

ARTICLE OPEN



Meta-analysis of epigenome-wide associations between DNA methylation at birth and childhood cognitive skills

Doretta Caramaschi ^{1,2,25}✉, Alexander Neumann^{3,4,25}, Andres Cardenas ⁵, Gwen Tindula⁶, Silvia Alemany^{7,8,9}, Lea Zillich¹⁰, Giancarlo Pesce ¹¹, Jari M. T. Lahti ¹², Alexandra Havdahl ^{1,13}, Rosa Mulder^{3,14,15}, Janine F. Felix ^{14,15}, Henning Tiemeier ³, Lea Sirignano ¹⁰, Josef Frank ¹⁰, Stephanie H. Witt ¹⁰, Marcella Rietschel ¹⁰, Michael Deuschle¹⁶, Karen Huen⁶, Brenda Eskenazi¹⁷, Tabea Sarah Send¹⁶, Muriel Ferrer⁷, Maria Gilles¹⁶, Maria de Agostini¹⁸, Nour Baiz¹¹, Sheryl L. Rifas-Shiman¹⁹, Tuomas Kvist¹², Darina Czamara ²⁰, Samuli T. Tuominen¹², Caroline L. Relton¹, Dheeraj Rai¹, Stephanie J. London ²¹, Katri Räikkönen¹², Nina Holland⁶, Isabella Annesi-Maesano ¹¹, Fabian Streit ¹⁰, Marie-France Hivert¹⁹, Emily Oken¹⁹, Jordi Sunyer^{7,8,9,22}, Charlotte A. M. Cecil^{3,23,24} and Gemma Sharp ¹

© The Author(s) 2022

Cognitive skills are a strong predictor of a wide range of later life outcomes. Genetic and epigenetic associations across the genome explain some of the variation in general cognitive abilities in the general population and it is plausible that epigenetic associations might arise from prenatal environmental exposures and/or genetic variation early in life. We investigated the association between cord blood DNA methylation at birth and cognitive skills assessed in children from eight pregnancy cohorts within the Pregnancy And Childhood Epigenetics (PACE) Consortium across overall (total $N = 2196$), verbal (total $N = 2206$) and non-verbal cognitive scores (total $N = 3300$). The associations at single CpG sites were weak for all of the cognitive domains investigated. One region near *DUSP22* on chromosome 6 was associated with non-verbal cognition in a model adjusted for maternal IQ. We conclude that there is little evidence to support the idea that variation in cord blood DNA methylation at single CpG sites is associated with cognitive skills and further studies are needed to confirm the association at *DUSP22*.

Molecular Psychiatry (2022) 27:2126–2135; <https://doi.org/10.1038/s41380-022-01441-w>

INTRODUCTION

The human brain starts developing prenatally in the third gestational week and its maturation extends postnatally into late adolescence and most likely adulthood. These processes occur as the result of genetic and environmental factors including the interplay between them [1]. General cognitive ability, or intelligence, often measured as intelligent quotient (IQ), shows considerable heritability, with estimates as high as 20% in infancy, increasing to more than 70% in adulthood [2]. Socioeconomic factors are also associated with cognitive skills and can moderate genetic influences

so that their effects vary across socioeconomic strata [3, 4]. It has been shown that cognitive skills predict important long-term outcomes such as higher educational attainment, later mortality, and better physical and mental health [5]. Already in childhood, cognitive functioning is strongly associated with higher educational attainment and later adulthood mortality [6, 7].

Since there is considerable environmental influence on cognitive ability in early childhood [1], it is plausible that environmental exposures that occur in early life, either prenatally or in early childhood, play a role in shaping children's cognitive

¹Medical Research Council Integrative Epidemiology Unit (MRC IEU), Bristol Medical School, Population Health Science, University of Bristol, Bristol, UK. ²Department of Psychology, College of Life and Environmental Sciences, University of Exeter, Exeter, UK. ³Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center Rotterdam, Rotterdam, The Netherlands. ⁴Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada. ⁵Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, Berkeley, CA, USA. ⁶Children's Environmental Health Laboratory, Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA, USA. ⁷ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain. ⁸Universitat Pompeu Fabra (UPF), Barcelona, Spain. ⁹CIBER Epidemiology and Public Health (CIBERESP), Madrid, Spain. ¹⁰Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Heidelberg, Germany. ¹¹Epidemiology of Allergic and Respiratory Diseases Team (EPAR), Institute Pierre Louis of Epidemiology and Public Health, UMR-S 1136 INSERM and Sorbonne Université, Paris, France. ¹²Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland. ¹³Department of Mental Disorders, Norwegian Institute of Public Health and Nic Waals Institute of Lovisenberg Diaconal Hospital, Oslo, Norway. ¹⁴The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. ¹⁵Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. ¹⁶Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. ¹⁷Center for Environmental Research and Children's Health (CERCH), School of Public Health, University of California, Berkeley, CA, USA. ¹⁸Inserm, Centre for Research in Epidemiology and Statistics (CRESS), Research Team on Early Life Origins of Health (EAROH), Villejuif, France. ¹⁹Division of Chronic Disease Research Across the Lifecourse, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA. ²⁰Department of Translational Research in Psychiatry, Max-Planck-Institute of Psychiatry, Munich, Germany. ²¹National Institute of Environmental Health Sciences, National Institutes of Health, Department of Health and Human Services, Research Triangle Park, NC, USA. ²²Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain. ²³Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. ²⁴Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. ²⁵These authors contributed equally: Doretta Caramaschi, Alexander Neumann. ✉email: d.caramaschi@exeter.ac.uk

Received: 29 April 2021 Accepted: 11 January 2022

Published online: 10 February 2022

development. Early life environmental exposures are known to be revealed in the epigenome, as reflected in changes in DNA methylation marks on cytosine nucleotides followed by guanine (CpG) across the genome. The total heritability of DNA methylation levels from single nucleotide polymorphisms has been estimated to be on average only 19% across the genome, suggesting a strong environmental component [8]. Moreover robust evidence indicate that maternal smoking and folate levels during pregnancy are associated with changes in the child's epigenome at birth [9, 10]. DNA methylation at birth can therefore reflect exposure to adverse factors, which in turn could have neurodevelopmental consequences.

A recent epigenome-wide association study (EWAS) of cognitive measures in older-aged adults across several cohorts, with final sample sizes ranging between 2557 and 6809 participants, identified associations of DNA methylation with global cognitive function at one intergenic CpG site on chromosome 12 and with phonemic verbal fluency at one CpG site on chromosome 10 in the *INPP5A* gene [11]. Another study of educational attainment in 10767 adults, although not specific to cognitive skills, revealed associations at nine CpG sites which are all known to be associated with smoking [12]. Methylation at each CpG explained 0.3–0.7% of the variance in educational attainment. At present, it is not known whether there is a prospective association between DNA methylation at birth and later cognitive functioning, and whether any association found might be indicative of prenatal exposures rather than environmental exposures across the lifetime such as own smoking habits. Other neurodevelopmental traits have been investigated in relation to DNA methylation at birth. Associations with attention-deficit hyperactivity disorder have been found at five CpG sites in cord blood in a robust meta-analysis [13] and at three sites for social communication trajectories, although the latter were cohort-specific [14]. There was less evidence of associations of DNA methylation at birth with autism spectrum disorder, although there was a strong association with genetic liability for autism [15, 16].

In this study, we extended the previous research on adulthood cognitive function to early life and we aimed to investigate whether DNA methylation was associated with cognitive skills already at birth. As we used a genome-wide approach that interrogates more than 400,000 CpG sites across the genome, we examined associations both independently at each CpG site and at CpG clusters, based on the knowledge that nearby CpG sites are often correlated and that methylation differences might be found across regions spanning multiple CpGs. Specifically we investigated whether: (1) DNA methylation at the single CpG site level in cord blood is prospectively associated with cognitive skills in childhood; (2) DNA methylation at the regional level in cord blood is associated with cognitive skills in childhood.

To assess the biological relevance of our findings and considering that DNA methylation is often tissue-specific and that we did not have access to brain tissue, we also performed a look-up of published brain data to examine the correspondence of DNA methylation between blood and brain at the relevant CpG sites and the potential consequences of DNA methylation differences at these sites on gene expression.

METHODS

Study sample

The data used in this study were previously obtained from the participants of eight longitudinal birth cohorts within the Pregnancy and Childhood Epigenetics (PACE) Consortium [17] who agreed to contribute to the meta-analysis. The final sample size is therefore determined opportunistically by the number of participants that had available data for this project in each cohort. The cohorts were: Avon Longitudinal Study of Parents and Children (ALSPAC) [18, 19], Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) [20, 21], Etude des Déterminants pré et

post natal du développement et de la santé de l'Enfant (EDEN) [22], Generation R [23], Infancia y Medio Ambiente project (INMA) [24], Prediction and Prevention of Preeclampsia and Intrauterine Growth Restriction study (PREDO) [25], Pre- Peri- and Postnatal Stress: Epigenetic impact on Depression (POSEIDON) [26], and Project Viva [27]. Participants were all singleton births. Ethical approval for each study was obtained by local committees and consent to use their data was obtained for all participants. Approved researchers with access to individual-level data for each cohort performed in-house analyses and shared only result files with the main analysts. Access to individual-level data is available only upon request to each cohort separately and following local procedures. For more information on each cohort, ethical approval and data access procedures please refer to the Supplementary Material.

Epigenetic data

For each cohort, cord blood was collected during delivery and DNA was isolated according to standard protocols. DNA was then bisulfite-treated according to standard protocols and loaded onto Infinium HumanMethylation450 BeadChip or Infinium MethylationEPIC arrays (Illumina, San Diego, CA). Array images were scanned and raw methylation intensities were normalized and subjected to quality control according to cohort-specific procedures. For some cohorts, batch correction was applied before analysis (PREDO, INMA and Project Viva), while for the others batch variables were included as covariates in the analysis. The epigenetic data were already pre-processed within each cohort at the time of the analysis, so it was not possible to standardize the procedure across all cohorts. Previous studies in the PACE consortium used this procedure successfully [28–30]. Full information on the methods used within each cohort is reported in the Supplementary Methods and Materials.

Cognitive skills

Cognitive skills were measured differently across cohorts, depending on available data at cohort-average ages ranging from 4 to 9 years (see Supplementary Material and Methods for more details). Main cognitive scores and subtests were then used to represent overall, verbal and non-verbal cognitive domains.

Overall cognitive skills were measured by:

- the Wechsler Intelligence Scale for Children, 4th edition, (WISC-IV) [31] full-scale IQ score in CHAMACOS and PREDO;
- the Wechsler Intelligence Scale for Children, 3rd edition, (WISC-III) [32] full-scale IQ score in ALSPAC;
- the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), 3rd edition [33] full-scale IQ score in EDEN and POSEIDON;
- the McCarthy Scales of Children's Abilities [34] general cognitive index in INMA;
- the Wide Range Assessment of Visual Motor Ability (WRAVMA) test [35] in Project Viva.

Verbal cognitive skills were measured by:

- the Wechsler Intelligence Scale for Children, 4th edition, (WISC-IV) [31] verbal comprehension index in CHAMACOS and PREDO;
- the Wechsler Intelligence Scale for Children, 3rd edition, (WISC-III) [32] verbal IQ scores (derived from the verbal comprehension index and working memory index, combined) in ALSPAC;
- the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), 3rd edition [33] verbal IQ score in EDEN and POSEIDON;
- the McCarthy Scales of Children's Abilities [34] verbal index in INMA;
- the Kaufman Brief Intelligence Test 2nd edition (KBIT-II) [36] verbal subtest in Project Viva.

Non-verbal cognitive skills were measured by:

- the Wechsler Intelligence Scale for Children, 4th edition, (WISC-IV) [31] perceptual reasoning index in CHAMACOS and PREDO;
- the Wechsler Intelligence Scale for Children, 3rd edition, (WISC-III) [32] performance IQ scores (derived from the perceptual organization index and processing speed index, combined) in ALSPAC;
- the Wechsler Preschool and Primary Scale of Intelligence (WPPSI), 3rd edition [33] performance IQ score in EDEN and POSEIDON;
- the McCarthy Scales of Children's Abilities [34] perceptual-performance index in INMA;
- the Kaufman Brief Intelligence Test 2nd edition (KBIT-II) [36] non-verbal subtest in Project Viva;

Table 1. Overview of the studies participating in the meta-analysis.

Cohort	Cognitive outcome	Age (years)	Overall cognitive abilities (N)			Verbal cognitive abilities (N)			Non-verbal cognitive abilities (N)					
			Main	Paternal education	Maternal IQ	Main	Paternal education	Maternal IQ	Main	Paternal education	Maternal IQ	PCs		
ALSPAC	WISC-III	9	780	739	-	684	784	743	-	687	783	742	-	687
CHAMACOS	WISC-IV	7	175	175	175	-	175	175	175	-	175	175	175	-
EDEN	WPPSI	6	157	157	-	-	157	157	-	-	157	157	-	-
Generation R	SON-R	6	-	-	-	-	-	-	-	-	1093	953	1058	1058
INMA	McCarthy	4	319	319	311	280	319	319	311	280	319	319	311	280
POSEIDON	WPPSI	4	199	198	-	198	201	200	-	200	203	202	-	202
PREDO	WISC-IV	9	285	192	-	285	285	192	-	285	285	192	-	285
Project Viva	KBIT-II	8	281	269	279	-	285	273	283	-	285	273	283	-
Total			2196	2049	765	1447	2206	2059	769	1452	3300	3013	1827	2512

The main model was adjusted for child age at cognitive testing, sex, maternal age at delivery, maternal education, birth weight, gestational age at birth, maternal smoking status during pregnancy, parity at delivery, batch covariates and cell type proportions. The other models included also paternal education, maternal IQ (in place of maternal education) and genomic PCs.

- the Snijders-Oomen non-verbal intelligence tests, revised (SON-R) [37], in Generation R.

The average cognitive skills scores across cohort methylation subsamples are reported in Supplementary Table ST1. Since the distribution and the scores used differed across the cohort methylation subsamples, the cognitive scores were transformed into standardized z-scores. Full information on the methods used within each cohort is reported in Supplementary Material and Methods.

Epigenome-wide association study (EWAS)

Prior to the analyses, a data analysis plan including details of the variables, the models to use, and a sample R code was distributed to the participating cohorts. The analysis plan and the code can be accessed upon request to the corresponding author. Untransformed DNA methylation beta values were used as the exposure variable. Extreme outliers (>3 × interquartile range, either side of the 25th and 75th percentiles) were removed.

The effect of DNA methylation at birth on childhood cognitive skills was estimated by linear regression models (*lm()* option in R) within each cohort for each CpG site individually. The main models included the covariates: child age at cognitive testing, sex, maternal age at delivery, maternal education (cohort-specific definition), birth weight, gestational age at birth, maternal smoking status during pregnancy (any smoking compared to no smoking), parity at delivery (1 or more previous children compared to none), batch covariates (cohort-specific definition) and proportions of seven blood cell types estimated using the Houseman algorithm using a published reference dataset for cord blood [38]. The covariates chosen for the main model were those that maximized the sample size as they were available in all cohorts.

As sensitivity analyses, three other models were run including other covariates in cohorts with the relevant data available: main model covariates + paternal education, main model covariates with maternal IQ in place of maternal education, and main model covariates + the first ten principal components from children’s genomic data. Maternal IQ was not included in the main model since not all the cohorts had this variable available. We excluded maternal education in the model with maternal IQ to avoid multicollinearity due to the high correlation between these two variables. Genomic PCs were not included in the main model as they were not available for all cohorts. Each model was run for overall, verbal and non-verbal cognitive skills. Details of the EWAS run within each cohort and specific information on the covariates used are in Supplementary Material and Methods. A summary of the models and the corresponding sample sizes is available in Table 1.

Meta-analysis

The results from the EWAS carried out in each cohort were subjected to an initial screening to check that results were comparable across cohorts using the *QCEWAS* R package [39]. Through visual inspection we assessed similarity of effect size distribution across cohorts and the presence of a linear relationship between precision (1/medianSE) and sample size (sqrt-transformed), as suggested by the package manual. A fixed-effect meta-analysis was then performed for each model and cognitive outcome using the inverse-variance weighted approach in the *metafor* R package [40]. An independent meta-analysis was conducted at a different research institution using the *metasoft* software [41] to confirm the results. Multiple comparisons were taken into account by setting a Bonferroni-corrected threshold of 1.02×10^{-7} for the 450 K arrays, which is conservative when taking into account the reduced number of CpGs common to all studies (i.e., excluding outlier and low detection probes and restricting to common probes between the 450 K and EPIC arrays). The number of CpG sites was 379445, 379445, 321854 for overall, verbal and non-verbal cognitive skills, respectively.

As some heterogeneity was expected due to the different studies and methods, we also performed a random effect meta-analysis using the DerSimonian-Laird estimator in the *metafor* R package on the 500 sites that had the lowest *p* values in the fixed effects meta-analysis. Statistical heterogeneity was assessed using the *I*² statistics for each CpG site [42]. As previously recommended [42], heterogeneity was described low, moderate and high around *I*² values of 25%, 50% and 75%, respectively. To explore heterogeneity and since the year of birth varied substantially across the studies, we also performed a random effect meta-regression on the top 500 CpG sites from the fixed effects meta-analysis by including year at the start of the cohort (centered to the mean) as a

covariate. To further account for heterogeneity, as cognitive skills were measured using different methods across the studies and at different ages, we also looked for Bonferroni-significant results within each individual study. Finally, we also meta-analyzed the results from the most homogenous group in terms of cognitive skills assessments. This was comprised of the ALSPAC, PREDO and CHAMACOS cohorts, where cognitive skills were assessed using the WISC instrument (3rd or 4th edition) and at similar ages (7–9 years).

To compare our results with the published literature, we looked up in our results the association of cognitive skills and DNA methylation at birth at candidate CpGs based on two previous EWAS studies, the EWAS of cognitive abilities in adulthood [11] and the EWAS of educational attainment [12]. Specifically, from the adult cognitive abilities study we looked up the 2 CpGs that were Bonferroni-significant across all seven cognitive tests and the 43 CpGs that were Bonferroni-significant within specific cognitive tests. From the educational attainment study we looked up the nine CpGs that were significant in adjusted models.

Differential methylation region (DMR) analysis

To assess the joint effect of blood DNA methylation across different sites on cognitive skills we performed a regional analysis. The results of the meta-analysis for each model were analyzed to detect differential methylation regions using the *dmrff* R package [43]. This method combines EWAS summary statistics from nearby CpG sites while taking into account their correlation by using an approach derived from the inverse-variance weighted meta-analysis. Genomic regions were defined as sets of CpGs less than 500 bp apart, nominal p values <0.05 and same sign of effect estimates. This analysis was performed for the main models on overall, verbal and non-verbal scores and for the sensitivity models with further adjustments for paternal education, maternal IQ and genetic PCs.

Brain DNA methylation and gene expression

Correlations between blood and brain methylation at CpG sites within the DMRs were obtained from two online comparison tools based on methylation data from adult participants (<https://epigenetics.essex.ac.uk/bloodbrain/>, date accessed 15-07-2020, and <https://redgar598.shinyapps.io/BECon/>, date accessed 01-12-2020). Since these reference datasets rely on post-mortem brain samples, we could not access data in younger individuals. Brain expression of genes that were in DMRs was examined using the GTEx online portal (<https://gtexportal.org/home/>, date accessed 15-07-2020) and the Braineac online tool (<http://www.braineac.org/>).

RESULTS

Sample characteristics

All cohorts included in the main meta-analysis (Supplementary Table ST1) included a mix of male and female children (average percentage of females across cohorts ranged from 41 to 54%). Testing for childhood cognitive skills was done at a range of average ages spanning from 4 to 9 years (Table 1 and Supplementary Table ST1). Maternal age at delivery was on average between 26 and 33 years overall. The children were born on average at 39–40 weeks of gestation. Maternal smoking during pregnancy varied across cohorts, with the most prevalent self-report in INMA (27%) and EDEN (25%) and the least prevalent in CHAMACOS (5%) and PREDO (4%). Maternal education was coded differently across the cohorts, with a low percentage of highly educated mothers in ALSPAC and CHAMACOS (20–21%) and highest in Generation R and Project Viva (66–78%). The percentage of children with at least one older sibling also varied across cohorts, ranging from 29 to 64%. Sample characteristics were similar for sensitivity models (Supplementary Tables ST2 and ST3). Average paternal education levels were similar to maternal ones within each cohort. Average maternal IQ was only available for 4 cohorts and it was the lowest in CHAMACOS (mean = 85.7) and the highest in Project Viva (mean = 111.4).

Epigenome-wide association study meta-analysis

The results of the meta-EWAS for the main models are plotted in Fig. 1. There was little evidence of an association between cord blood DNA methylation and childhood cognitive skills, either in

terms of overall, verbal, or non-verbal scores. No associations passed the Bonferroni-adjusted p value cut-off of 1.02×10^{-7} . The strongest associations, although not significant when taking into account multiple comparisons, are reported in Table 2. One CpG site, cg00573504 which is located in an intergenic region on chromosome 5, showed a similar association with the overall ($\beta = 3.71$, $p = 4.98E-06$) and the non-verbal cognitive scores ($\beta = 3.38$, $p = 3.32E-06$), whereas the other top sites ($p < 10^{-5}$) were not overlapping across the different cognitive measures. Effect sizes for the top sites were in the range of 0.02–3.6 z-score changes per 10% change in methylation, whereas heterogeneity at those sites between contributing studies was low for most sites and moderate for cg10620273 and cg04783204, with $I^2 < 60$. There was some genomic inflation in the overall score model ($\lambda = 1.11$). When we performed random effect meta-analysis on the top 500 sites, the effect sizes showed very little change and none of them had p values $< 1.02 \times 10^{-7}$ (see Supplementary Table ST4 for the top sites). The meta-regression analysis showed very little evidence of an effect of year of birth. The meta-regression p values for 500 CpGs were all higher than the Bonferroni-corrected threshold $0.05/500 = 10^{-4}$ (see Supplementary Table ST5 for meta-regression results for the top CpG sites identified in the EWAS at $p < 10^{-5}$).

The results of the sensitivity models are summarized in Supplementary Figs. SF1 and SF2 and in Supplementary Tables ST6–ST8. None of these models revealed associations with cognitive skills at p values lower than 1.02×10^{-7} . Regarding the top sites from the main model (Supplementary Table ST9), additionally adjusting for paternal education, maternal IQ or genomic PCs did not substantially change the results.

When EWAS results for each cohort were examined individually across the main models using the same p value cut-off of 1.02×10^{-7} which takes multiple comparisons across the total number of CpGs in the array within each cohort, there was evidence of association of lower non-verbal cognitive skills with DNA methylation at cg26664492 on chromosome 8 (intergenic location) in the ALSPAC cohort ($p = 8.72 \times 10^{-8}$) (Supplementary Fig. SF3). When we performed meta-analysis across ALSPAC, PREDO and CHAMACOS (WISC assessment of cognitive skills at age 7–9) we did not observe associations at p values $< 1.02 \times 10^{-7}$ (see Supplementary Table ST10 for top sites).

Using the results of the main models, we looked up findings from a recent EWAS meta-analysis of adult blood cells DNA methylation and cognitive abilities in adulthood. The two main CpG sites from the EWAS meta-analysis did not replicate in our study (p value > 0.05). Amongst the 43 CpGs that were significant within the different cognitive measures reported in Supplementary Table 3 in the cognitive abilities EWAS study [11], we were able to replicate the association at cg17759224 (intergenic region on chromosome 1) methylation in cord blood with non-verbal scores (nominal $p = 0.00065$ lower than $p = 0.05/34 = 0.001$), out of the 34 sites that were available in our meta-analysis (Supplementary Table ST11). Since cognitive development is strongly associated with educational attainment [6], we also checked if cord blood DNA methylation was associated with childhood cognitive skills at CpG sites previously found to be linked to educational attainment [12]. None of the 9 CpG sites from the educational attainment EWAS had cord blood DNA methylation levels associated with childhood cognitive outcomes in our meta-analysis (Supplementary Table ST12).

DMR analysis

We performed a DMR analysis to investigate the association of DNA methylation at clusters of CpG sites with cognitive skills and found little evidence of clusters in the main models across the cognitive scores (no regions at corrected $p < 0.05$). In the sensitivity model adjusted for maternal IQ (3 cohorts for overall and verbal cognitive skills, 4 cohorts for non-verbal cognitive skills), methylation within a region comprising 5 CpG sites on

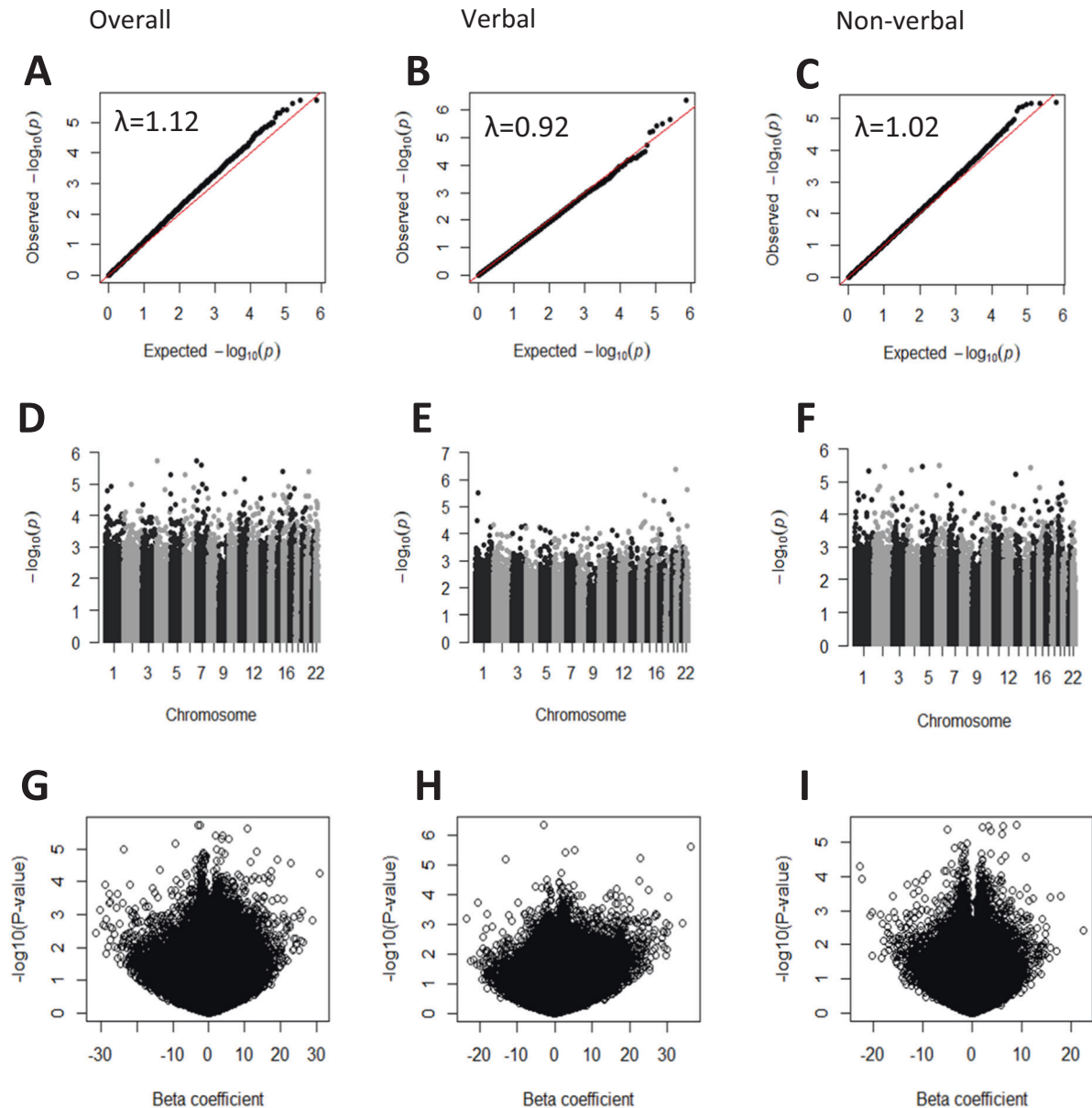


Fig. 1 Results from the meta-analysis of epigenome-wide association studies of cognitive skills in childhood and DNA methylation in cord blood. **A–C** QQ plots showing the observed vs expected probabilities per CpG site and λ index of genomic inflation. **D–F** Manhattan plots showing the $-\log_{10}(p)$ values at each CpG site according to chromosome location. **G–I** Volcano plots showing the effect sizes (difference in IQ z-score units per change from 0 to 1 proportion methylated) and probability values for each CpG site. Models were adjusted for age at testing, sex, maternal age at delivery, maternal education, birth weight, gestational age, maternal smoking status during pregnancy, parity, batch covariates and cell proportions.

chromosome 6 was associated with slightly higher non-verbal scores (DMR adjusted p value = 0.002, Supplementary Table ST13). Even considering the multiple tests across the main model for non-verbal and the three sensitivity models with further adjustments, this p value was lower than a Bonferroni-corrected p value of $0.05/4 = 0.0125$. The heterogeneity at all five sites was low for the maternal IQ model ($I^2 = 0-5$). The positive association between methylation and cognitive skills was similar in the main model and in other sensitivity models, although attenuated, with heterogeneity increased to moderate in some of the sites ($I^2 = 9-49$) (Supplementary Table ST14). A search within two mQTL databases (www.mqtldb.org and mqtldb.godmc.org.uk/) showed only *trans*-mQTLs for these CpG sites, therefore we could not perform any Mendelian randomization analyses to investigate if these associations are causal.

Brain DNA methylation and gene expression

The DMR is located within the *DUSP22* gene, a phosphatase that is expressed across examined tissues including blood and several brain areas (Supplementary Fig. SF4). DNA methylation in blood within this region highly correlates with brain methylation at all CpGs across all brain regions and across the two datasets interrogated, with $r > 0.9$, $p < 10^{-32}$ in the prefrontal cortex, entorhinal cortex, cerebellum and superior temporal gyrus (Table 3), and r values of 0.4–0.8, reaching the 90th percentile, for BA10, BA20 and BA7 (Supplementary Fig. SF5).

DISCUSSION

We have performed the largest EWAS of cognitive skills in cord blood, by running individual EWAS in eight cohorts and

Table 2. Top CpG sites ($p_{\text{meta}} < 10^{-5}$), not significant after Bonferroni correction for multiple comparisons, from the meta-analysis of epigenome-wide association studies of overall, verbal and non-verbal cognitive skills in childhood and DNA methylation in cord blood.

	CpG site	N	Beta ^a	SE	p value ^b	I ^{2c}	Chr.	Position	Gene ^d
Overall	cg05827775	2193	-3.05	0.64	1.84E-06	0	4	9762166	
	cg00213080	2190	-2.44	0.51	1.89E-06	0	7	6204521	CYTH3
	cg26599274	2181	10.83	2.30	2.46E-06	15.1	7	66205733	RABGEF1
	cg23789148	2190	2.03	0.44	3.88E-06	0	15	97321146	
	cg18622281	2188	3.96	0.86	3.91E-06	13.8	20	43977112	SDC4
	cg00573504	2181	3.71	0.81	4.98E-06	0	5	1123809	
	cg09535605	2174	5.57	1.22	5.09E-06	5.1	6	17281327	RBM24
	cg21735491	2190	-9.38	2.09	6.77E-06	0	11	66749665	
Verbal	cg18075761	2152	-23.68	5.36	9.84E-06	0	7	75617028	TMEM120A
	cg03568675	2201	-3.00	0.59	4.36E-07	12.3	20	979279	RSPO4
	cg12361663	2190	36.38	7.70	2.30E-06	0	22	38142561	TRIOBP
	cg17223866	2188	5.21	1.12	3.14E-06	0	1	44794158	ER13
	cg11005998	2199	2.94	0.64	3.87E-06	0	14	101514642	MIR655
	cg10620273	2200	22.72	5.00	5.65E-06	51.0	16	3096488	MMP25
	cg16047144	2200	-13.16	2.92	6.33E-06	0	17	62097953	ICAM2
Non-verbal	cg04783204	3286	8.90	1.91	3.11E-06	53.3	6	44191600	SLC29A1
	cg04229103	3278	6.34	1.36	3.31E-06	0	2	145090268	GTDC1
	cg00573504	3282	3.38	0.73	3.32E-06	0	5	1123809	
	cg25990848	3276	2.07	0.45	3.59E-06	17.8	14	105517573	GPR132
	cg08529049	3254	-4.95	1.08	4.30E-06	25.6	4	48988038	CWH43
	cg03332597	3297	3.83	0.84	4.80E-06	37.3	1	185125704	C1orf25
	cg07489869	3256	5.94	1.31	5.85E-06	1.8	13	20806730	GJB6

^aBeta coefficient from the regression indicating the change in IQ score standard deviations per change in methylation proportion from 0 to 1. Models were adjusted for age at testing, sex, maternal age at delivery, maternal education, birth weight, gestational age, maternal smoking status during pregnancy, parity, batch covariates and cell proportions.

^bUnadjusted p value.

^cHeterogeneity statistics.

^dGene annotation from the Illumina 450 K manifest file.

Table 3. Correlation of DNA methylation levels between brain and blood in adult samples ($N = 71-75, 40-105$ years old) in the *DUSP22* differentially methylated region (<https://epigenetics.essex.ac.uk/bloodbrain/>).

CpG	Prefrontal cortex		Entorhinal cortex		Cerebellum		Superior temporal gyrus	
	r	p value	r	p value	r	p value	r	p value
cg03395511	0.981	3.83e-53	0.965	5.05e-42	0.984	6.30e-57	0.975	5.60e-47
cg07332563	0.946	5.55e-37	0.940	7.93e-34	0.951	4.76e-39	0.927	3.38e-31
cg15383120	0.971	1.67e-46	0.939	9.17e-34	0.961	2.26e-42	0.937	2.59e-33
cg18110333	0.986	6.57e-58	0.981	7.78e-51	0.986	4.74e-58	0.970	4.82e-44
cg21548813	0.986	5.33e-58	0.975	7.80e-47	0.988	9.75e-61	0.986	2.11e-55

combining the results through meta-analysis. We hypothesized that, at birth, we would be able to identify methylation variation that was associated with variation in cognitive skills in childhood. Overall, the evidence at single CpG sites was weak across all models to confirm an association. In one region spanning 5 CpGs on chromosome 6 methylation was positively correlated with small increases in non-verbal scores after adjusting for maternal IQ. This association was revealed only in a subgroup sensitivity analysis (3-4 cohorts instead of 7-8). Cord blood DNA methylation in this region was highly correlated with brain methylation, although we could not identify cis-mQTLs to perform further analyses to establish if methylation in this region is causal to variation in cognitive skills.

The main strength of this study is the large sample size achieved by analyzing data from eight cohorts and combining the results in a meta-analysis allowing to identify only robust results. By using the same protocol and script in all the cohorts we have reduced bias due to heterogeneity.

Our study also has some limitations. Cognitive skills were assessed using different instruments across the cohorts investigated. Although Wechsler scales were used in five of the eight cohorts and despite the evidence of strong correlations between composite measures from cognitive tests [44-46], the different tests used could have contributed to the heterogeneity and reduced precision in our estimates. The covariates also differed across cohorts in the way they were measured, for instance in

terms of maternal education and smoking. There was also some heterogeneity in terms of year of birth across cohorts, with some starting and ending in the 90s, others in the 00s and others across both decades. However, the meta-regression results did not show a strong effect of year of birth. Although the sample size was large, some cohorts could not participate in some of the models, such as Generation R only having data on non-verbal skills and ALSPAC not having maternal IQ data. Our study consisted mostly of participants that identified themselves as white and lived either in Europe or the US. Therefore, the results of this study are not generalizable globally to other ethnic groups and countries. Moreover, the 450 K methylation arrays only capture 2% of the CpGs in the human genome.

Despite the limitations, the lack of association of DNA methylation at single CpGs at birth with cognitive skills in childhood suggests that DNA methylation does not capture prenatal influences (genetic and environmental) on cognitive development, unlike ADHD and social communication development. This denotes a specificity of DNA methylation at birth for certain neurodevelopmental pathways. Our study replicated the association of one of the 43 subthreshold CpG sites identified in a previous EWAS of cognitive skills in adulthood [11], but none of the nine sites found in an EWAS of educational attainment [12]. In those previous studies, both carried out in adulthood, blood cells methylation patterns seemed to be an indication of lifestyle characteristics such as smoking and BMI. Since our study looked for methylation patterns at birth, we would not necessarily expect to see the effect of direct exposures that would occur later in life. However, maternal smoking is highly associated with DNA methylation at birth [28] and since maternal smoking was included as a covariate in the current analyses, we cannot rule out effects of maternal smoking on cognitive skills via DNA methylation at birth. Further studies using formal mediation models should verify this. If our null results are true, our study suggests that the relationship between cognitive skills and blood DNA methylation seen previously are reflective of exposures after birth, rather than in the prenatal period. Another explanation is that blood methylation marks arise from a gene-environment interaction and they appear only later in life, due to the cumulative effect of environmental exposures that are moderated by genetic variation.

It is also possible that the effect of prenatal exposures on childhood cognitive skills are associated with brain DNA methylation patterns that are not captured by cord blood DNA methylation. Largely due to the inaccessibility of brain tissue, most molecular studies of brain-related traits and disorders rely on blood samples, including cord blood in studies of children at birth. Although DNA methylation patterns are often tissue-specific [47], there are strong cross-tissue correlations at specific sites, and studies on blood-brain correspondence allow us to make comparisons that are relevant to brain phenotypes [48–50]. Moreover, despite blood not being the main target tissue for neurodevelopmental conditions, epigenetic associations have been found in peripheral blood e.g., for schizophrenia and autism [51, 52] and correspondence with brain is not a prerequisite for functionally-relevant DNA methylation changes, for example, where immune-related effects on the brain are potential mechanisms. Additionally, peripheral tissue can still serve as a biomarker, when stable signals are identified.

The differentially methylated region on chromosome 6 is located in the *DUSP22* gene, coding for a phosphatase expressed ubiquitously, including in blood and across brain regions. From previously published data [53], there is evidence that, in whole blood, DNA methylation within this region is associated with increased *DUSP22* expression (Supplementary Table ST15). We also observed that DNA methylation correlates highly between peripheral blood and brain tissue across two different datasets

and across brain regions. Although we could only investigate brain-blood correlation in adulthood, we expect some correlation to be present at earlier ages. Moreover, methylation at the *DUSP22* gene in brain tissue has been previously implicated in schizophrenia, Parkinson's and Alzheimer's disease, albeit in different directions depending on the brain region investigated [54–56]. More generally, dual-specific phosphatases, including *DUSP22*, are implicated in a number of neural functions, as shown by several *in vivo* and *in vitro* preclinical studies across species, and in a range of mental and neurological disorders (see for review An et al. [57]). These findings, altogether, suggest that DNA methylation may affect brain functioning via changes in *DUSP22* expression. However, in our study this association was observed in only one of the sensitivity models, with a smaller sample size than the other models and without adjustment for maternal education. Although the association in the other models was in the same direction, the evidence was much weaker, suggesting that this association might be sample-specific or confounded by maternal education. Alternatively, this association could be specific for the effect of maternal IQ on offspring's IQ, which is revealed only when adjusting for maternal IQ. It has been previously shown that maternal IQ affects children cognitive development independently from socioeconomic status and from maternal education [58, 59]. This suggests that maternal IQ could affect the association of methylation at birth with child cognition independently from maternal education and more generally socioeconomic status. Furthermore, this association could be due to residual confounding that was not accounted for in our models (e.g., for aspects of family environment other than socioeconomic status) and needs to be confirmed by replication in independent studies or by using causal inference methods.

In conclusion, we have conducted the largest epigenome-wide scan at birth for cognitive functioning in childhood. Overall, the evidence does not suggest that cord blood DNA methylation at the single CpGs investigated could be an indication of later cognitive skills, either overall, verbal or non-verbal. Most likely, any variation in DNA methylation associated with cognition in peripheral blood arise later in life or are stochastic. Further studies are needed to replicate these results across more ethnically diverse cohorts, in larger samples with more homogenous measurements of cognitive function or in the timing of the cohorts, with data on maternal IQ, and using higher resolution arrays.

DATA AVAILABILITY

Meta-analysis results files will be deposited in the EWAS Catalog data repository (<http://ewascatalog.org/>) upon publication. Individual-level data are available upon request to the cohorts involved and according to their procedures.

REFERENCES

1. Stiles J, Jernigan TL. The basics of brain development. *Neuropsychol Rev.* 2010;20:327–48.
2. Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. *Mol Psychiatry.* 2015;20:98–108.
3. Rask-Andersen M, Karlsson T, Ek WE, Johansson A. Modification of heritability for educational attainment and fluid intelligence by socioeconomic deprivation in the UK biobank. *Am J Psychiatry.* 2021;178:625–34.
4. Tucker-Drob EM, Bates TC. Large cross-national differences in gene x socioeconomic status interaction on intelligence. *Psychol Sci.* 2016;27:138–49.
5. Deary IJ. Intelligence. *Annu Rev Psychol.* 2012;63:453–82.
6. Peet ED, McCoy DC, Danaei G, Ezzati M, Fawzi W, Jarvelin MR, et al. Early childhood development and schooling attainment: longitudinal evidence from British, Finnish and Philippine birth cohorts. *PLoS ONE.* 2015;10:e0137219.
7. Osler M, Andersen AM, Due P, Lund R, Damsgaard MT, Holstein BE. Socio-economic position in early life, birth weight, childhood cognitive function, and adult mortality: a longitudinal study of Danish men born in 1953. *J Epidemiol Community Health.* 2003;57:681–6.

8. van Dongen J, Nivard MG, Willemsen G, Hottenga JJ, Helmer Q, Dolan CV, et al. Genetic and environmental influences interact with age and sex in shaping the human methylome. *Nat Commun.* 2016;7:11115.
9. Joubert BR, Haberg SE, Bell DA, Nilsen RM, Vollset SE, Midttun O, et al. Maternal smoking and DNA methylation in newborns: in utero effect or epigenetic inheritance? *Cancer Epidemiol Biomark Prev.* 2014;23:1007–17.
10. Joubert BR, den Dekker HT, Felix JF, Bohlin J, Ligthart S, Beckett E, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat Commun.* 2016;7:10577.
11. Marioni RE, McAra AF, Bressler J, Colicino E, Hannon E, Li S, et al. Meta-analysis of epigenome-wide association studies of cognitive abilities. *Mol Psychiatry.* 2018;23:2133–44.
12. Karlsson Linner R, Marioni RE, Rietveld CA, Simpkin AJ, Davies NM, Watanabe K, et al. An epigenome-wide association study meta-analysis of educational attainment. *Mol Psychiatry.* 2017;22:1680–90.
13. Neumann A, Walton E, Alemany S, Cecil C, Gonzalez JR, Jima DD, et al. Association between DNA methylation and ADHD symptoms from birth to school age: a prospective meta-analysis. *Transl Psychiatry.* 2020;10:398.
14. Rijlaarsdam J, Cecil CAM, Relton CL, Barker ED. Epigenetic profiling of social communication trajectories and co-occurring mental health problems: a prospective, methylome-wide association study. *Dev Psychopathol.* 2021;1–10. <https://doi.org/10.1017/S0954579420001662>
15. Hannon E, Schendel D, Ladd-Acosta C, Grove J, iPSYCH-Broad ASD Group, Hansen CS, et al. Elevated polygenic burden for autism is associated with differential DNA methylation at birth. *Genome Med.* 2018;10:19.
16. Massrali A, Brunel H, Hannon E, Wong C, iPSYCH-MINERVA Epigenetics Group, Baron-Cohen S, et al. Integrated genetic and methylomic analyses identify shared biology between autism and autistic traits. *Mol Autism.* 2019;10:31.
17. Felix JF, Joubert BR, Baccarelli AA, Sharp GC, Almqvist C, Annesi-Maesano I, et al. Cohort profile: pregnancy and childhood epigenetics (PACE) consortium. *Int J Epidemiol.* 2018;47:22–3.
18. Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort profile: the avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013;42:97–110.
19. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: the ‘children of the 90s’—the index offspring of the avon longitudinal study of parents and children. *Int J Epidemiol.* 2013;42:111–27.
20. Eskenazi B, Bradman A, Gladstone EA, Jaramillo S, Birch K, Holland N. CHAMACOS, a longitudinal birth cohort study: lessons from the fields. *J Children’s Health.* 2003;1:3–27.
21. Eskenazi B, Harley K, Bradman A, Weltzien E, Jewell NP, Barr DB, et al. Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. *Environ Health Perspect.* 2004;112:1116–24.
22. Heude B, Forhan A, Slama R, Douhaud L, Bedel S, Saurel-Cubizolles MJ, et al. Cohort Profile: The EDEN mother-child cohort on the prenatal and early postnatal determinants of child health and development. *Int J Epidemiol.* 2016;45:353–63.
23. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, et al. The Generation R Study: design and cohort update 2017. *Eur J Epidemiol.* 2016;31:1243–64.
24. Guxens M, Ballester F, Espada M, Fernandez MF, Grimalt JO, Ibarluzea J, et al. Cohort profile: the INMA-Infancia y Medio Ambiente-(Environment and Childhood) project. *Int J Epidemiol.* 2012;41:930–40.
25. Girchenko P, Lahti M, Tuovinen S, Savolainen K, Lahti J, Binder EB, et al. Cohort profile: prediction and prevention of preeclampsia and intrauterine growth restriction (PREDO) study. *Int J Epidemiol.* 2017;46:1380–1g.
26. Nieratschker V, Massart R, Gilles M, Luoni A, Suderman MJ, Krumm B, et al. MORC1 exhibits cross-species differential methylation in association with early life stress as well as genome-wide association with MDD. *Transl Psychiatry.* 2014;4:e429.
27. Oken E, Baccarelli AA, Gold DR, Kleinman KP, Litonjua AA, De Meo D, et al. Cohort profile: project viva. *Int J Epidemiol.* 2015;44:37–48.
28. Joubert BR, Felix JF, Yousefi P, Bakulski KM, Just AC, Breton C, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *Am J Hum Genet.* 2016;98:680–96.
29. Kupers LK, Monnerau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. *Nat Commun.* 2019;10:1893.
30. Sammallahti S, Cortes Hidalgo AP, Tuominen S, Malmberg A, Mulder RH, Brunst KJ, et al. Maternal anxiety during pregnancy and newborn epigenome-wide DNA methylation. *Mol Psychiatry.* 2021;26:1832–45.
31. Wechsler D. Wechsler Intelligence Scale for Children—Fourth Edition (WISC-IV) Administration and Scoring Manual. San Antonio, TX: Harcourt Assessment Incorporated; 2003.
32. Wechsler D, Golombok S, Rust J. WISC-IIIUK Wechsler Intelligence Scale for Children—Third Edition UK Manual. Sidcup, UK: The Psychological Corporation; 1992.
33. Wechsler D. WPPSI-III: Administration and scoring manual. San Antonio, TX: The Psychological Corporation; 2002.
34. McCarthy AM, Cool VA, Petersen M, Bruene DA. Cognitive behavioral pain and anxiety interventions in pediatric oncology centers and bone marrow transplant units. *J Pediatr Oncol Nurs.* 1996;13:3–12.
35. Adams W, Sheslow D. Wide range assessment of visual motor abilities. Wilmington, DE: Wide Range, Inc; 1995.
36. Chin CE, Ledesma HM, Cirino PT, Sevcik RA, Morris RD, Frijters JC, et al. Relation between Kaufman Brief Intelligence Test and WISC-III scores of children with RD. *J Learn Disabil.* 2001;34:2–8.
37. Tellegen PJ, Winkel M, Wijnberg-Williams BJ. Snijders-Oomen Niet-verbale Intelligentietest-Revisie SON-R 2½–7. Lisse: Swets & Zeitlinger; 1996.
38. Bakulski KM, Feinberg JI, Andrews SV, Yang J, Brown S, LM S, et al. DNA methylation of cord blood cell types: applications for mixed cell birth studies. *Epigenetics.* 2016;11:354–62.
39. Van der Most PJ, Kupers LK, Snieder H, Nolte I. QCEWAS: automated quality control of results of epigenome-wide association studies. *Bioinformatics.* 2017;33:1243–1245.
40. Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw.* 2010;36:1–48.
41. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet.* 2011;88:586–98.
42. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327:557–60.
43. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. dmrrf: identifying differentially methylated regions efficiently with power and control. *bioRxiv* 2018:508556.
44. Hagmann-von Arx P, Lemola S, Grob A. Does IQ = IQ? Comparability of intelligence test scores in typically developing children. *Assessment.* 2018;25:691–701.
45. Floyd RG, Clark MH, Shadish WR. The exchangeability of IQs: implications for professional psychology. *Professional Psychol Res Pract.* 2008;39:414–23.
46. Faust DS, Oakes Hollingsworth J. Concurrent validation of the wechsler preschool and primary scale of intelligence-revised (WPPSI-R) with two criteria of cognitive abilities. *J Psychoeducational Assess.* 1991;9:224–9.
47. Bakulski KM, Halladay A, Hu VW, Mill J, Fallin MD. Epigenetic research in neuropsychiatric disorders: the “Tissue Issue”. *Curr Behav Neurosci Rep.* 2016;3:264–74.
48. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10:1024–32.
49. Walton E, Hass J, Liu J, Roffman JL, Bernardoni F, Roessner V, et al. Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research. *Schizophr Bull.* 2016;42:406–14.
50. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kober MS. BECON: a tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry.* 2017;7:e1187.
51. Montano C, Taub MA, Jaffe A, Briem E, Feinberg JI, Trygvadottir R, et al. Association of DNA methylation differences with schizophrenia in an epigenome-wide association study. *JAMA Psychiatry.* 2016;73:506–14.
52. Wong CC, Meaburn EL, Ronald A, Price TS, Jeffries AR, Schalkwyk LC, et al. Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioural traits. *Mol Psychiatry.* 2014;19:495–503.
53. Zhernakova DV, Deelen P, Vermaat M, van Iterson M, van Galen M, Arindart W, et al. Identification of context-dependent expression quantitative trait loci in whole blood. *Nat Genet.* 2017;49:139–45.
54. Young JI, Sivasankaran SK, Wang L, Ali A, Mehta A, Davis DA, et al. Genome-wide brain DNA methylation analysis suggests epigenetic reprogramming in Parkinson disease. *Neuro Genet.* 2019;5:e342.
55. Boks MP, Houtepen LC, Xu Z, He Y, Ursini G, Maihofer AX, et al. Genetic vulnerability to DUSP22 promoter hypermethylation is involved in the relation between in utero famine exposure and schizophrenia. *NPJ Schizophr.* 2018;4:16.
56. Sanchez-Mut JV, Aso E, Heyn H, Matsuda T, Bock C, Ferrer I, et al. Promoter hypermethylation of the phosphatase DUSP22 mediates PKA-dependent TAU phosphorylation and CREB activation in Alzheimer’s disease. *Hippocampus.* 2014;24:363–8.
57. An N, Bassil K, Al Jowf GI, Steinbusch HWM, Rothermel M, de Nijs L, et al. Dual-specificity phosphatases in mental and neurological disorders. *Prog Neurobiol.* 2021;198:101906.
58. Tong S, Baghurst P, Vimpani G, McMichael A. Socioeconomic position, maternal IQ, home environment, and cognitive development. *J Pediatrics.* 2007;151:284–8.e281.

59. LeWinn KZ, Bush NR, Batra A, Tylavsky F, Rehkopf D. Identification of modifiable social and behavioral factors associated with childhood cognitive performance. *JAMA Pediatrics*. 2020;174:1063–72.

ACKNOWLEDGEMENTS

We are extremely grateful to all the families who took part in the study, the midwives for their help in recruiting them, interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses, general practitioners, hospitals and pharmacies. In particular, we thank: Dr. S.M. Ring, Dr. W. McArdle, Dr. M. Suderman (ALSPAC); Mr. Michael Verbiest, Ms. Mila Jhamai, Ms. Sarah Higgins, Mr. Marijn Verkerk, Dr. Lisette Stolk and Dr. A.Teumer (Generation R); Silvia Fochs and Nuria Pey (INMA); L. Douhaud, S. Bedel, B. Lortholary, S. Gabriel, M. Rogeon, M. Malinbaum. J.Y. Bernard, J. Botton, M.A. Charles, P. Dargent-Molina, B. de Lauzon-Guillain, P. Ducimetière, B. Foliguet, A. Forhan, X. Fritel, A. Germa, V. Goua, R. Hankard, B. Heude, M. Kaminski, B. Laroque, N. Lelong, J. Lepeule, G. Magnin, L. Marchand, C. Nabet, F. Pierre, R. Slama, M.J. Saurel-Cubizolles, M. Schweitzer, O. Thiebaugeorges (EDEN); E. Hamäläinen, E. Kajantie, H. Laivuori, P.M. Villa, A.-K. Pesonen, A. Aitokallio-Tallberg, A.-M. Henry, V.K. Hiilesmaa, T. Karipohja, R. Meri, S. Sainio, T. Saisto, S. Suomalainen-Konig, V.-M. Ulander, T. Vaitilo L. Keski-Nisula, M.-R. Orden, E. Koistinen, T. Walle, R. Solja, M. Kurkinen, P. Taipale. P. Staven and J. Uotila (PREDO). Please see the Supplementary Material for more detailed acknowledgements.

AUTHOR CONTRIBUTIONS

DCa: study design, ALSPAC analysis, meta-analysis, interpretation of results, manuscript drafting and revision; AN: shadow meta-analysis, Generation R analysis, interpretation of results, critical revision of manuscript; AC: Project Viva analysis and interpretation; GT: CHAMACOS analysis, manuscript review; SA: INMA analysis, manuscript revision; LZ: POSEIDON analysis, manuscript revision; GP: EDEN analysis, manuscript revision; JMTL: provided funding, PREDO data acquisition, quality control, cohort description, reviewed manuscript; AH: contributed to analysis plan and revision of drafts and manuscript; RM: Generation R data acquisition, data analysis, critical revision of the manuscript; JFF: interpretation of results, critical revision of manuscript; HT: Generation R data acquisition, reviewing drafts, provided funding, quality control and supervision of IQ data; LS: data acquisition, manuscript revision; JF: POSEIDON data acquisition, manuscript revision; SHW: POSEIDON data acquisition, manuscript revision; MR: POSEIDON data acquisition, manuscript revision; MD: POSEIDON data acquisition, manuscript revision; KH: POSEIDON manuscript revision; BE: manuscript revision; TS: POSEIDON data acquisition, manuscript revision; MF: INMA data collection, manuscript revision; MG: POSEIDON data acquisition, manuscript revision; MdeA: EDEN cohort data collection; NB: EDEN cohort data analysis; SLR-S: Project Viva analysis and interpretation; TK: data acquisition, PREDO data collection, reviewed manuscript; DCz: PREDO data acquisition and cleaning, manuscript review; CLR: ALSPAC study funding and data acquisition, interpretation of findings; DR: study design and analysis plan; STT: PREDO data analyses, reviewed manuscript; SJL: interpretation and manuscript review; NH: manuscript review; KR: provided funding, planned the PREDO data collection, reviewed manuscript; IA-M: EDEN cohort data collection and funding acquisition; FS: manuscript review; M-FH: Project Viva analysis and interpretation; EO: Project Viva analysis and interpretation; JS: funding, manuscript revision; CAMC: study design, manuscript drafting; GS: study design, manuscript drafting.

FUNDING

The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). This research was specifically funded by the BBSCR (BB1025751/1 and BB1025263/1). GWAS data were generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe. DCa is funded by the MRC (MC_UU_00011/1 and MC_UU_00011/5). GS is financially supported by the MRC [New Investigator Research Grant, MR/S009310/1] and the European Joint Programming Initiative “A Healthy Diet for a Healthy Life” (JPI HDHL, NutriPROGRAM project, UK MRC MR/S036520/1). AH is supported by the South-Eastern Norway Regional Health Authority (2020022) and the Research Council of Norway (274611 and 288083). The POSEIDON work was supported by the German Research Foundation [DFG; grant FOR2107; RI908/11-2 and WI3429/3-2], the German Federal Ministry of Education and Research (BMBF) through the Integrated Network IntegraMent, under the auspices of the e:Med Programme [01ZX1314G; 01ZX1614G] through grants 01EE1406C, 01EE1409C and through ERA-NET NEURON, “SynSchiz—Linking synaptic dysfunction to disease mechanisms in schizophrenia—a

multilevel investigation” [01EW1810], through ERA-NET NEURON “Impact of Early life MetaBolic and psychosocial stress on susceptibility to mental Disorders; from converging epigenetic signatures to novel targets for therapeutic intervention” [01EW1904] and by a grant of the Dietmar-Hopp Foundation. The general design of the Generation R Study is made possible by financial support from Erasmus Medical Center, Rotterdam, Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw) and the Ministry of Health, Welfare and Sport. The EWAS data were funded by a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA; project nr. 050-060-810), by funds from the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by a grant from the National Institute of Child and Human Development (R01HD068437). AN and HT are supported by a grant of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant No. 024.001.003, Consortium on Individual Development). AN is also supported by a Canadian Institutes of Health Research team grant. The work of HT is further supported by a NWO-VICI grant (NWO-ZonMW: 016.VICI.170.200). JFF has received funding from the European Joint Programming Initiative “A Healthy Diet for a Healthy Life” (JPI HDHL, NutriPROGRAM project, ZonMw the Netherlands no. 529051022) and the European Union’s Horizon 2020 research and innovation programme (733206, LifeCycle; 633595, DynaHEALTH). The work of CAMC, JF and RM has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No. 848158 (EarlyCause Project). Main funding of the epigenetic studies in INMA were grants from Instituto de Salud Carlos III (Red INMA G03/176, CB06/02/0041, CP18/00018), Spanish Ministry of Health (FIS-PI04/1436, FIS-PI08/1151 including FEDER funds, FIS-PI11/00610, FIS-FEDER-PI06/0867, FIS-FEDER-PI03-1615) Generalitat de Catalunya-CIRIT 19995GR 00241, Fundació La marató de TV3 (090430), EU Commission (261357-MeDALL: Mechanisms of the Development of ALLergy), and European Research Council (268479-BREATHE: BRain dEvelopment and Air polluTion ultrafine particles in sChool childrEn). ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya. SA is funded by a Juan de la Cierva—Incorporación Postdoctoral Contract awarded by Ministry of Economy, Industry and Competitiveness (IJCI-2017-34068). The Project Viva work was supported by the US National Institutes of Health grants R01 HD034568, UH3 OD023286 and R01 ES031259. The CHAMACOS project was supported by grants from the Environmental Protection Agency [R82670901 and RD83451301], the National Institute of Environmental Health Science (NIEHS) [P01 ES009605, R01ES021369, R01ES023067, R24ES028529, F31ES027751], the National Institute on Drug Abuse (NIDA) [R01DA035300], the National Institutes of Health (NIH) [UG3OD023356] and the National Institute of Mental Health (NIMH) [T32MH112510]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the EPA, NIEHS, or NIH. We thank all funding sources for the EDEN study (not allocated for the present study but for the cohort): Foundation for medical research (FRM), National Agency for Research (ANR), National Institute for Research in Public health (IRESP: TGIR cohorte santé 2008 program), French Ministry of Health (DGS), French Ministry of Research, INSERM Bone and Joint Diseases National Research (PRO-A) and Human Nutrition National Research Programs, Paris-Sud University, Nestlé, French National Institute for Population Health Surveillance (InVS), French National Institute for Health Education (INPES), the European Union FP7 programs (FP7/2007-2013, HELIX, ESCAPE, ENRIECO, Medall projects), Diabetes National Research Program (in collaboration with the French Association of Diabetic Patients (AFD), French Agency for Environmental Health Safety (now ANSES), Mutuelle Générale de l’Education Nationale complementary health insurance (MGEN), French national agency for food security, French speaking association for the study of diabetes and metabolism (ALFEDIAM), grant # 2012/51290-6 Sao Paulo Research Foundation (FAPESP), EU funded MeDALL project. PREDO: The PREDO Study has been funded by the Academy of Finland (JL: 311617 and 269925, KR: 1312670 ja 128789 1287891), EraNet Neuron, EVO (a special state subsidy for health science research), University of Helsinki Research Funds, the Signe and Ane Gyllenberg foundation, the Emil Aaltonen Foundation, the Finnish Medical Foundation, the Jane and Aatos Erkko Foundation, the Novo Nordisk Foundation, the Päivikki and Sakari Sohlberg Foundation, Juho Vainio foundation, Yrjö Jahnsson foundation, Jalmari and Rauha Ahokas foundation, Sigrid Juselius Foundation granted to members of the Predo study board. Methylation assays were funded by the Academy of Finland (269925). SJL is supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences.

COMPETING INTERESTS

MD served as PI in phase II and III studies of Johnson & Johnson, Lilly and Roche. He participated in advisory boards of Johnson & Johnson and received speaker fees from Mundipharma. GT received a Student/New Investigator Travel Award of \$750.00 to attend and present at the 2019 Environmental Mutagenesis and Genomics Society (EMGS) meeting in Washington DC from September 19–23, 2019. All the other co-authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41380-022-01441-w>.

Correspondence and requests for materials should be addressed to Doretta Caramaschi.

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

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



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

© The Author(s) 2022