one of the strongest predictors of decreased neuronal processing capacity and speed (Aleman et al., 2005). In terms of impaired neurodevelopment, multiple studies have reported an association between tobacco smoke exposure and reduced intelligence (Huijbregts et al., 2006; Lawlor et al., 2006). Others report that children whose mother continues to smoke during pregnancy had a higher risk of behavior problems (Roza et al., 2009), slower reaction times (Mezzacappa et al., 2011), and decreased motor abilities (Polanska et al., 2013). However, all of these studies were conducted in high-income countries where exposure to tobacco smoke is typically low. In a country where exposure to tobacco smoke is much higher, could result in even stronger association with reduced neurological function.

Evidence accumulated so far strongly indicates that air pollution negatively disrupts proper CNS function and may be casually associate with neurodevelopment impairment. While little is known about the biology underlying this association, epigenetics has been discussed as a plausible mechanism by which environmental exposures might regulate the activity of genes relevant to child health. Epigenetic changes can affect gene expression in multiple different ways and possible types include DNA methylation, histone modification, and non-coding RNA function. DNA methylation modifies gene expression and provides a mechanism for propagating epigenetic information through DNA replication and cell division (Harman & Martín, 2020; Meng et al., 2019). Epigenetic mechanisms are critical regulators of development and proper functioning of the nervous system (Gapp et al., 2014). If epigenetic changes were to influence DNA methylation in developing infants, this could result in a deviation from actual gestational age and this deviation may influence neurodevelopment due to the temporal-dependent and systematic nature of brain formation. To the best of our knowledge, there is not much known about epigenetic alteration mechanisms in the context of neurodevelopment. Therefore, this study is aiming to also investigate epigenetics as a possible biological mechanism.

In this study, we aim to analyze 1) associations between prenatal exposure to indoor air pollution (particulate matter with diameters ≤10 micrometers, PM_{10}) and tobacco smoke

on neurodevelopment at 2 years of age and 2) to determine if these associations are mediated by deviations of epigenetic gestational age from chronological gestational age (ΔGA).

Methods

Study population and participants

The Drakenstein Child Health Study (DCHS) is located in the Drakenstein sub-district of the Western Cape, South Africa, a peri-urban area 60 km outside Cape Town (Zar et al., 2015). A detailed description of the enrollment process, inclusion criteria, exclusion criteria, assessment of home environment and indoor air pollution (IAP) exposures, measurement processing, and ethics of the study have been previously published (Vanker et al., 2015, 2017; Zar et al., 2015). We did a prospective longitudinal study of children enrolled in the Drakenstein Child Health Study (DCHS), a birth cohort study in a peri-urban area of South Africa that included follow-up through the first two years of life (Zar et al., 2015). Consenting pregnant women were enrolled at 20–28 weeks' gestation at two public primary health clinics serving different populations: Mbekweni (serving a predominantly black African population) and Newman (serving a predominantly mixed-race population) from March 1, 2012, to March 31, 2015 (Vanker et al., 2017; Zar et al., 2015). We excluded participants who were younger than 18 years, who did not attend study clinics for postnatal care (and thus could not be readily followed up), or who were intending to move out of the district within 2 years after the infant's birth (Vanker et al., 2017; Zar et al., 2015). All children were born at Paarl Hospital (Paarl, South Africa) (Vanker et al., 2017; Zar et al., 2015). In the original cohort, mother and infant pairs were followed at 6–10 weeks, 14 weeks, and 6, 9, and 12 months after birth (Vanker et al., 2017). For the purpose of this study, the follow up postnatal information from 6-10 weeks was used. Study questionnaires and clinical data were collected at enrollment and at each follow-up visit. 1137 mothers with 1143 livebirths were originally enrolled in DCHS, but only one livebirth per mother was included in this study's analysis to ensure independence among observations.

The DCHS study was approved by the Ethics Committee of the Faculty of Health Sciences, University of Cape Town, by Stellenbosch University and the Western Cape Provincial Research committee. Written informed consent was provided by the mothers for herself and her infant and is renewed annually.

Assessment of indoor air pollution (IAP) and tobacco smoke

A prenatal (within 4 weeks of enrollment) and postnatal (between 6-10 weeks of the infant's life) home visit was undertaken to measure IAP; the most common pollutants and by-products of combustion were measured (Vanker et al., 2017). IAP measurements included particulate matter of diameter 10 μm or less (PM_{10}), carbon monoxide, nitrogen dioxide, sulfur dioxide, and the volatile organic compounds benzene and toluene. Additionally, the source of heating and type of stove in each participant's house was noted. Most notably were the use of fossil fuels for cooking or heating based

on associations with benzene exposure (Vanker et al., 2015) and the use of paraffins for cooking or heating which produce volatile organic compounds on combustion (Vanker et al., 2017). During the postnatal home visit, these same measurements were repeated.

Mothers also provided urine within four weeks of enrollment for analysis of cotinine levels as a measure of smoke exposure (Vanker et al., 2015). Urine cotinine levels were classified as <10 ng/ml (non-smoker), 10–499 ng/ml (passive smoker), or ≥500 ng/ml (active smoker) (Vanker et al., 2015). In addition, we used a dichotomized smoking variable indicating if the pregnant women were exposed to any (≥10 ng/ml) versus no (<10 ng/ml) tobacco smoke; these participants were categorized as exposed or non-exposed, respectively. For the postnatal assessment of tobacco smoke exposure, mothers filled out a questionnaire approximately 6-10 weeks after giving birth asking current smoking status. If the mothers indicated they were a current smoker, a secondary question asked about how many cigarettes are smoked on average per day.

Epigenetic age predictors

DNA was isolated from cord blood samples that were collected at time of delivery (Morin et al., 2017). DNA methylation data quantified by the Illumina HumanMethylation450K BeadChip (450K) and Infinium MethylationEPIC BeadChip (EPIC) were combined, followed by quality control and normalization. Pre-processing and statistics were done using R 3.5.1 (R Core Team, 2018). Raw iDat files were imported to RStudio where intensity values were converted into beta values. The 450K and EPIC datasets were then combined using the minfi package (Aryee et al., 2014) resulting in 316 samples and 453,093 probes that were available on both arrays. Background subtraction, color correction and normalization were performed using the preprocessFunnorm function (Fortin et al., 2014). After sample and probe filtering, 273 samples and 409,033 probes remained for downstream analyses (Hüls et al., 2020).

The Bohlin-London predictor for each participant was calculated using predictGA function (predictGA package) in R. The Bohlin-London predictor is a weighted sum of beta values of 96 CpG sites with weights as the coefficients of these CpG sites obtained from perdition models trained in 1204 cord blood DNA (Bohlin et al., 2016). Among the 96 CpG sites necessary to compute Bohlin-London predictor, 8 were missing only among EPIC array and 3 were missing in both 450K and EPIC array. Median beta values from 450K array were used to impute the beta values of CpG sites missing only among EPIC array. Median beta values of CpG sites cg23457357, cg24366564 and cg17022232 from published DNA methylation data (GSE69176 and GSE97628) were used to impute missing values of corresponding CpG sites.

The Knight-Clock predictor were calculated using R codes provided from github (https://github.com/akknight/PredictGestationalAge). This predictor is a weighted sum of beta values of 142 CpG sites, 6 of which were missing and were imputed by k-nearest neighbors method implemented in the tool (Knight et al., 2016).

The Bohlin-London predictor was determined to be a better predictor of gestational age in our study based on how closely correlated the predicted epigenetic gestational age was to the actual gestational age. The correlation value of the Knight-Clock and Bohlin-London predictor was 0.387 and 0.641, respectively (**Figures S6, S7**). After merging phenotypic data with gestational age data, 191 cases remained for the mediation analysis.

Assessment of neurodevelopment

At two years of age, children in the DCHS underwent a neurodevelopment assessment using the Bayley Scales of Infant and Toddler Development, Third Edition (Bayley, 2006).The Bayley Scales are the most frequently used objective tests in infant developmental assessment by comparing neurodevelopmental delay across domains, in both clinical and research settings (Albuquerque et al., 2018; Anderson & Burnett, 2017; Johnson et al., 2014; Kaya-Kara et al., 2019). The Bayley Scales are composed of four domains – Cognitive, Language, Motor, and Adaptive Behavior. Trained assessors administered the assessment to generate scores for each domain via direct observation of the children (Donald et al., 2018). Composite scores have been used in previous research and validated for use in a South African setting (Ballot et al., 2017; Rademeyer & Jacklin, 2013). Additionally, for this study, an overall composite score was calculated as the average of the four domain sub-scores. If a child did not complete an assessment for each domain, the overall composite score was calculated as the average of the composite sub-scores available. 734 out of the 1137 children originally enrolled completed at least one Bayley Scales domain assessment and were subsequently included in the analysis. The results of the Bayley Scales composite sub-scores and the overall composite score followed a normal distribution (**Figures S1, S2, S3, S4, S5**).

Statistical Analysis

We used descriptive statistics to characterize the study population, summarizing categorical data as N (%) and continuous data as median (IQR). We calculated the Pearson correlations between all indoor air pollutants; no pollutants were heavily correlated (**Tables S5**). We used linear regression models to explore associations of indoor air pollution and maternal smoke exposure on neurodevelopment adjusting for demographic and socioeconomic characteristics (maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status); results are presented as β-estimates and 95% confidence intervals (CI). Results showing the association between PM_{10} and the Bayley Scale overall composite score are standardized by the IQR (54.40 $\mu g/m^3$). We estimated the association of prenatal PM_{10} and maternal tobacco smoke exposure with the overall composite Bayley Scales score as well as each Bayley Scale composite sub-score to assess if specific neurodevelopmental domains are more effected by indoor air pollution or tobacco smoke exposure than others. In addition to the single pollutant models, we conducted two- and multi-pollutant models. In the two-pollutant models, associations between PM_{10}

and the Bayley scores were additionally adjusted for smoking and associations between smoking and the Bayley scores were additionally adjusted for PM_{10}. In the multi-pollutant models, associations were additionally adjusted for benzene, toluene, SO_2, NO_2, CO, fossil stove, fossil heating, paraffin stove and paraffin heating. Lastly, a sensitivity analysis was conducted to identify susceptible time window of exposure. In the sensitivity analysis, the linear regression models estimating the association of prenatal PM_{10} and maternal tobacco smoke exposure with the overall composite Bayley Scales score as well as each Bayley Scale composite sub-score were compared to linear regression models estimating the association of postnatal PM_{10} and maternal tobacco smoke exposure with the overall composite Bayley Scales score as well as each Bayley Scale composite sub-score, respectively. We used R (version 4.0.2) for all data analysis.

A mediation analysis using the mediation package (Tingley et al., 2014) was conducted to analyze if the associations were mediated by deviations of epigenetic gestational age from chronological gestational age (ΔGA) using DNA methylation measurements from the cord blood samples. Since the Bohlin-London was determined to be a more accurate predictor of gestational age, this predictor was used for the mediation analysis. The mediator ((ΔGA) was calculated by subtracting the predicted epigenetic gestational age of the Bohlin-London predictor from the actual gestational age. Then, a mediator model was fitted where the measure of ΔGA was modeled as a linear regression function of the exposure variable (PM_{10} or maternal tobacco smoke) and confounding variables (maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status). Next, the outcome variable (Bayley Scales overall composite score, Bayley Scales cognition sub-score, Bayley Scales language sub-score, Bayley Scales motor sub-score, or Bayley Scales adaptive behavior sub-score) was modeled as a linear regression function of the mediator, exposure variable, and confounding variables. Then the mediate function (mediation package) was used to estimate the average causal mediation effects (Mediated), the average direct effects (Direct), the total effects of the mediator and outcome models (Total), and the proportion of the outcome mediated by the mediator. Results of the mediation analysis determined the mediated, direct, and total effects of the exposure on the outcome with the mediated effect accounting for the influence of the mediator, (ΔGA).

Results

Description of Study Participants (**Table 1**)

734 children from the South African Drakenstein Child Health Study were analyzed. Overall, about half of the children were male and of colored ancestry. Participants were categorized into quartiles as lowest, low-to-moderate, moderate-to-high, or highest socioeconomic status based on a composite socioeconomic status score (le Roux et al., 2015; Vanker et al., 2015, 2017; Zar et al., 2015). Overall, there was about an equal percent of participants of each SES category, but when the cohort was split into tertiles based on the Bayley scale overall composite score, the first tertile's largest groups of

participants were in the lowest and low-moderate SES categories, and the third tertiles's largest groups of participants were in the moderate- high and highest SES categories; the second tertile's largest groups were in the low-moderate and moderate-high SES categories. Birth weight and birth length remained similar across all three tertiles and the majority of mothers were exposed to tobacco smoke in all three tertiles.

PM_{10} and neurodevelopment

Among the 1137 mothers and livebirths, at least one Bayley Scales sub-score was completed for 734 children and therefore, these 734 participants were included in the statistical analysis. An increase of one IQR (54.4) in PM_{10} was not significantly associated with a lower composite score (β-estimate [95% confidence interval]: -0.514 [-1.087, -0.066]) once adjusted with confounding variables, nor when additionally adjusted for maternal tobacco smoke exposure in a two-pollutant model (-0.537 [-1.109, 0.034]) (**Figure 1A, Table S1A**). However, it was significantly associated with the Cognition sub-score (-0.011 [-0.219, -0.0003]) (**Figure 2A, Table S1B**) after adjusting for confounding variables and the composite score when included in a crude multi-pollutant model (-0.843 [-1.544, -0.131]), adjusted multi-pollutant model (-0.897 [-1.589, -0.203]), and extended multi-pollutant model (-0.906 [-1.603, -0.209]) (**Figure 1B, Table S1C**).. The crude multi-pollutant model assessed the association between multiple IAPs (PM_{10}, benzene, toluene , SO_2, NO_2, CO, fossil stove, fossil heating, paraffin stove, paraffin heating) and the Bayley Scales overall composite score, the adjusted model adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status, and the extended model additionally adjusted for maternal smoking.

Tobacco smoke and neurodevelopment

Being exposed to maternal smoking was determined to be significantly associated with a lower composite score in the crude model (-2.168 [-3.805, -0.532]) and after adjusting for confounding variables (-1.839 [-3.522, -0.158]), but associations were not significant when additionally adjusted for PM_{10} exposure in a two-pollutant model (-1.818 [-3.872, 0.236]) (**Figure 1C, Table S2A**). It was also significantly associated with a lower Adaptive Behavior sub-score (-3.386 [-5.632, -1.1389]) when adjusted for confounding variables (**Figure 2B, Table S2B**) and in a crude multi-pollutant model (-2.56 [-4.846, -0.279]) (**Figure 1D, Table S2C**). However, it was not significantly associated with a lower composite score when included in an adjusted multi-pollutant model (-1.945 [-4.362, -0.473]) or an extended multi-pollutant model (-1.945 [-4.362, 0.473]) (**Figure 1D, Table S2C**). The crude multi-pollutant model assessed the association between multiple IAPs (smoking, benzene, toluene , SO_2, NO_2, CO, fossil stove, fossil heating, paraffin stove, paraffin heating) and the Bayley Scales overall composite score, the adjusted model adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status, and the extended model additionally adjusted for PM_{10} exposure.

Mediation by ΔGA

In a subsequent mediation analysis, epigenetic gestational age was used as a biomarker of fetal development and used to analyze if the association between indoor air pollution and neurodevelopment is mediated by ΔGA using DNA methylation from cord blood. Analysis determined these associations are not significantly mediated by ΔGA. The proportion of the effect of PM_{10} on the overall composite score, cognition, language, motor, and adaptive behavior that goes through the mediator is about 2% [-107, 167%], 1% [-12, 27%], 0% [-136, 164%], 0% [-80, 110%], and 7% [-192, 235%], respectively (**Figure 3A, 3B, 3C, 3D, 3E**). The proportion of the effect of maternal smoking on the overall composite score, cognition, language, motor, and adaptive behavior that goes through the mediator is about 2% [-53, 55%], 1% [-42, 52%], 0% [-27, 36%], 0% [-98, 65%], and 6% [-82, 113%], respectively (**Figure 3F, 3G, 3H, 3I, 3J**).

Identification of susceptible time window of exposure

In order to investigate if postnatal exposure to indoor air pollution or tobacco smoke was also impacting the results of the Bayley scores, a sensitivity analysis was conducted. Due to the longitudinal aspect of the study, it is imperative to determine the time window of exposure at which cognitive development is most susceptible. This can be achieved by comparing linear regression models assessing the association of the Bayley Scales overall composite score with prenatal and postnatal exposure to the same IAP. In contrast to prenatal exposure, postnatal exposure to neither PM_{10} (**Figure 4A and 4B, Table S3**) nor tobacco smoke (**Figure 4C and 4D, Table S4**) was significantly associated with the Bayley Scales overall composite score or any of the four domain sub-scores.

Discussion

In our prospective longitudinal cohort study of pregnant women whose children underwent a neurodevelopment assessment at 2 years of age, our findings suggest that in-utero exposure to indoor air pollution and tobacco smoke may be negatively associated with neurodevelopment. Exposure to PM_{10} was associated with reduced cognitive function based on the Bayley Scales assessment, and exposure to maternal tobacco smoke was associated with reduced overall neurodevelopment, especially in regards to adaptive behavior, based on the Bayley Scales assessment. This indicates that prenatal indoor air pollution exposure effects different neurodevelopmental domains to varying extents. Subsequently, our findings indicated that the observed significance of indoor air pollution exposure on neurodevelopment is primarily the result of the prenatal, in-utero exposure, not postnatal early-life exposure. Previous studies have reported that the second trimester was the most sensitive time window for behavioral and activetone development (Chen et al., 2020), and therefore, this trimester could also be highly sensitive to air pollution exposure as well. However, this study did not investigate varying levels of indoor air pollution in the first, second, and third trimester, so future studies would be needed to assess this association.

In our study, association of prenatal exposure to PM_{10} and tobacco smoke on neurodevelopment were not significantly mediated by epigenetic gestational age (measured by DNA methylation in cord blood). Current research suggests that numerous indicators of aging are due to epigenetic mechanisms (Campisi & Vijg, 2009; Oberdoerffer & Sinclair, 2007). Because of this, it would not be unreasonable to suggest a single indicator of neurodevelopment, such as cognitive function, is influenced by multiple epigenetic mechanisms. Therefore, although deviations between chronological gestational age and epigenetic age did not show significant mediating effects of DNA methylation, other epigenetic factors could be influencing the role that air pollution has on neurodevelopment.

To our knowledge, this is the first study to assess prenatal exposure to indoor air pollutants on neurodevelopment in a large cohort study, and therefore, it fills a gap in epidemiologic and environmental health research. Moreover, this is also the first study to include research on the epigenetic mechanisms that may be influencing neurodevelopment in the context of exposure to indoor air pollutants. Although our findings determined these associations are not significantly mediated by ΔGA, other epigenetic mechanisms could be influencing neurodevelopment upon in-utero exposure to air pollution. Additionally, our findings of the sensitivity analysis determined these findings are predominantly due to prenatal exposure not postnatal exposure, indicating the importance of preventing exposure during pregnancy.

The DCHS is a unique birth cohort that enables assessment of various exposure assessments over time to investigate potential associations between in-utero air pollution and neurodevelopment due to prenatal recruitment and postnatal follow-ups. Prenatal recruitment allows assessment of exposure during gestations and after birth, use of a population-based sample helps to eliminate selection biases inherent in case-based approaches, and the cohort design ensures clear identification of the time-order of associations (Zar et al., 2015). Additional strengths include a large sample size that evaluated high-risk infants using the Bayley-III, a well- established assessment tool. The limitations of this study include the short PM_{10} exposure record time. Although PM_{10} was recorded prenatally and postnatally and therefore, highlighted multiple time points, it was only recorded for 24 hours at each assessment. Therefore, this measurement could potentially incorrectly reflect actual indoor air pollution exposure. Additionally, co-occurring medical diseases or illnesses in the mothers during pregnancy were not included in this analysis. A subsequent study of Aleman cohort (Aleman et al., 2005) determined that cardiovascular disease, one of the results of smoking, was associated with lower memory performance (Muller et al.). Therefore, if a mother had an undocumented or undiagnosed disease or illness while pregnant, this may affect the infant's neurodevelopment, especially if that illness or disease has been associated with smoking or previously linked to impaired cognitive function. Another limitation was the cross-sectional aspect of the neurodevelopment assessment. Neurodevelopment occurred over multiple years, but the Bayley scores assessment was only applied once when the child was two years of age.

Conclusion

In summary, our findings show a significant association between prenatal exposure to indoor air pollution (PM_{10}) and tobacco smoke on neurodevelopment at 2 years of age. However, the mediating effect of epigenetic gestational age does not significantly impact these associations. Further research to understand underlying biological mediators is needed.

References

Ailshire, J. A., & Crimmins, E. M. (2014). Fine Particulate Matter Air Pollution and Cognitive Function Among Older US Adults. *American Journal of Epidemiology*, *180*(4), 359–366. https://doi.org/10.1093/aje/kwu155

Ailshire, Jennifer A., & Clarke, P. (2015). Fine Particulate Matter Air Pollution and Cognitive Function Among U.S. Older Adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *70*(2), 322–328. https://doi.org/10.1093/geronb/gbu064

Albuquerque, P. L. de, Guerra, M. Q. de F., Lima, M. de C., & Eickmann, S. H. (2018). Concurrent validity of the Alberta Infant Motor Scale to detect delayed gross motor development in preterm infants: A comparative study with the Bayley III. *Developmental Neurorehabilitation*, *21*(6), 408–414. https://doi.org/10.1080/17518423.2017.1323974

Aleman, A., Muller, M., de Haan, E. H. F., & van der Schouw, Y. T. (2005). Vascular risk factors and cognitive function in a sample of independently living men. *Neurobiology of Aging*, *26*(4), 485–490. https://doi.org/10.1016/j.neurobiolaging.2004.05.005

Anderson, P. J., & Burnett, A. (2017). Assessing developmental delay in early childhood - concerns with the Bayley-III scales. *The Clinical Neuropsychologist*, *31*(2), 371–381. https://doi.org/10.1080/13854046.2016.1216518

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

Ballot, D. E., Ramdin, T., Rakotsoane, D., Agaba, F., Davies, V. A., Chirwa, T., & Cooper, P. A. (2017). Use of the Bayley Scales of Infant and Toddler Development, Third Edition, to Assess Developmental Outcome in Infants and Young Children in an Urban Setting in South Africa. *International Scholarly Research Notices*, *2017*, 1631760. https://doi.org/10.1155/2017/1631760

Bayley, N. (2006). No Title. *Bayley Scales of Infant and Toddler Development, Technical Manual. 2006; 3rd Ed.*

Bohlin, J., Håberg, S. E., Magnus, P., Reese, S. E., Gjessing, H. K., Magnus, M. C., Parr, C. L., Page, C. M., London, S. J., & Nystad, W. (2016). Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biology*, *17*(1), 207. https://doi.org/10.1186/s13059-016-1063-4

Cacciottolo, M., Wang, X., Driscoll, I., Woodward, N., Saffari, A., Reyes, J., Serre, M. L., Vizuete, W., Sioutas, C., Morgan, T. E., Gatz, M., Chui, H. C., Shumaker, S. A., Resnick, S. M., Espeland, M. A., Finch, C. E., & Chen, J. C. (2017). Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. *Translational Psychiatry*, *7*(1), e1022–e1022. https://doi.org/10.1038/tp.2016.280

Campisi, J., & Vijg, J. (2009). Does Damage to DNA and Other Macromolecules Play a Role in Aging? If So, How? *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *64A*(2), 175–178. https://doi.org/10.1093/gerona/gln065

Chen, B., Huang, S., He, J., He, Q., Chen, S., Liu, X., Peng, S., Luo, D., & Duan, Y. (2020). Sex-specific influence of prenatal air pollutant exposure on neonatal neurobehavioral development and the sensitive window. *Chemosphere*, *254*, 126824. https://doi.org/10.1016/j.chemosphere.2020.126824

Costa, L. G., Aschner, M., Vitalone, A., Syversen, T., & Soldin, O. P. (2004). Developmental neuropathology of environmental agents. *Annual Review of Pharmacology and Toxicology*, *44*, 87–110. https://doi.org/10.1146/annurev.pharmtox.44.101802.121424

Costa, L. G., Cole, T. B., Dao, K., Chang, Y.-C., & Garrick, J. M. (2019). Developmental impact of air pollution on brain function. *Neurochemistry International*, *131*(1), 104580. https://doi.org/10.1016/j.neuint.2019.104580

Cullen, B., Newby, D., Lee, D., Lyall, D. M., Nevado-Holgado, A. J., Evans, J. J., Pell, J. P., Lovestone, S., & Cavanagh, J. (2018). Cross-sectional and longitudinal analyses of outdoor air pollution exposure and cognitive function in UK Biobank. *Scientific Reports*, *8*(1), 12089. https://doi.org/10.1038/s41598-018-30568-6

Delgado-Saborit, J. M., Guercio, V., Gowers, A. M., Shaddick, G., Fox, N. C., & Love, S. (2021). A critical review of the epidemiological evidence of effects of air pollution on dementia, cognitive function and cognitive decline in adult population. *The Science of the Total Environment*, *757*, 143734. https://doi.org/10.1016/j.scitotenv.2020.143734

Donald, K. A., Hoogenhout, M., du Plooy, C. P., Wedderburn, C. J., Nhapi, R. T., Barnett, W., Hoffman, N., Malcolm-Smith, S., Zar, H. J., & Stein, D. J. (2018). Drakenstein Child Health Study (DCHS): investigating determinants of early child development and cognition. *BMJ Paediatrics Open*, *2*(1), e000282. https://doi.org/10.1136/bmjpo-2018-000282

Fortin, J.-P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., Greenwood, C. M., & Hansen, K. D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biology*, *15*(12), 503. https://doi.org/10.1186/s13059-014-0503-2

Gapp, K., Woldemichael, B. T., Bohacek, J., & Mansuy, I. M. (2014). Epigenetic regulation in neurodevelopment and neurodegenerative diseases. *Neuroscience*, *264*, 99–111. https://doi.org/10.1016/j.neuroscience.2012.11.040

Harman, M. F., & Martín, M. G. (2020). Epigenetic mechanisms related to cognitive decline during aging. *Journal of Neuroscience Research*, *98*(2), 234–246. https://doi.org/10.1002/jnr.24436

Huijbregts, S. C. J., Séguin, J. R., Zelazo, P. D., Parent, S., Japel, C., & Tremblay, R. E. (2006). Interrelations between maternal smoking during pregnancy, birth weight

and sociodemographic factors in the prediction of early cognitive abilities. *Infant and Child Development*, *15*(6), 593–607. https://doi.org/10.1002/icd.480

Hüls, A., Wedderburn, C., Groenewold, N., Gladish, N., Jones, M., Koen, N., Maclsaac, J., Lin, D., Ramadori, K., Epstein, M., Donald, K. A., Kobor, M. S., Zar, H. J., & Stein, D. J. (2020). Newborn Differential DNA Methylation and Subcortical Brain Volumes as Early Signs of Severe Neurodevelopmental Delay in a South African Birth Cohort Study. *MedRxiv*. https://doi.org/https://doi.org/10.1101/2020.03.13.20035501

Johnson, S., Moore, T., & Marlow, N. (2014). Using the Bayley-III to assess neurodevelopmental delay: which cut-off should be used? *Pediatric Research*, *75*(5), 670–674. https://doi.org/10.1038/pr.2014.10

Jung, C.-R., Lin, Y.-T., & Hwang, B.-F. (2013). Air pollution and newly diagnostic autism spectrum disorders: a population-based cohort study in Taiwan. *PloS One*, *8*(9), e75510. https://doi.org/10.1371/journal.pone.0075510

Kalkbrenner, A. E., Windham, G. C., Serre, M. L., Akita, Y., Wang, X., Hoffman, K., Thayer, B. P., & Daniels, J. L. (2015). Particulate matter exposure, prenatal and postnatal windows of susceptibility, and autism spectrum disorders. *Epidemiology (Cambridge, Mass.)*, *26*(1), 30–42. https://doi.org/10.1097/EDE.0000000000000173

Kaya-Kara, Ö., Kerem-Günel, M., & Yiğit, Ş. (2019). Correlation of the Bayley scales of infant-toddler development-3rd edition and neuro-sensory motor assessment in preterm infants during the first year of life. *The Turkish Journal of Pediatrics*, *61*(3), 399–406. https://doi.org/10.24953/turkjped.2019.03.012

Knight, A. K., Craig, J. M., Theda, C., Bækvad-Hansen, M., Bybjerg-Grauholm, J., Hansen, C. S., Hollegaard, M. V, Hougaard, D. M., Mortensen, P. B., Weinsheimer, S. M., Werge, T. M., Brennan, P. A., Cubells, J. F., Newport, D. J., Stowe, Z. N., Cheong, J. L. Y., Dalach, P., Doyle, L. W., Loke, Y. J., … Smith, A. K. (2016). An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biology*, *17*(1), 206. https://doi.org/10.1186/s13059-016-1068-z

Lawlor, D. A., Najman, J. M., Batty, G. D., O'Callaghan, M. J., Williams, G. M., & Bor, W. (2006). Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. *Paediatric and Perinatal Epidemiology*, *20*(2), 148–162. https://doi.org/10.1111/j.1365-3016.2006.00704.x

le Roux, D. M., Myer, L., Nicol, M. P., & Zar, H. J. (2015). Incidence and severity of childhood pneumonia in the first year of life in a South African birth cohort: the Drakenstein Child Health Study. *The Lancet. Global Health*, *3*(2), e95–e103. https://doi.org/10.1016/S2214-109X(14)70360-2

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., AlMazroa, M. A., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., … Ezzati, M. (2012). A comparative risk assessment of

burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, *380*(9859), 2224–2260. https://doi.org/10.1016/S0140-6736(12)61766-8

Meng, S., Zhou, H., Feng, Z., Xu, Z., Tang, Y., & Wu, M. (2019). Epigenetics in Neurodevelopment: Emerging Role of Circular RNA. *Frontiers in Cellular Neuroscience*, *13*, 327. https://doi.org/10.3389/fncel.2019.00327

Mezzacappa, E., Buckner, J. C., & Earls, F. (2011). Prenatal cigarette exposure and infant learning stimulation as predictors of cognitive control in childhood. *Developmental Science*, *14*(4), 881–891. https://doi.org/10.1111/j.1467-7687.2011.01038.x

Morin, A. M., Gatev, E., McEwen, L. M., MacIsaac, J. L., Lin, D. T. S., Koen, N., Czamara, D., Räikkönen, K., Zar, H. J., Koenen, K., Stein, D. J., Kobor, M. S., & Jones, M. J. (2017). Maternal blood contamination of collected cord blood can be identified using DNA methylation at three CpGs. *Clinical Epigenetics*, *9*(1), 75. https://doi.org/10.1186/s13148-017-0370-2

Oberdoerffer, P., & Sinclair, D. A. (2007). The role of nuclear architecture in genomic instability and ageing. *Nature Reviews Molecular Cell Biology*, *8*(9), 692–702. https://doi.org/10.1038/nrm2238

Polanska, K., Hanke, W., Sobala, W., Trzcinka-Ochocka, M., Ligocka, D., Brzeznicki, S., Strugala-Stawik, H., & Magnus, P. (2013). Developmental Effects of Exposures to Environmental Factors: The Polish Mother and Child Cohort Study. *BioMed Research International*, *2013*, 1–11. https://doi.org/10.1155/2013/629716

R Core Team. (2018). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna*. https://www.r-project.org

Rademeyer, V., & Jacklin, L. (2013). A study to evaluate the performance of black South African urban infants on the Bayley Scales of Infant Development III. *South African Journal of Child Health*, *7*(2), 54. https://doi.org/10.7196/sajch.547

Richards, M., Jarvis, M. J., Thompson, N., & Wadsworth, M. E. J. (2003). Cigarette smoking and cognitive decline in midlife: evidence from a prospective birth cohort study. *American Journal of Public Health*, *93*(6), 994–998. https://doi.org/10.2105/ajph.93.6.994

Rodier, P. M. (1995). Developing brain as a target of toxicity. *Environmental Health Perspectives*, *103 Suppl*, 73–76. https://doi.org/10.1289/ehp.95103s673

Roza, S. J., Verhulst, F. C., Jaddoe, V. W., Steegers, E. A., Mackenbach, J. P., Hofman, A., & Tiemeier, H. (2009). Maternal smoking during pregnancy and child behaviour problems: the Generation R Study. *International Journal of Epidemiology*, *38*(3), 680–689. https://doi.org/10.1093/ije/dyn163

Starr, J. M., Deary, I. J., Fox, H. C., & Whalley, L. J. (2007). Smoking and cognitive change from age 11 to 66 years: a confirmatory investigation. *Addictive Behaviors*,

32(1), 63–68. https://doi.org/10.1016/j.addbeh.2006.03.020

Talbott, E. O., Arena, V. C., Rager, J. R., Clougherty, J. E., Michanowicz, D. R., Sharma, R. K., & Stacy, S. L. (2015). Fine particulate matter and the risk of autism spectrum disorder. *Environmental Research*, *140*, 414–420. https://doi.org/10.1016/j.envres.2015.04.021

Tallon, L. A., Manjourides, J., Pun, V. C., Salhi, C., & Suh, H. (2017). Cognitive impacts of ambient air pollution in the National Social Health and Aging Project (NSHAP) cohort. *Environment International*, *104*, 102–109. https://doi.org/10.1016/j.envint.2017.03.019

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation : R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5). https://doi.org/10.18637/jss.v059.i05

Tonne, C., Elbaz, A., Beevers, S., & Singh-Manoux, A. (2014). Traffic-related air pollution in relation to cognitive function in older adults. *Epidemiology (Cambridge, Mass.)*, *25*(5), 674–681. https://doi.org/10.1097/EDE.0000000000000144

Vanker, A., Barnett, W., Nduru, P. M., Gie, R. P., Sly, P. D., & Zar, H. J. (2015). Home environment and indoor air pollution exposure in an African birth cohort study. *The Science of the Total Environment*, *536*, 362–367. https://doi.org/10.1016/j.scitotenv.2015.06.136

Vanker, A., Barnett, W., Workman, L., Nduru, P. M., Sly, P. D., Gie, R. P., & Zar, H. J. (2017). Early-life exposure to indoor air pollution or tobacco smoke and lower respiratory tract illness and wheezing in African infants: a longitudinal birth cohort study. *The Lancet. Planetary Health*, *1*(8), e328–e336. https://doi.org/10.1016/S2542-5196(17)30134-1

Weuve, J., Puett, R., & Schwartz, J. (2012). Exposure to Particulate Air Pollution and Cognitive Decline in Older Women. *Archives of Internal Medicine*, *172*(3), 219. https://doi.org/10.1001/archinternmed.2011.683

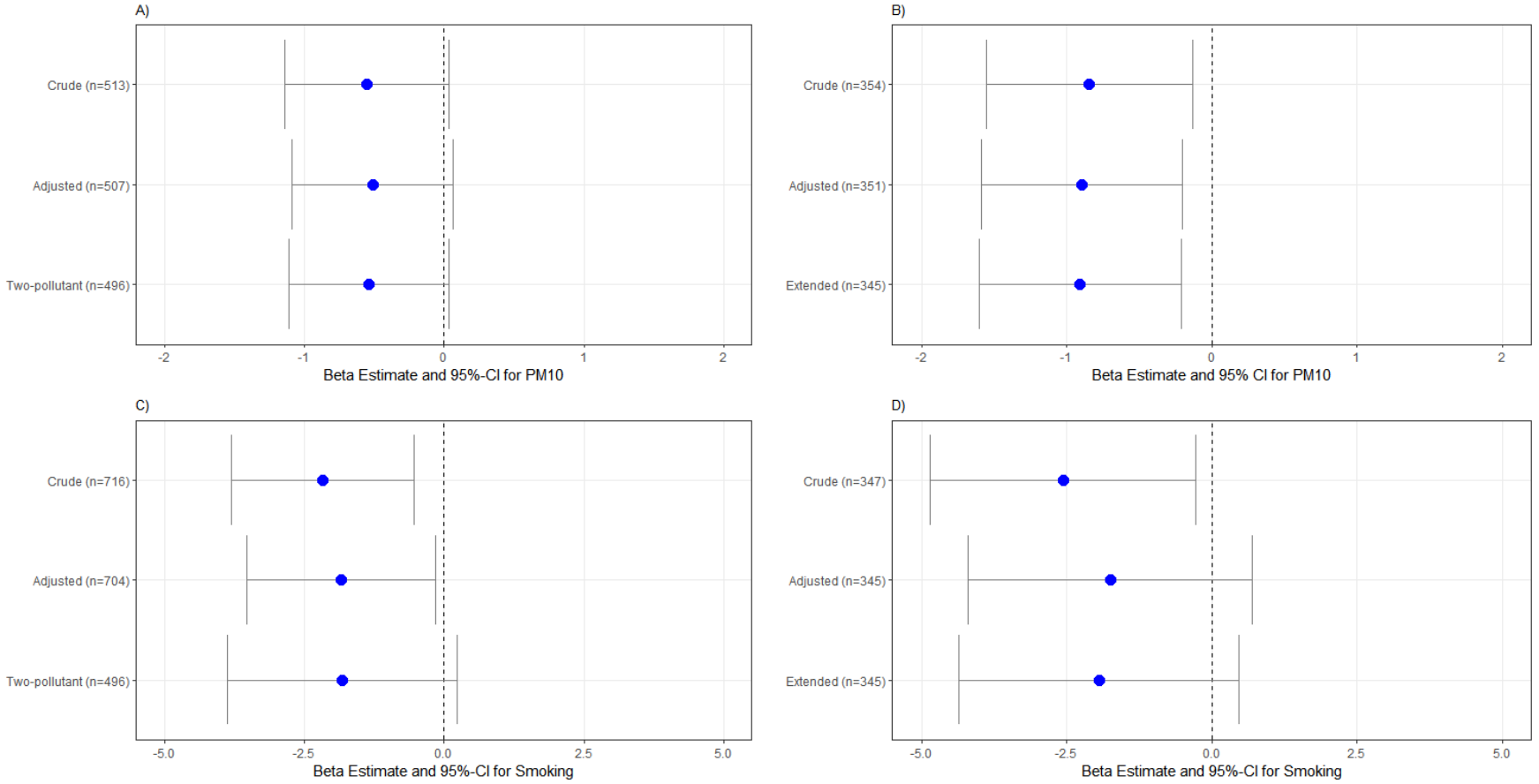
Zar, H. J., Barnett, W., Myer, L., Stein, D. J., & Nicol, M. P. (2015). Investigating the early-life determinants of illness in Africa: the Drakenstein Child Health Study. *Thorax*, *70*(6), 592–594. https://doi.org/10.1136/thoraxjnl-2014-206242

Tables and Figures

Table 1. Characteristics of birth cohort and prenatal IAP measurements

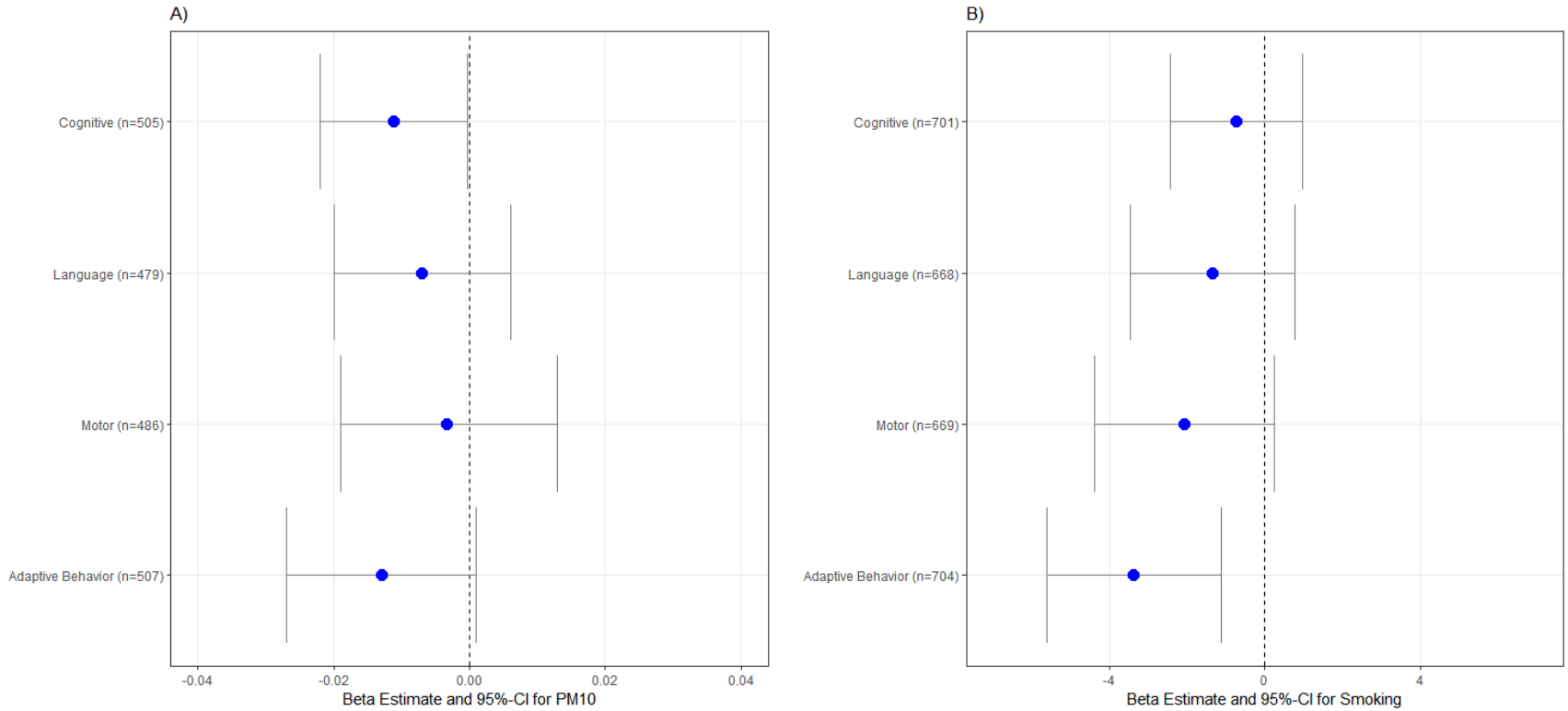
	Overall	First Tertile of Overall Composite Score	Second Tertile of Overall Composite Score	Third Tertile of Overall Composite Score
N=	734	246	243	245
Maternal Age (yrs) (median [IQR])	26.59 [22.49, 31.35]	26.72 [22.65, 31.90]	26.32 [22.69, 31.66]	26.38 [22.21, 30.70]
Infant Sex = Male (%)	379 (51.6)	158 (64.2)	113 (46.5)	108 (44.1)
Gestational Age (wks) (median [IQR])	39.00 [38.00, 40.00]	39.00 [37.00, 40.00]	39.00 [38.00, 40.00]	39.00 [38.00, 40.00]
Birth Weight (lbs) (median [IQR])	3.10 [2.71, 3.41]	2.99 [2.56, 3.30]	3.09 [2.74, 3.37]	3.21 [2.80, 3.50]
Birth Length (cm) (median [IQR])	50.00 [48.00, 52.00]	50.00 [47.00, 52.00]	50.00 [48.00, 52.00]	50.00 [48.00, 52.00]
Ancestry = Colored (%)	348 (47.4)	102 (41.5)	115 (47.3)	131 (53.5)
Parental Socioeconomic status (%)				
Lowest SES	180 (24.5)	83 (33.7)	54 (22.2)	43 (17.6)
Low-moderate SES	196 (26.7)	73 (29.7)	65 (26.7)	58 (23.7)
Moderate-high SES	193 (26.3)	53 (21.5)	66 (27.2)	74 (30.2)
Highest SES	165 (22.5)	37 (15.0)	58 (23.9)	70 (28.6)
Maternal Smoking = Exposed (%)	559 (78.1)	199 (83.3)	180 (76.3)	180 (74.7)
Maternal Smoking (%)				
Non-Smoker	157 (21.9)	40 (16.7)	56 (23.7)	61 (25.3)
Passive Smoker	308 (43.0)	110 (46.0)	95 (40.3)	103 (42.7)
Active Smoker	251 (35.1)	89 (37.2)	85 (36.0)	77 (32.0)
Benzene Concentration ($\mu g/m^3$) (median [IQR])	3.67 [1.45, 9.71]	4.34 [1.67, 10.46]	3.10 [1.07, 7.95]	3.98 [1.65, 9.98]
Toluene Concentration ($\mu g/m^3$) (median [IQR])	16.57 [6.79, 46.24]	19.51 [8.09, 57.28]	13.47 [5.82, 37.83]	17.49 [8.17, 38.99]
SO_2 Concentration ($\mu g/m^3$) (median [IQR])	0.00 [0.00, 0.28]	0.00 [0.00, 0.23]	0.00 [0.00, 0.49]	0.00 [0.00, 0.18]
NO_2 Concentration ($\mu g/m^3$) (median [IQR])	6.61 [3.36, 12.59]	6.32 [2.79, 12.61]	7.03 [3.88, 13.22]	6.58 [3.43, 11.38]
PM_{10} Concentration ($\mu g/m^3$) (median [IQR])	35.51 [13.17, 67.57]	43.26 [17.80, 72.64]	31.17 [10.24, 61.20]	32.73 [13.22, 62.89]
CO Concentration (mg/m^3) (median [IQR])	0.00 [0.00, 2.86]	0.00 [0.00, 2.56]	0.00 [0.00, 5.42]	0.00 [0.00, 2.65]
Fossil Stove = Yes (%)	107 (19.3)	35 (18.5)	40 (22.7)	32 (16.8)
Fossil Heating = Yes (%)	87 (15.7)	28 (14.8)	35 (19.9)	24 (12.6)
Paraffin Heating = Yes (%)	77 (13.9)	24 (12.7)	31 (17.6)	22 (11.7)
Paraffin Stove = Yes (%)	76 (13.7)	25 (13.2)	31 (17.6)	20 (10.5)
Bayley Scales-Cognitive (median [IQR])	85.00 [80.00, 90.00]	80.00 [70.00, 85.00]	90.00 [85.00, 90.00]	90.00 [90.00, 95.00]
Bayley Scales-Language(median [IQR])	83.00 [77.00, 91.00]	74.00 [71.00, 79.00]	83.00 [79.00, 89.00]	94.00 [89.00, 100.00]
Bayley Scales-Motor(median [IQR])	92.50 [85.00, 100.00]	85.00 [79.00, 88.00]	91.00 [88.00, 97.00]	100.00 [97.00, 110.00]
Bayley Scales-Adaptive Behavior (median [IQR])	84.00 [72.25, 92.00]	69.00 [65.00, 76.00]	84.00 [78.00, 88.50]	93.00 [88.00, 99.00]
Bayley Scales-Overall Composite (median [IQR])	86.25 [80.25, 92.25]	77.50 [73.25, 80.25]	86.25 [84.29, 88.50]	94.25 [92.25, 97.75]

Figure 1.



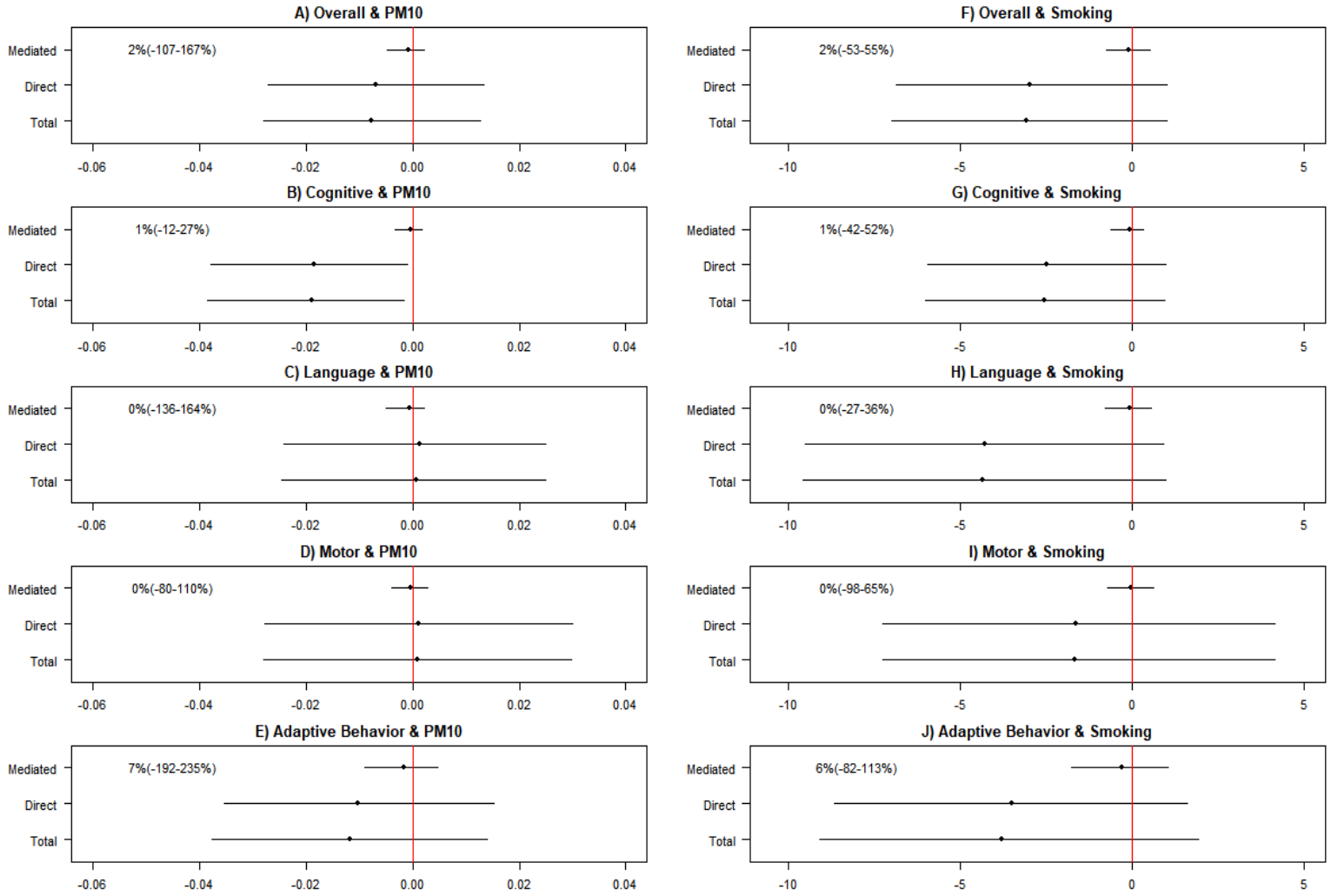
A. PM_{10} and Bayley Scales overall composite score (single pollutant models, standardized by IQR [54.4]). Crude: crude association between PM_{10} and Bayley score (-0.544 [-1.141, 0.033]); Adjusted: adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status (-0.514 [-1.087, 0.066]); Two-pollutant: additionally adjusted for maternal_smoking (-0.537 [-1.109, 0.034]). **B. PM_{10} and Bayley Scales overall composite score (multi-pollutant models, standardized by IQR [54.4]).** Crude: crude association between PM_{10} and Bayley score adjusted for IAPs (benzene, toluene , SO2, NO2, CO, fossil stove, fossil heating, paraffin stove, paraffin heating) (-0.843 [-1.544, -0.131]); Adjusted: additionally adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status (-0.897 [-1.589, -0.203]); Extended: additionally, adjusted for maternal smoking (-0.906 [-1.603, -0.209]). **C. Smoking and Bayley Scales overall composite score (single pollutant models).** Crude: crude association between maternal tobacco smoke and Bayley score (-2.168 [-3.805, -0.532]); Adjusted: adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status (-1.839 [-3.522, -0.158]); Two-pollutant: additionally adjusted for PM_{10} (-1.818 [-3.872, 0.236]). **D. Smoking and Bayley Scales overall composite score (multi-pollutant models).** Crude: crude association between smoking and Bayley score adjusted for IAPs (benzene, toluene , SO2, NO2, CO, fossil stove, fossil heating, paraffin stove, paraffin heating) (-2.56 [-4.846, -0.279]); Adjusted: additionally adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status (-1.945 [-4.193, 0.701]); Extended: additionally adjusted for PM_{10} (-1.945 [-4.362, 0.473]).

Figure 2.



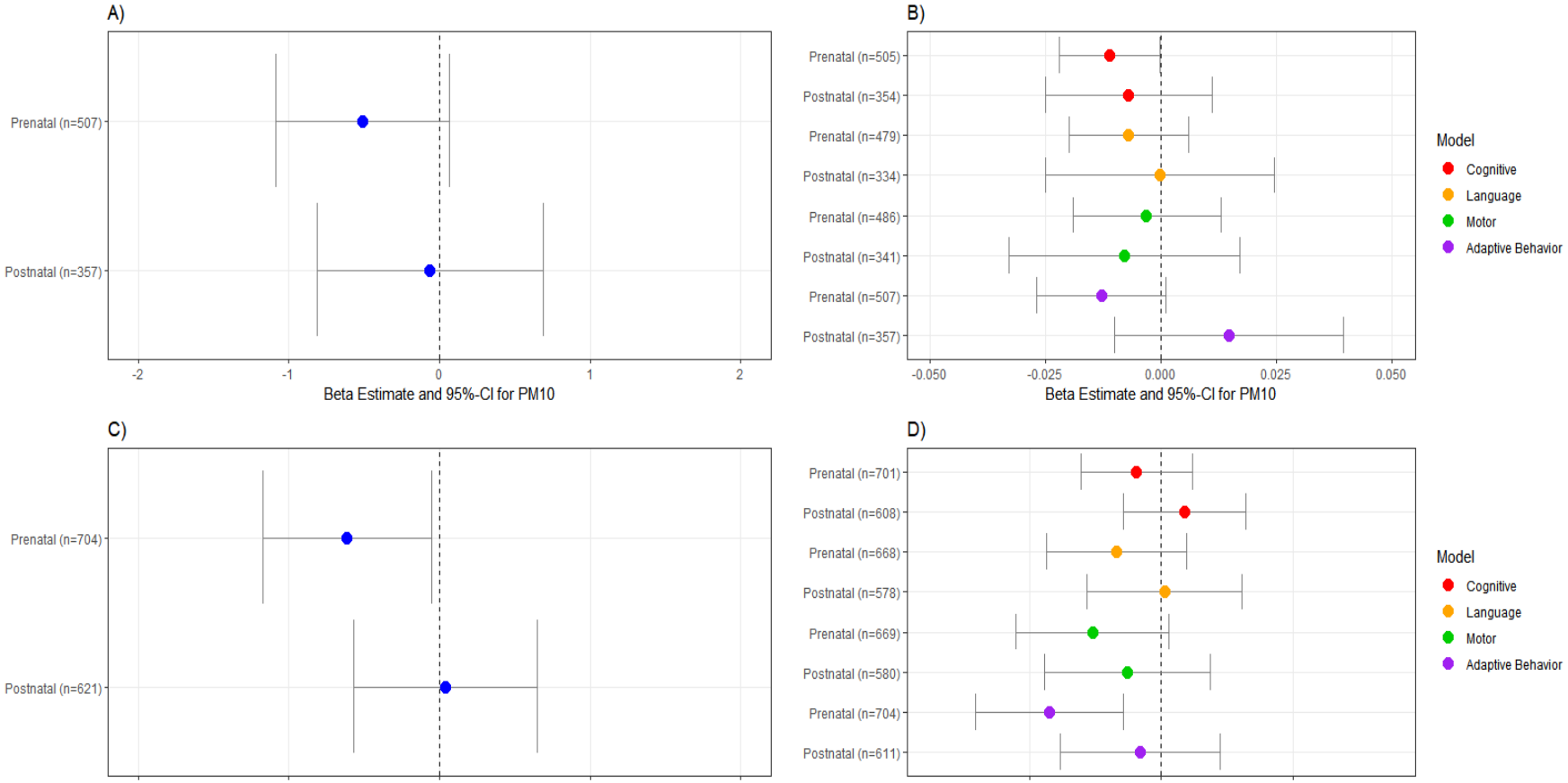
A. Single pollutant analysis of PM_{10} exposure on the Bayley Scales four composite sub-scores adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status; Cognitive (-0.011 [-0.022, -0.0003]), Language (-0.007 [-0.021, 0.006]), Motor (-0.003 [-0.019, 0.013]), Adaptive Behavior (-0.013 [-0.027, 0.001]). **B.** Single pollutant analysis of maternal tobacco smoke exposure on the Bayley Scales four composite sub-scores adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status; Cognitive (-0.745 [-2.440, 0.951]), Language (-1.352 [-3.485, 0.779]), Motor (-2.080 [-4.404, 0.244]), Adaptive Behavior (-3.386 [-5.632, -1.139]).

Figure 3.



All figures show the average causal mediation effects (Mediated), the average direct effects (Direct), the total effects of the mediator and outcome models (Total), and the proportion of the outcome mediated by the mediator. **A.** Mediation analysis of ΔGA and PM_{10} exposure on Bayley Scales overall composite score adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status. **B-E.** Mediation analysis of ΔGA and PM_{10} exposure on Bayley Scales four composite sub-scores adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status. **F.** Mediation analysis of ΔGA and maternal tobacco smoke exposure on Bayley Scales overall composite score adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status. **G-J.** Mediation analysis of ΔGA and maternal tobacco smoke exposure on Bayley Scales four composite sub-scores adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status.

Figure 4.



Identification of the most vulnerable time window of exposure. **A. PM_{10} and Bayley Scales overall composite score.** Prenatal (-0.514 [-1.087, 0.066]), results standardized by IQR (54.4); Postnatal (-0.063 [-0.811, 0.685]), results standardized by IQR (40.55). **B. PM_{10} and Bayley Scales composite sub-scores.** Cognitive: Prenatal (-0.011 [-0.022, -0.0003]), Postnatal (-0.007 [-0.025, 0.011]); Language: Prenatal (-0.007 [-0.021, 0.006]), Postnatal (-0.0002 [-0.025, 0.024]); Motor: Prenatal (-0.003 [-0.019, 0.013]), Postnatal (-0.008 [-0.033, 0.017]); Adaptive Behavior: Prenatal (-0.013 [-0.027, 0.001]), Postnatal (0.015 [-0.009, 0.039]). **C. Smoking and Bayley Scales overall composite score.** Prenatal (-1.839 [-3.522, -0.158]); Postnatal (-2.305 [-6.023, 1.412]). **D. Smoking and Bayley Scales composite sub-scores.** Cognitive: Prenatal (-0.745 [-2.440, 0.951]), Postnatal (0.708 [-1.160, 2.570]); Language: Prenatal (-1.352 [-3.485, 0.779]), Postnatal (0.100 [-2.258, 2.458]); Motor: Prenatal (-2.080 [-4.404, 0.244]), Postnatal (-1.028 [-3.536, 1.479]); Adaptive Behavior: Prenatal (-3.386 [-5.632, -1.139]), Postnatal (-0.637 [-3.065, 1.792]). Prenatal: Prenatal PM_{10} or tobacco smoke exposure (within 4 weeks of enrollment), Postnatal: Postnatal PM_{10} or tobacco smoke exposure (between 6-10 weeks of the infant's life); All associations were adjusted for maternal age, gestational age, sex of child, ancestry, birth weight, birth length, and parental socioeconomic status

Table S1. PM_{10}

A. Association of PM_{10} and Bayley Scale Overall Composite Score			
	β-Estimate	95% CI	Pr(> |t|)
Crude (n=513)*	-0.554	(-1.141, 0.033)	0.065
Adjusted (n=351)*	-0.514	(-1.087, 0.066)	0.083
Two-pollutant (n=496)*	-0.537	(-1.109, 0.034)	0.065

B. Association of PM_{10} and Bayley Scale Sub-scores~			
Cognitive (n=505)	-0.011	(-0.022, -0.0003)	0.045
Language (n=479)	-0.007	(-0.021, 0.006)	0.295
Motor (n=486)	-0.003	(-0.019, 0.013)	0.682
Adaptive Behavior (n=507)	-0.013	(-0.027, 0.001)	0.076

C. Multi-pollutant model: Association of PM_{10} and Bayley Scale Overall Composite Score			
Crude (n=354)*	-0.843	(-1.554, -0.131)	0.021
Adjusted (n=351)*	-0.897	(-1.589, -0.203)	0.011
Extended (n=345)*	-0.906	(-1.603, -0.209)	0.011

*These values have been standardized by the IQR (54.4)
~The adjusted single-pollutant model was used

Table S2. Smoking

A. Association of Maternal Smoking (Unexposed, Exposed) and Bayley Scale Overall Composite Score			
	β-Estimate	95% CI	Pr(> |t|)
Crude (n=716)	-2.168	(-3.805, -0.532)	0.009
Adjusted (n=704)	-1.839	(-3.522, -0.158)	0.032
Two-pollutant (n=496)	-1.818	(-3.872, 0.236)	0.083
B. Association of Maternal Smoking and Bayley Scale Sub-scores~			
Cognitive (n=701)	-0.745	(-2.440, 0.951)	0.389
Language (n=668)	-1.352	(-3.485, 0.779)	0.213
Motor (n=669)	-2.080	(-4.404, 0.244)	0.079
Adaptive Behavior (n=704)	-3.386	(-5.632, -1.139)	0.003
C. Multi-pollutant Model: Association of Maternal Smoking (Unexposed, Exposed) and Bayley Scales Composite Score			
Crude (n=347)	-2.560	(-4.846, -0.279)	0.028
Adjusted (n=345)	-1.945	(-4.193, 0.701)	0.115
Extended (n=345)	-1.945	(-4.362, 0.473)	0.115

~The adjusted single-pollutant model was used

Table S3. Identification of the most vulnerable time window of exposure: PM_{10}

A. Association of PM_{10} and Bayley Scale Overall Composite Score			
	β-Estimate	95% CI	Pr(> |t|)
Prenatal (n=704)*	-0.514	(-1.087, 0.066)	0.083
Postnatal (n=357)**	-0.063	(-0.811, 0.685)	0.868
B. Association of PM_{10} and Bayley Scale Sub-scores~			
Cognitive			
Prenatal (n=505)	-0.011	(-0.022, -0.0003)	0.045
Postnatal (n=354)	-0.007	(-0.025, 0.011)	0.447
Language			
Prenatal (n=479)	-0.007	(-0.021, 0.006)	0.295
Postnatal (n=334)	-0.0002	(-0.025, 0.024)	0.984
Motor			
Prenatal (n=486)	-0.003	(-0.019, 0.013)	0.682
Postnatal (n=341)	-0.008	(-0.033, 0.017)	0.528
Adaptive Behavior			
Prenatal (n=507)	-0.013	(-0.027, 0.001)	0.076
Postnatal (n=357)	0.015	(-0.009, 0.039)	0.241

*These values have been standardized by the IQR (54.4)
**These values have been standardized by the IQR (40.55)
~The adjusted single-pollutant model was used

Table S4. Identification of the most vulnerable time window of exposure: Smoking

A. Association of Maternal Smoking (Unexposed, Exposed) and Bayley Scale Overall Composite Score			
	β-Estimate	95% CI	Pr(> |t|)
Prenatal (n=704)	-1.839	(-3.522, -0.158)	0.032
Postnatal (n=621)	-2.305	(-6.023, 1.412)	0.222
Avg # of Cigarettes per Day (n=621)	-0.0008	(-0.003, 0.001)	0.412
B. Association of Maternal Smoking and Bayley Scale Sub-scores~			
Cognitive			
Prenatal (n=701)	-0.745	(-2.440, 0.951)	0.389
Postnatal (n=608)	0.708	(-1.160, 2.57)	0.456
Language			
Prenatal (n=668)	-1.352	(-3.485, 0.779)	0.213
Postnatal (n=578)	0.100	(-2.258, 2.458)	0.933
Motor			
Prenatal (n=669)	-2.080	(-4.404, 0.244)	0.079
Postnatal (n=580)	-1.028	(-3.536, 1.479)	0.421
Adaptive Behavior			
Prenatal (n=704)	-3.386	(-5.632, -1.139)	0.0032
Postnatal (n=611)	-0.637	(-3.065, 1.792)	0.607

~The adjusted single-pollutant model was used

Table S5. Pearsons Correlations of IAP

	Benzene	NO_2	SO_2	Toluene	PM_{10}	CO
Benzene						
NO_2	0.07					
SO_2	-0.02	0.52				
Toluene	0.58	0.03	-0.02			
PM_{10}	0.01	0.12	-0.01	0.03		
CO	0.02	0.16	-0.02	-0.04	0.33	
Bayley Overall Composite Score	-0.04	0.01	0.07	-0.04	-0.08	0.03

Figure S1.Histogram for Cognitive Skill of Bayley Scale

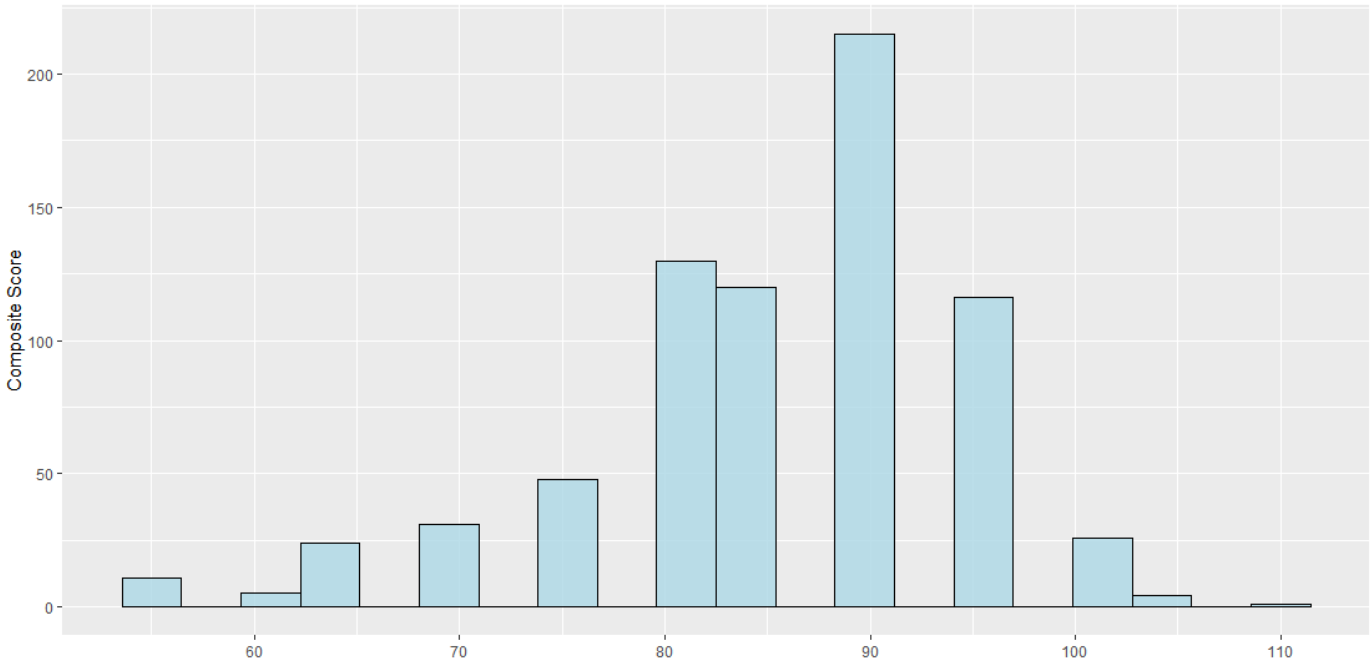


Figure S2.Histogram for Language Skill of Bayley Scale

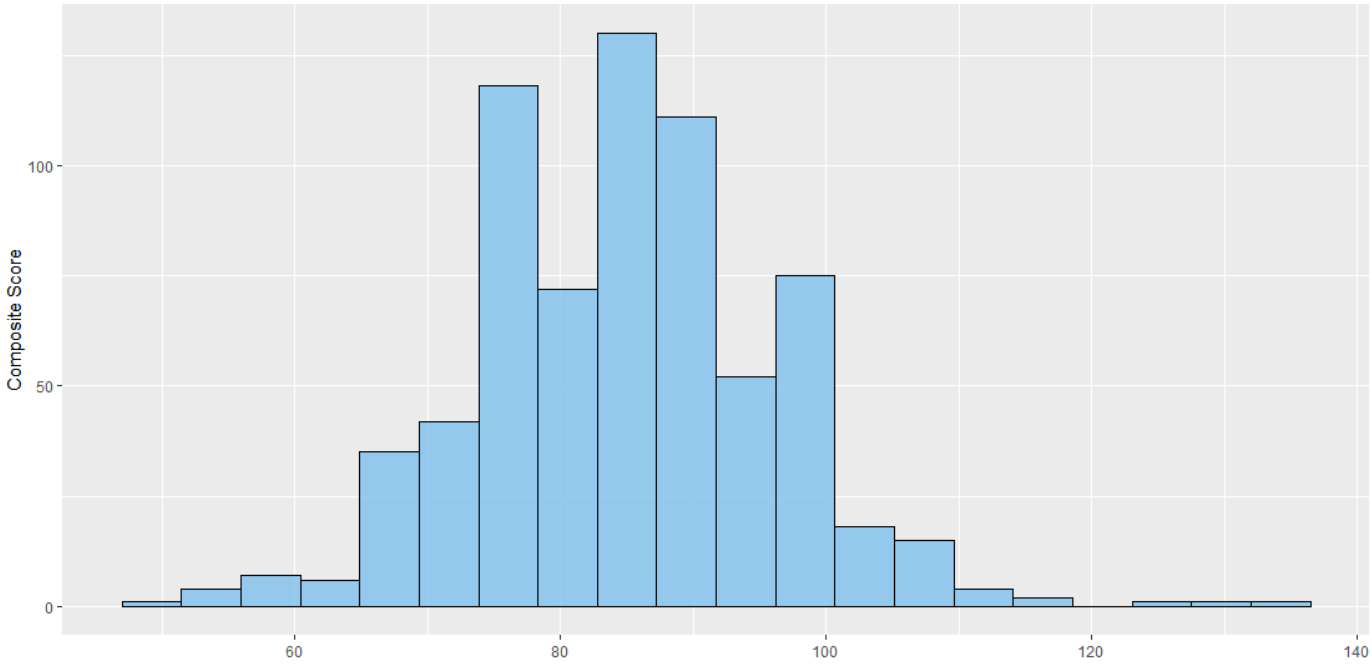


Figure S3.Histogram for Motor Skill of Bayley Scale

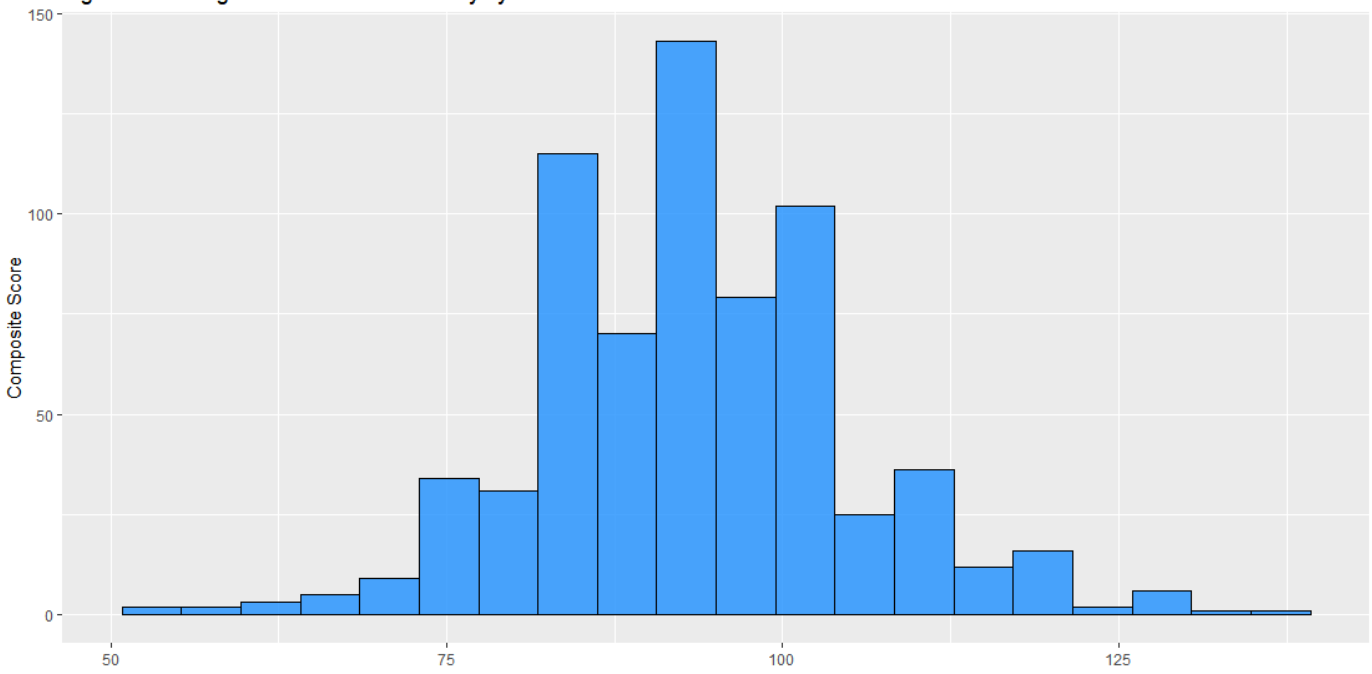


Figure S4.Histogram for Adaptive Behavior Skill Bayley Scale

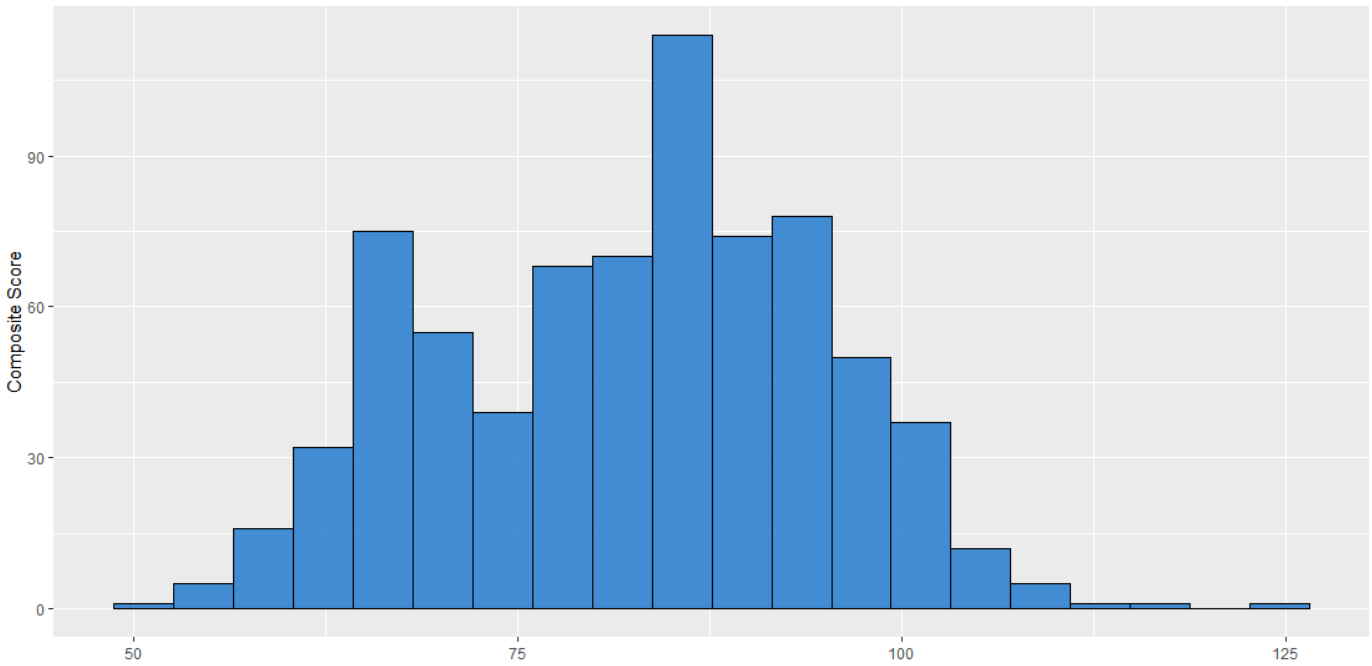


Figure S5.Histogram for Overall Composite of Bayley Scales

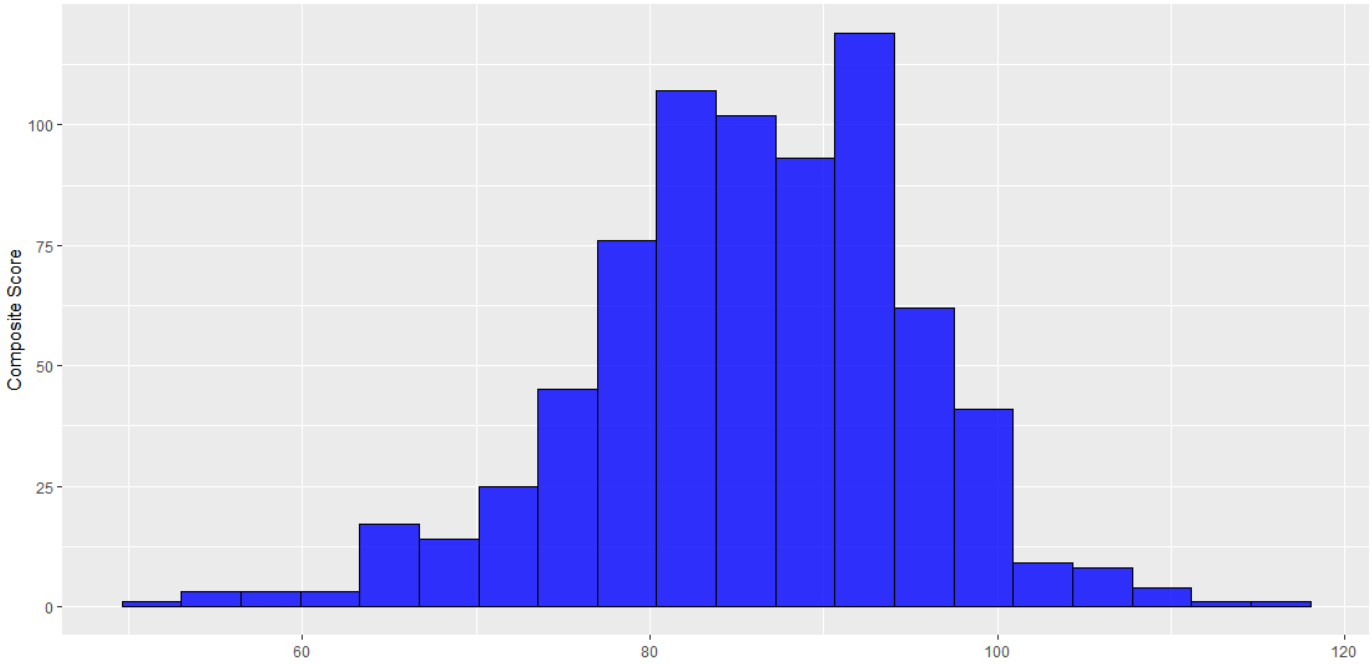
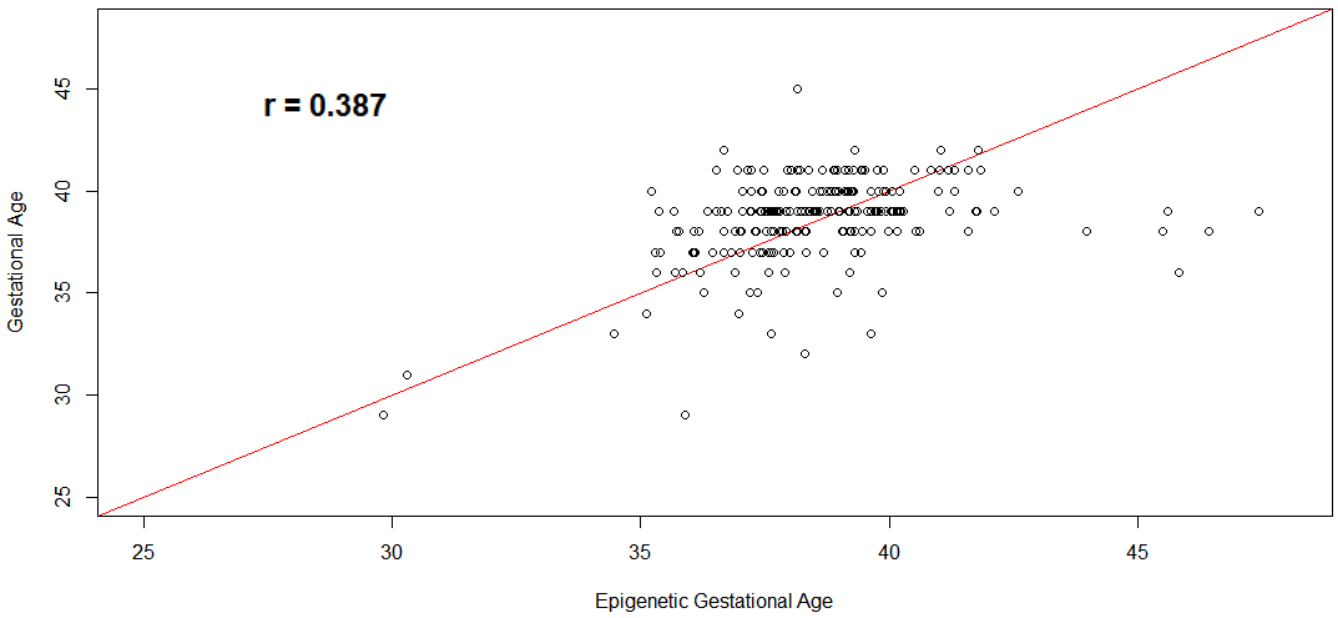
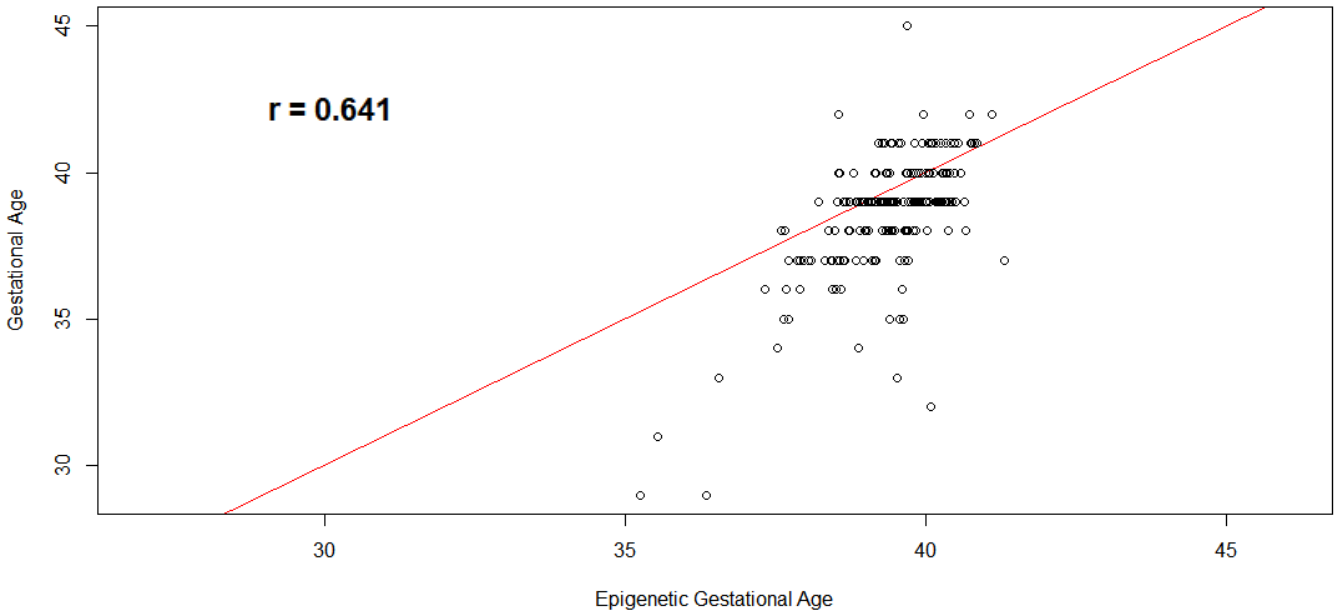


Figure S6.Knight-Clock Correlation



This graph shows the epigenetic gestational age predicted by the Knight-Clock predictor versus the actual gestational age retrieved from medical records. The Pearson's correlation coefficient (r) is an indicator of how closely accurate the predictor was able to predict the gestational age (r = 0.387).

Figure S7.Bohlin-London Correlation



This graph shows the epigenetic gestational age predicted by the Bohlin-London predictor versus the actual gestational age retrieved from medical records. The Pearson's correlation coefficient (r) is an indicator of how closely accurate the predictor was able to predict the gestational age (r = 0.641).