View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by University of Bath Research Portal

Citation for published version:
Walton, E 2020, 'Association between DNA methylation and ADHD symptoms from birth to school age: A prospective meta-analysis', *Translational Psychiatry*.

Publication date:
2020

Document Version
Peer reviewed version

Link to publication

Publisher Rights
CC BY

University of Bath

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 29. Jul. 2020

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

TITLE PAGE

Association between DNA methylation and ADHD symptoms from birth to school age: A prospective meta-analysis

AUTHORS' NAMES

Alexander Neumann, PhD[1,2,17]*; Esther Walton, PhD[3,4]*; Silvia Alemany, PhD[5,6,7]*; Charlotte Cecil, PhD[1]; Juan Ramon González, PhD[5,6,7]; Dereje Demissie Jima, MSc[8,9]; Jari Lahti, PhD[10,11]; Samuli T. Tuominen, MSc[11] ; Edward D. Barker, PhD[12,13]; Elisabeth Binder, MD, PhD[14,16]; Doretta Caramaschi, PhD[3]; Ángel Carracedo, PhD[15]; Darina Czamara, PhD[16]: Jorunn Evandt, PhD[17]; Janine F. Felix, PhD[18,19]; Bernard F. Fuemmeler, PhD[20,21]; Kristine B. Gutzkow, PhD[22]; Cathrine Hoyo, PhD[23,24]; Jordi Julvez, PhD[5,6,7]; Eero Kajantie, MD, PhD[25,26]; Hannele Laivuori, MD, PhD[27,28]; Rachel Maguire, MPH[23]; Léa Maitre, PhD[5,6,7]; Susan K. Murphy[29]; Mario Murcia, PhD[7,30]; Pia M. Villa, MD, PhD[27]; Gemma Sharp, PhD[13]; Jordi Sunyer, PhD, MD[5,6,7]; Katri Raikkönen, PhD[11]; Marian Bakermans-Kranenburg, PhD[31]; Marinus van IJzendoorn, PhD[32]; Mònica Guxens, MD, PhD[1,5,6,7]; Caroline L. Relton, PhD[3]; Henning Tiemeier, MD, PhD[1,33]

* These authors contributed equally to this work

AUTHORS' AFFILIATIONS

1 Department of Child and Adolescent Psychiatry/Psychology, Erasmus University Medical Center Rotterdam, the Netherlands

2 Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, QC, Canada

3 Medical Research Council Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, UK

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

4 Department of Psychology, University of Bath, UK

5 ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain

6 Universitat Pompeu Fabra (UPF), Barcelona, Spain

7 CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain.

8 Center for Human Health and the Environment, NCSU, USA

9 Bioinformatics Research Center, NCSU, USA

10 Turku Institute for Advanced Studies, University of Turku, Turku, Finland

11 Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Finland

12 Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

13 Centre for Population Neuroscience and Stratified Medicine (PONS), MRC Social, Genetic and Developmental Psychiatry (SGDP) Centre, London, UK

14 Department of Psychiatry and Behavioral Sciences, Emory University School of Medicine, USA

15 Genomic Medicine Group, CIBERER, University of Santiago de Compostela and Galician Foundation of Genomic Medicine (SERGAS)Compostela

16 Dept. Translational Research in Psychiatry, Max-Planck-Institute of Psychiatry, Munich, Germany

17 Department of Air Pollution and Noise, Norwegian Institute of Public Health, Oslo, Norway

18 The Generation R Study Group, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

19 Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

20 Department of Health Behavior and Policy, Virginia Commonwealth University, Richmond, VA, USA

21 Massey Cancer Center, Virginia, Commonwealth Univeristy

22 Department of Molecolar Biology, Norwegian institute of Public Health, Norway

23 Department of Biological Sciences, North Carolina State University

24 Center for Human Health and the Environment, NCSU, USA

25 Chronic Disease Prevention Unit, National Institute for Health and Welfare, Helsinki, Finland

26 Hospital for Children and Adolescents, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland

27 Medical and Clinical Genetics, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

28 Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science, University of Helsinki, Helsinki, Finland

29 Department of Obstetrics and Gynecology, Duke University Medical Center, Durham, North Carolina

30 Joint Research Unit of Epidemiology and Environmental Health, FISABIO–Universitat Jaume I–Universitat de València, Valencia, Spain

31 Clinical Child & Family Studies, Vrije Universiteit Amsterdam, the Netherlands

32 School of Clinical Medicine, University of Cambridge, UK

33 Department of Social and Behavioral Science, Harvard TH Chan School of Public Health, Boston, MA, USA

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

CORRESPONDING AUTHOR

Henning Tiemeier, Erasmus University Medical Center Rotterdam, Department of Child and Adolescent Psychiatry/Psychology, Room KP-2822, PO-Box 2060, 3000 CB Rotterdam, The Netherlands. Tel.: (31) 10-7043489; fax: (31) 10-7044657.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

ABSTRACT

Attention-deficit and hyperactivity disorder (ADHD) is a common childhood disorder with a substantial genetic component. However, the extent to which epigenetic mechanisms play a role in the etiology of the disorder is not known. We performed epigenome-wide association studies (EWAS) within the Pregnancy And Childhood Epigenetics (PACE) Consortium to identify DNA methylation sites associated with ADHD symptoms at two methylation assessment periods: birth and school-age. We examined associations of DNA methylation in cord blood with repeatedly assessed ADHD symptoms (age range 4-15 years) in 2477 children from five cohorts and DNA methylation at school-age with concurrent ADHD symptoms (age 7-11 years) in 2374 children from ten cohorts. The regression estimates at nominal significance of the two EWAS (birth and school-age methylation) correlated (ρ=0.30) between both time points. At birth, we identified nine probes that were associated with later ADHD symptoms at genome-wide significance, including ERC2 and CREB5. Peripheral blood DNA methylation at one of these probes (cg01271805 located in the promotor region of *ERC2*, which regulates neurotransmitter release) was previously associated with brain methylation. Another genome-wide significant probe (cg25520701) lies within the gene body of *CREB5*, which was associated with neurite outgrowth and an ADHD diagnosis in previous studies. In contrast, no probes reached genome-wide significance when ADHD was associated with school-age DNA methylation. In conclusion, DNA methylation at birth is associated with ADHD. Future studies are needed to confirm the utility of methylation variation as biomarker and its involvement in causal pathways.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Introduction

Attention-deficit and hyperactivity disorder (ADHD) is a common neurodevelopmental disorder characterized by impulsivity, excessive activity and attention problems. Symptoms often become apparent during school-age with a world-wide prevalence of 5-7.5%.[1] Genetic heritability is estimated between 64%-88%.[2,3] Additionally several environmental factors are suspected to impact ADHD, e.g. prenatal maternal smoking or lead exposure.[4–7] However, the genetics and environmental pathways that contribute to ADHD risk remain unclear. A possibility is that DNA methylation, an epigenetic mechanism that regulates gene expression, may mediate genetic or environmental effects.

Several studies have investigated DNA methylation in relation to ADHD diagnoses or symptoms using either a candidate approach or epigenome-wide association studies (EWAS) in peripheral blood and saliva tissue.[8,9] A leading hypothesis concerning the etiology of ADHD suggests that deficiencies in the dopamine system of the brain have an impact on ADHD development.[4,10] Consequently, candidate studies have primarily focused on genes related to dopamine function. For instance, DNA methylation alterations in *DRD4*[11–13], *DRD5*[12], and *DAT1*[12,14] genes have been associated with ADHD diagnoses or symptoms, though not consistently[15]. Beyond the candidate gene approach, three studies tested DNA methylation across the whole genome. One study performed an EWAS in school-aged children using a case-control design.[16] The study identified differentially methylated probes in *VIPR2*, a gene expressed in the caudate and previously associated with psychopathology. Another EWAS investigated DNA methylation at birth and at 7 years of age.[17] At birth 13 probes located in *SKI*, *ZNF544*, *ST3GAL3* and *PEX2* were found to be associated with ADHD trajectories from age 7 to 15 years. Interestingly, cross-sectionally

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

the methylation status of these probes at age 7 was not associated with ADHD. An EWAS in adulthood failed to find any differentially methylated sites.[18]

While considerable research has begun to investigate DNA methylation in relation to ADHD, large multi-center epigenome-wide studies, which allow for increased power and generalizability, are lacking in childhood. Here we performed the first epigenome-wide prospective meta-analysis to identify DNA methylation sites associated with childhood ADHD symptoms employing cohorts from the Pregnancy And Childhood Epigenetics (PACE) Consortium[19]. Since the temporal stability of methylation potentially associated with ADHD symptoms is unclear, we tested DNA methylation both at birth using cord blood and in school-age (age 7-9 years) using DNA derived from peripheral whole blood. In the analyses of methylation at birth, the aim was to predict ADHD symptoms between ages 4 and 15 years. We took advantage of the fact that many participating cohorts assessed ADHD repeatedly and employed a repeated measures design to increase precision. Furthermore, we utilized data in childhood to examine cross-sectional DNA methylation patterns associated with ADHD symptoms at school age.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Methods

This study consists of two parts: the birth methylation EWAS and the school-age methylation EWAS described successively below.

Birth Methylation EWAS

Participants

Five cohorts (Avon Longitudinal Study of Parents and Children (ALSPAC),[20–22] Generation R (GENR),[23] INfancia y Medio Ambiente (INMA),[24] Newborn Epigenetic Study (NEST)[25,26] and Prediction and prevention of preeclampsia and intrauterine growth restriction (PREDO)[27]) in the PACE consortium had information on DNA methylation in cord blood and ADHD symptoms. These cohorts have a combined sample size of 2477 (Table 1). Participants were mostly of European ancestry, except for NEST, an American cohort which also included participants of African ancestry. Both ancestries were treated as separate cohorts in the meta-analysis to account for heterogeneity.

DNA Methylation and QC

DNA methylation was measured at birth in cord blood. The Illumina Infinium HumanMethylation450K BeadChip was used to interrogate CpG probes in all cohorts. (Table S1). Methylation levels outside of the lower quartile minus 3*interquartile or upper quartile plus 3*interquartile range were removed. Each cohort ran the EWAS separately and results were meta-analyzed. The distribution of the regression estimates and p-values were examined for each study and pooled results. Deviations from a normal distribution of regression estimates or a higher number of low p-values than expected by chance may be

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

signs of residual confounding, but may also be the result of a true poly-epigenetic signal. To help in interpretation of the results, we used the BACON method.[28] BACON analyzes the distribution of regression coefficients and estimates an empirical null distribution. Results can then be compared against the empirical null, which already includes biases, rather than the theoretical null. We excluded CpG probes, that were available in less than four cohorts, fewer than 1000 participants, and allosomal probes, due to the complex interpretation of dosage compensation.

ADHD Symptoms

ADHD symptoms were measured when children were 4-15 years old (depending on the cohort) with parent-rated instruments, specifically the Behavior Assessment System for Children (BASC),[29] Child Behavior Checklist (CBCL),[30,31] Conners[32] and the Development and Well-Being Assessment (DAWBA)[33] (Table S2). If a cohort had ADHD symptoms measured repeatedly (3 cohorts), every assessment wave was jointly analyzed in a mixed model (see statistical analysis). The repeated measure design increased the precision of the ADHD severity estimate and sample size, since missing data in one or two of the assessments can be handled with maximum likelihood. Given the variety of instruments used within and across cohorts, all ADHD scores were z-score standardized to enable meta-analysis.

Statistical analysis

Cohorts with repeated ADHD assessment were analyzed using linear mixed models, with z-scores of ADHD symptoms as the outcome and methylation (in betas) as the main predictor. Each CpG probe was analyzed separately and adjusted for multiple correction using Bonferroni adjustment. We used a random intercept on the participant and batch level, to account for clustering due to repeated measures and batch effects. The following

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

potential confounders were included as fixed effects: maternal age, educational level, smoking status (yes vs no during pregnancy), gestational age, sex, and estimated white blood cell proportions (Bakulski reference estimated with the Houseman method).[34] Mixed models were fitted using restricted maximum likelihood. We used R[35] with the lme4[36] package to estimate the models. Cohorts with a single ADHD assessment wave used a model without random effects or only on batch level, depending on analyst preference.

Meta-analysis was performed using the Han and Eskin random effects model.[37] This model does not assume that true effects are homogeneous between cohorts, however, it does assume that null effects are homogeneous. This modified version of the random effect model has comparable power to a fixed effects analysis, while better accounting for study heterogeneity, such as ancestry differences, in simulation studies.[37] Genome-wide significance was defined at the Bonferroni-adjustment threshold of p<1E-07, suggestive significance at p<1E-05, and nominal significance at p<0.05.

Follow-up analyses

We performed several look-ups of genome-wide significant probes. We used the BECon database[38] to check the correlation between peripheral and brain methylation levels in post-mortem tissue. To test genetic influence we interrogated the genome-wide significant probes in MeQTL[39] and twin heritability databases.[40] We also attempted to replicate genome-wide significant probes reported in a previous EWAS from the ALSPAC study.[17] For replication we reran the meta-analysis without the ALSPAC cohort. To quantify the variance explained by genome-wide significant probes, we predicted ADHD scores at age 8 in Generation R by all meta-analytically genome-wide significant probes with 10-fold cross-validation and 100 repetitions.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Pathway Analysis

We performed pathway enrichment analysis with the missMethylpackage[41] on suggestive probes (P<1E-05). We used as references: gene ontology (GO), KEGG and curated gene sets (C2; http://software.broadinstitute.org/gsea/msigdb/collections.jsp#C2) from the Broad Institute Molecular signatures database[42]. P-values were adjusted for the number of CpGs associated with each gene[43] and false discovery rate.

To test enrichment for regulatory features (gene relative position, CpG island relative position and blood chromatin states) we applied χ^2 tests. Enrichment tests were performed for all CpGs, and for hypo and hypermethylated CpGs separately. CpG annotation was performed with the IlluminaHumanMethylation450kanno.ilmn-12.hg19 R package.[44] Annotation to 15 chromatin states was retrieved from 27 blood cell types from the Roadmap Epigenomics Project web portal (https://egg2.wustl.edu/roadmap/web_portal/). Each CpG in the array was annotated to one or several chromatin states by defining a state as present in those loci if it was described in at least 1 of 27 cell types. See supplementary information 1 for full description.

School-age methylation EWAS

Participants

Four cohorts (ALSPAC, GENR, HELIX[45] and GLAKU[46]) with a combined sample size of 2374 joined the school-age methylation EWAS (Table 1). HELIX consists of six sub-cohorts, which were pre-processed and analyzed jointly.[45] All cohorts had participants of European ancestry, except HELIX, which also included participants with a Pakistani background living in the UK, which were treated as a separate cohort in the meta-analysis.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Fifty-three percent of participants in the school-age EWAS were also part of the birth EWAS.

DNA Methylation and QC

DNA methylation was measured at ages 7-12 using peripheral whole blood. The Illumina Infinium HumanMethylation450K BeadChip and Infinium MethylationEPIC Kit (GLAKU) were used to interrogate CpG probes. QC steps were identical to the birth methylation EWAS.

ADHD Symptoms

ADHD symptoms were measured at the same age as DNA methylation (age 7-11 years) using the parent-rated measures DAWBA and CBCL (Table S2). In contrast to the birth methylation EWAS only the assessment closest to the DNA methylation assessment age was analyzed.

Statistical analysis

The statistical model was similar to the model used in the birth methylation EWAS. However, cell counts were estimated with the Houseman method using the Reinius reference.[47] Since the outcome was not repeated, there was no random effect on the participant level. We also added assessment age as covariate. The meta-analysis methods were identical to the birth methylation EWAS.

Follow-up analyses

As we observed an overall low signal, we did not perform follow-up analyses. However, we attempted to replicate six probes identified as most suggestive in a previous case-control EWAS in school-age.[16]

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Results

Birth Cord Blood Methylation

EWAS Quality Check

Four out of the six cohorts showed a high λ, indicating larger number of low p-values than expected under the null (Table 1). BACON analysis suggested that the majority of the inflation was due to a true signal, as indicated by inflation values clearly lower than λ. To test the impact of sample size on λ, we restricted the GENR sample randomly to 900 and 1100 participants, resulting 812 and 991 participants due to missing covariates. The lambdas were 0.96, 1.21, 1.51 for 812, 991, and 1191 participants. We thus conclude that the over-representation of low p-values is mostly due to power.

The BACON analyses also indicated a trend towards positive/negative values in some of the datasets, which might indicate confounding, e.g. by population stratification. To test this, we added principal components of ancestry in GENR and ALSPAC, but these did not meaningfully change results.

We conducted the meta-analysis under the assumption that any such biases will be corrected in the pooled analysis, since they were not homogeneous across cohorts. Indeed, the pooled estimates did not show a trend towards positive or negative regression estimates (Median=0.02), only an overrepresentation of low p-values (λ=1.86, see QQ Plot in Figure 1). The BACON estimates for inflation, however, suggested that these are mostly due to a true signal (Inflation=1.1).

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Single Probe Analysis

After QC, 472,817 CpG sites remained for analysis. Results of the meta-analysis of cord blood EWAs are shown in Figure 2. Nine CpG sites showed genome-wide significance ($p<1E-07$, Table 2). ADHD symptoms were between 0.16SD (SE=0.03) and 0.44SD (SE= 0.12) higher with 10% lower methylation at these probes. Eight probes out of nine that were available in the BECon database[38] are typically methylated in both whole blood as well as in the brain (Figures 3, S1 and S2). A lookup in the BECon database revealed that the CpG site cg01271805 in the promoter region of gene *ERC2* shows variable methylation in three brain regions (BA10, BA20, BA7). Importantly, methylation levels in the brain are moderately correlated with whole blood methylation ($\rho=0.33-0.46$) (Figure 3), suggesting that peripheral cg01271805 methylation levels could be a useful marker for brain methylation levels. The other seven genome-wide significant probes showed less consistent correlations between blood and brain tissues and associated genes had less specificity for expression in the brain, based on GTEx[48] data. No SNP was associated with our nine top CpG probes when accounting for linkage disequilibrium according to the MeQTL database[39]. Furthermore, all nine probes had a twin heritability below 20% in a previous study (Table S3).[40] After adjusting for inflation and bias with BACON, only one CpG remained statistically significant (cg25520701, CREB5, ß =-3.54, SE = 0.66, p = 9.59E-08). It should be noted, that the BACON adjusted p-values rely on statistics from the traditional random effects model. With the traditional model, only cg25520701, cg09762907 and cg22997238 remained genome-wide significant. Thus the difference in p-value is not solely the result of adjustment for the inflation, but also the use of more conservative tests. In Generation R, the joint explained variance of ADHD scores at age 8 by the genome-wide significant probes was 1.8% (R^2 from 10-fold repeated cross-validation).

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Pathway Analysis

Two-hundred forty-nine probes showed suggestive (P<1E-05) associations and were annotated to 182 unique genes. In gene-based analyses no pathway survived multiple testing correction. See supplementary materials for full results.

Suggestive CpGs were enriched in intergenic regions. Of these, hypomethylated probes were enriched for 3'UTR regions and depleted for TSS200 and first exon regions, open sea, north shelf and south shelf regions, south shore and islands. Regarding chromatin states, hypomethylated probes showed an enrichment for transcription (Tx and TxWk), quiescent positions and depletion for transcription start site positions (TSSA, TxFlnk, TxFlnk), bivalent (EnhBiv) and repressor (ReprPC) positions. Hypermethylated probes showed the opposite chromatine state patterns. See supplementary information 1 for full results.

Replication of previous EWAS

We attempted to replicate findings for 13 CpGs, at which DNA methylation at birth was associated with ADHD trajectories.[17] However, no probe survived multiple-testing correction with incosistent effect directions. (Table S4).

School-age methylation

EWAS Quality Checks

The beta regression distribution showed no signs of errors, but three out of the five cohorts showed a trend towards positive associations in separate analyses (Table 1). The lambda was below 1.11 for all cohorts. BACON suggested no inflation of the test statistics due to confounding or other biases, though the trend towards positive associations remained. The pooled results showed a low lambda (0.96), no inflation (0.92) but a slight over-representation of positive associations (0.14).

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Single Probe Meta-Analysis

We associated DNA methylation at school-age in whole-blood at 466,574 CpG sites with ADHD symptoms at the same age. No CpG reached genome-wide significance (all $p>4.96E-06$, Figure 2). Furthermore, none of the loci at which DNA methylation at birth was significantly associated with ADHD symptoms, also showed a cross-sectional association at school-age ($p>0.33$).

Replication of previous EWAS

We attempted to replicate the six most suggestive EWAS probes of a previous case-control study.[16] While all but one showed a consistent direction, none of the probes were statistically significant. (Table S5)

Stability of methylation association across age

The associations between methylation at birth with ADHD symptoms and methylation at school-age with ADHD symptoms were largely consistent for nominally significant probes. The regression estimates from those CpG sites, which had nominally significant associations at birth ($p<0.05$, $n=73{,}057$) correlated with the regression estimates of the school-age EWAS ($\rho=0.45$). When restricting the school-age methylation EWAS to those cohorts, which were not featured in the birth methylation EWAS (thus excluding overlaps), the correlation remained ($\rho=0.30$). Vice versa, when filtering for probes which were nominally significant at school-age, 23,770 probes remained of which 4075 overlapped with nominally significant probes at birth. The correlation for this set was very similar, $\rho=0.47$ among all cohorts and $\rho=0.35$ between independent cohorts (47%).

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Discussion

In this population-based study, we performed the first epigenome-wide meta-analysis of ADHD symptoms in childhood, using two DNA methylation assessments (birth and school-age), as well as repeated measures of ADHD symptoms. DNA methylation at birth, but not at school-age, was associated with later development of ADHD symptoms with genome-wide significance at nine loci. Interestingly all the identified probes showed a pattern of a high average rate of methylation in cord blood, where lower levels of methylation in an individual were associated with more ADHD symptoms in childhood. DNA methylation in cord blood reflects the effects of genetics and the intrauterine environment. The results thus suggest that cord blood DNA methylation is a marker for some of the ADHD risk factors present before birth or functions as a potential mediator of these risk factors. While not impossible, reverse causality at this age is unlikely to explain our results, as ADHD only manifests at a later stage of development.

We analyzed DNA methylation in cord and peripheral blood, which may not correspond to the methylation status in the brain. While DNA methylation may affect behavior via other pathways, DNA methylation in the brain arguably has the strongest a priori likelihood of representing causal mechanisms. Seven out of eight significant probes did not show consistent correlation between methylation status in whole blood and post-mortem brain tissue in a previous study, i.e. DNA methylation levels in blood may not represent brain levels and thus associations with ADHD may be different.[38] However, one of the eight probes (cg01271805) is an exception in that methylation levels in whole blood are associated with methylation levels in various brain regions. Importantly, this probe lies in the promoter region of the gene *ERC2*, that is highly expressed in brain tissue. *ERC2* regulates calcium dependent neurotransmitter release in the axonal terminal.[50] Specifically, *ERC2* is suspected to increase the sensitivity of voltage dependent calcium channels to

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

hyperpolarization, resulting in higher neurotransmitter release. SNPs in the *ERC2* locus were previously shown to distinguish schizophrenia and bipolar disorder patients[51] and have been suggested to impact cognitive functioning[52]. *ERC2* is especially expressed in Broadmann area 9 of the frontal cortex.[48] Previous imaging studies have demonstrated differential activation in this area when children with or without ADHD performed various cognitive tasks.[53,54] The correlation with brain methylation, the location in a promoter and gene expression in the brain make cg01271805 a plausible candidate locus, where reduced methylation may be mechanistically involved in ADHD development. We hypothesize, that lower methylation levels at cg01271805 increases the expression of *ERC2*, which in turn increases neurotransmitter release, with an adverse impact on the development of ADHD symptoms. Another gene with a genome-wide significant probe and high relevance for neural functioning is *CREB5* (cg25520701). *CREB5* is expressed in fetal brain and the prefrontal cortex, and has been previously related to neurite outgrowth. Moreover, SNPs in this gene were associated with ADHD in two recent GWAS.[55,56] Thus, it is plausible that differences in DNA methylation at this locus may modify ADHD risk during development.

While the birth methylation EWAS identified several loci, associating school-age methylation with concurrent ADHD symptoms revealed no genome-wide significant associations. Furthermore, the overall association signal was lower, despite similar sample sizes. None of the probes, which were significantly associated at birth showed any association when measured at school-age. Given that sample sizes were comparable, this difference must come from changes in the epigenome or study heterogeneity, rather than differences in statistical power. In terms of instrument heterogeneity, the school-age EWAS was more homogeneous, almost exclusively using CBCL. Additionally, as both EWAS feature a mix of several cohorts selected based on the same criteria and around half of the

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

participants were represented at both time points, study heterogeneity appears to be an unlikely explanation. The stronger signal in the birth EWAS may be considered surprising given that typically two measures are typically more strongly associated if measured in closer temporal proximity. However, in line with our results Walton et al. also observed in a previous EWAS,[17] that birth methylation may be a better predictor of later ADHD symptoms than childhood methylation, possibly reflecting sensitive periods. This may also explain why a previous EWAS using DNA methylation in adulthood did not find associations with adult ADHD symptoms, despite a much larger sample size compared to this EWAS.[18] Whether DNA methylation in cord blood has stronger causal effects or is a better marker for early life factors cannot be concluded from the present study. Alternatively, tissue differences between cord blood and whole blood may account for the differences in association pattern. Finally, it is possible that interventions in childhood and other environmental influences reduced the initial epigenetic differences at birth between children with higher and lower ADHD symptoms.

That said, we observed consistency in the associations of methylation at both timepoints with ADHD symptoms. The regression estimates of both EWAS correlated on a genome-wide level. This held true, even when cohorts featured in both EWAS (3 cohorts) were removed from the analysis suggesting that the association between DNA methylation at birth and ADHD symptoms remain in school-age to some extent and is consistent across independent cohorts.

Strengths of this study include the large sample size, repeated outcome measures, extensive control for potential confounders and the use of DNA methylation at two different time-points, enabling us to characterize both prospective and cross-sectional associations with ADHD symptoms. However, several limitations need to be discussed as well. A causal interpretation of our findings is challenged by the possibility of residual confounding and

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

reverse causality. DNA methylation might be a marker for untested adverse environmental factors that could affect ADHD via independent pathways. In addition, children with higher ADHD symptoms may evoke a particular environment, which might shape the epigenome and be a substantial factor for the cross-sectional analyses in school-age. It is also likely that many more CpG sites are associated with ADHD than identified in this study, given the necessity for stringent multiple testing correction. Larger sample sizes are necessary to detect further methylation sites. As is typical for (epi-)genetic studies, the effect size of individual top probes was rather small: the joint effect of the genome-wide probes was estimated below 2%. However, the strong genome-wide epigenetic signal suggests a potential for the development of epigenetic-scores based on birth methylation, which could lead to early prevention efforts before ADHD symptoms arise.

In summary, we identified nine CpG sites for which lower methylation status at birth is associated with later development of ADHD symptoms. The results suggest that DNA methylation in *ERC2* and *CREB5* may exert an influence on ADHD symptoms, potentially via modification of neurotransmitter functioning or neurite outgrowth.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Acknowledgments

We thank all the children and families who took part in this study, as well as the support of hospitals, midwives and pharmacies.

ALSPAC

We are grateful to the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (MRC) and Wellcome (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors and E Walton will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf). Methylation data in the ALSPAC cohort were generated as part of the UK BBSRC funded (BB/I025751/1 and BB/I025263/1) Accessible Resource for Integrated Epigenomic Studies (ARIES, http://www.ariesepigenomics.org.uk). Edward Barker is supported by an ESRC grant ES/S013229/1. Doretta Caramaschi, Gemma Sharp and Caroline L. Relton work in a department supported by the UK MRC (MC_UU_00011/5). Gemma Sharp is further supported by an 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, MRC, MR/S036520/1).

GENR

The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the Erasmus University Rotterdam, Faculty of Social Sciences, the Municipal Health Service Rotterdam area, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of general practitioners, hospitals, midwives and pharmacies in Rotterdam. The generation and management of the Illumina 450K methylation array data (EWAS data) for the Generation R Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We thank Mr. Michael Verbiest, Ms. Mila Jhamai, Ms. Sarah Higgins, Mr. Marijn Verkerk and Dr. Lisette Stolk for their help in creating the EWAS database. We thank Dr. A.Teumer for his work on the quality control and normalization scripts.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 was 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). A. Neumann and H. Tiemeier 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). A. Neumann is also supoerted by a Canadian Institutes of Health Research team grant. The

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

work of H. Tiemeier is further supported by a NWO-VICI grant (NWO-ZonMW: 016.VICI.170.200). M.H. van IJzendoorn and M.J. Bakermans-Kranenburg were supported by the Netherlands Organization for Scientific Research (SPINOZA, VICI), and M.J. Bakermans-Kranenburg was supported by the European Research Council (AdG 669249). J.F.F. 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). This project received funding from the European Union’s Horizon 2020 research and innovation programme (733206, LifeCycle; 633595, DynaHEALTH.

GLAKU

We thank all the research nurses, research assistants, and laboratory personnel involved in the GLAKU study. The study has been supported by Academy of Finland, University of Helsinki, Hope and Optimism Initiative, Finnish Foundation for Pediatric Research, Sigrid Juselius Foundation, Jalmari and Rauha Ahokas Foundation, Signe and Ane Gyllenberg Foundation, Yrjo Jahnsson Foundation, Juho Vainio Foundation, Emil Aaltonen Foundation, and Ministry of Education and Culture, Finland.

INMA

A full roster of the INMA Project Investigators can be found at http://www.proyectoinma.org/presentacion-inma/listado-investigadores/en_listado-investigadores.html. Silvia Alemany is funded by a Juan de la Cierva-Incorporación fellowship (IJCI-2017-34068) awarded by the Spanish Ministerio de Economía, Industria y Competitividad. Mònica Guxens is funded by a Miguel Servet fellowship (MS13/00054, CP18/00018) awarded by the Spanish Institute of Health Carlos III. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya. Jordi Julvez is funded by a Miguel Servet fellowship (MS14/00108, CP14/00108) awarded by the Spanish Institute of Health Carlos III.

HELIX

The authors would like to thank allpractitioners and researchers in the six countries who took part in this study. The authors would like to thank Sonia Brishoual, Angelique Serre and Michele Grosdenier (Poitiers Biobank, CRB BB-0033-00068, Poitiers, France) for biological sample management and Professor Frederic Millot (Principal Investigator), Elodie Migault, Manuela Boue and Sandy Bertin (Clinical Investigation Center, Inserm CIC1402, CHU de Poitiers, Poitiers, France) for planning and investigational actions. The authors would like to thank Veronique Ferrand-Rigalleau, Céline Leger and Noella Gorry (CHU de Poitiers, Poitiers, France) for administrative assistance (EDEN). The authors would like to thank Silvia Fochs, Nuria Pey, Cecilia Persavente and Susana Gross for field work, sample management and overall management in INMA. The authors would like to thank Georgia Chalkiadaki and Danai Feida for biological sample management, to Eirini Michalaki, Mariza Kampouri, Anny Kyriklaki and Minas Iakovidis for field study performance and to Maria Fasoulaki for administrative assistance (RHEA). The authors would also like to thank Ingvild Essén and Jorunn Evandt for thorough field work, Heidi Marie Nordheim for biological sample management and the MoBa administrative unit (MoBa). The contribution of the Spanish National Gneotyping Center (CEGEN-PRB2) is also acknowledged.

NEST

The NEST cohort has been supported by the National Institute of Environmental Health Sciences (R01ES016772 [CH], R21ES014947 [CH], P30ES025128 [CH, DDJ, RM], R01MD011746 [CH, SKM], P30ES011961 pilot project [SKM], P01ES022831 [SKM, BFF]), the US Environmental Protection Agency (RD-83543701 [SKM, BFF]), the National

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

Institute of Diabetes and Digestive and Kidney Diseases (R01DK085173 [CH, SKM]), the National Institute of Aging (R21AG041048 [BFF]) and the Duke Cancer Institute [CH, SKM]. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health or the United States Environmental Protection Agency (US EPA). Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication.

PREDO

The PREDO Study has been funded by the Academy of Finland, 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, the Sigrid Juselius Foundation granted to members of the Predo study board. Methylation assays were funded by the Academy of Finland. The PREDO study would not have been possible without the dedicated contribution of the PREDO study group members: A-K Pesonen (Department of Psychology and Logopedics, University of Helsinki; Finland), A Aitokallio-Tallberg, A-M Henry, VK Hiilesmaa, T Karipohja, R Meri, S Sainio, T Saisto, S Suomalainen-Konig, V-M Ulander, T Vaitilo (Department of Obstetrics and Gynaecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland), L Keski-Nisula, Maija-Riitta Orden (Kuopio University Hospital, Kuopio Finland), E Koistinen, T Walle, R Solja (Northern Karelia Central Hospital, Joensuu, Finland), M Kurkinen (Päijät-Häme Central Hospital, Lahti, Finland), P.Taipale. P Staven (Iisalmi Hospital, Iisalmi, Finland), J Uotila (Tampere University Hospital, Tampere, Finland). We also thank all the research nurses, research assistants, and laboratory personnel involved in the PREDO study.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

References

1 Thomas R, Sanders S, Doust J, Beller E, Glasziou P. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics* 2015; **135**: e994-1001.

2 Bergen SE, Gardner CO, Kendler KS. Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood. *Twin Res Hum Genet* 2007; **10**: 423–433.

3 Larsson H, Chang Z, D'Onofrio BM, Lichtenstein P. The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan. *Psychol Med* 2014; **44**: 2223–2229.

4 Swanson JM, Kinsbourne M, Nigg J, Lanphear B, Stefanatos GA, Volkow N *et al.* Etiologic Subtypes of Attention-Deficit/Hyperactivity Disorder: Brain Imaging, Molecular Genetic and Environmental Factors and the Dopamine Hypothesis. *Neuropsychol Rev* 2007; **17**: 39–59.

5 Marceau K, Cinnamon Bidwell L, Karoly HC, Evans AS, Todorov AA, Palmer RH *et al.* Within-Family Effects of Smoking during Pregnancy on ADHD: the Importance of Phenotype. *J Abnorm Child Psychol* 2018; **46**: 685–699.

6 Daneshparvar M, Mostafavi SA, Jeddi MZ, Yunesian M, Mesdaghinia A, Mahvi AH *et al.* The role of lead exposure on Attention-Deficit/Hyperactivity Disorder in children: A systematic review. *Iran J Psychiatry* 2016; **11**: 1–14.

7 Lam J, Lanphear BP, Bellinger D, Axelrad DA, McPartland J, Sutton P *et al.* Developmental pbde exposure and IQ/ADHD in childhood: A systematic review and meta-analysis. *Environ Health Perspect* 2017; **125**. doi:10.1289/EHP1632.

8 Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *J Child Psychol Psychiatry* 2017; **4**: 303–322.

9 Dall'Aglio L, Muka T, Cecil CAM, Bramer WM, Verbiest MMPJ, Nano J *et al.* The Role of Epigenetic Modifications in Neurodevelopmental Disorders: A Systematic Review. *Neurosci Biobehav Rev* 2018; **94**: 17–30.

10 Pappa I, Mileva-Seitz VR, Szekely E, Verhulst FC, Bakermans-Kranenburg MJ, Jaddoe VWV *et al.* DRD4 VNTRs, observed stranger fear in preschoolers and later ADHD symptoms. *Psychiatry Res* 2014; **220**: 982–986.

11 van Mil NH, Steegers-Theunissen RPM, Bouwland-Both MI, Verbiest MMPJ, Rijlaarsdam J, Hofman A *et al.* DNA methylation profiles at birth and child ADHD symptoms. *J Psychiatr Res* 2014; **49**: 51–9.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

12 Xu Y, Chen X-T, Luo M, Tang Y, Zhang G, Wu D *et al.* Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *J Psychiatr Res* 2015; **64**: 40–50.

13 Dadds MR, Schollar-Root O, Lenroot R, Moul C, Hawes DJ. Epigenetic regulation of the DRD4 gene and dimensions of attention-deficit/hyperactivity disorder in children. *Eur Child Adolesc Psychiatry* 2016; **25**: 1081–1089.

14 Adriani W, Romano E, Pucci M, Pascale E, Cerniglia L, Cimino S *et al.* Potential for diagnosis versus therapy monitoring of attention deficit hyperactivity disorder: a new epigenetic biomarker interacting with both genotype and auto-immunity. *Eur Child Adolesc Psychiatry* 2018; **27**: 241–252.

15 Ding K, Yang J, Reynolds GP, Chen B, Shao J, Liu R *et al.* DAT1 methylation is associated with methylphenidate response on oppositional and hyperactive-impulsive symptoms in children and adolescents with ADHD. *World J Biol Psychiatry* 2017; **18**: 291–299.

16 Wilmot B, Fry R, Smeester L, Musser ED, Mill J, Nigg JT. Methylomic analysis of salivary DNA in childhood ADHD identifies altered DNA methylation in VIPR2. *J Child Psychol Psychiatry Allied Discip* 2016; **57**: 152–160.

17 Walton E, Pingault J-B, Cecil CAM, Gaunt TR, Relton CL, Mill J *et al.* Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Mol Psychiatry* 2017; **22**: 250–256.

18 van Dongen J, Zilhão NR, Sugden K, Heijmans BT, ’t Hoen PAC, van Meurs J *et al.* Epigenome-wide Association Study of Attention-Deficit/Hyperactivity Disorder Symptoms in Adults. *Biol Psychiatry* 2019; : 1–9.

19 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**: 22-23u.

20 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**: 111–127.

21 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* 2012; **42**: 97–110.

22 Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H *et al.* Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int J Epidemiol* 2015; **44**: 1181–1190.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

23 Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IJzendoorn MH *et al.* The Generation R Study: design and cohort update 2017. *Eur J Epidemiol* 2016; **31**: 1243–1264.

24 Guxens M, Ballester F, Espada M, Fernández MF, Grimalt JO, Ibarluzea J *et al.* Cohort profile: The INMA-INfancia y Medio Ambiente-(environment and childhood) project. *Int J Epidemiol* 2012; **41**: 930–940.

25 Hoyo C, Murtha AP, Schildkraut JM, Jirtle R, Demark-Wahnefried W, Forman MR *et al.* Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. *Epigenetics* 2011; **6**: 928–936.

26 Liu Y, Murphy SK, Murtha AP, Fuemmeler BF, Schildkraut J, Huang Z *et al.* Depression in pregnancy, infant birth weight and DNA methylation of imprint regulatory elements. *Epigenetics* 2012; **7**: 735–746.

27 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**: 1380–1381.

28 Maarten van Iterson, Erik van Zwet , the BIOS Consortium PES and BT, Heijmans. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *bioRxiv* 2016; **38**: 1–16.

29 Reynolds CR. Behavior assessment system for children. *Corsini Encycl Psychol* 2010; : 1–2.

30 Achenbach TM, Rescorla LA. *Manual for the ASEBA Preschool Forms and Profiles.* Burlington, VT, 2000.

31 Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms and Profiles.* . Burlington, VT, 2001.

32 Conners CK, Sitarenios G, Parker JDA, Epstein JN. The revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity. *J Abnorm Child Psychol* 1998; **26**: 257–268.

33 Goodman R, Ford T, Richards H, Gatward R, Meltzer H. The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *J child Psychol psychiatry* 2000; **41**: 645–655.

34 Bakulski KM, Feinberg JI, Andrews S V., Yang J, Brown S, L. McKenney S *et al.* DNA methylation of cord blood cell types: Applications for mixed cell birth studies. *Epigenetics* 2016; **11**: 354–362.

35 R Core Team. R: A Language and Environment for Statistical Computing. 2016.https://www.r-project.org/.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

36 Bates D, Mächler M, Bolker BM, Walker SC. Fitting Linear Mixed-Effects Models using lme4. *J Stat Softw* 2014; : 1–51.

37 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**: 586–598.

38 Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: A tool for interpreting DNA methylation findings from blood in the context of brain. *Transl Psychiatry* 2017; **7**: e1187-10.

39 Smith AK, Kilaru V, Kocak M, Almli LM, Mercer KB, Ressler KJ *et al.* Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics* 2014; **15**. doi:10.1186/1471-2164-15-145.

40 Hannon E, Knox O, Sugden K, Burrage J, Wong CCY, Belsky DW *et al.* Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genet* 2018; **14**: 1–27.

41 Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* 2015; **32**: 286–288.

42 Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011; **27**: 1739–1740.

43 Geeleher P, Hartnett L, Egan LJ, Golden A, Raja Ali RA, Seoighe C. Gene-set analysis is severely biased when applied to genome-wide methylation data. *Bioinformatics* 2013; **29**: 1851–1857.

44 IlluminaHumanMethylation450kanno HK. ilmn12. hg19: Annotation for Illumina's 450k methylation arrays. *R Packag version 02* 2014; **1**.

45 Maitre L, de Bont J, Casas M, Robinson O, Aasvang GM, Agier L *et al.* Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* 2018; **8**: e021311.

46 Strandberg TE, Järvenpää AL, Vanhanen H, McKeigue PM. Birth outcome in relation to licorice consumption during pregnancy. *Am J Epidemiol* 2001; **153**: 1085–1088.

47 Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; **13**. doi:10.1186/1471-2105-13-86.

48 GTEx-Portal. GTEx Portal. 2017.https://www.gtexportal.org/home/ (accessed 11 May2017).

49 Wu S, Hivert MF, Cardenas A, Zhong J, Rifas-Shiman SL, Agha G *et al.* Exposure to low levels of lead in utero and umbilical cord blood DNA methylation in project viva: An epigenome-wide association study. *Environ Health Perspect* 2017; **125**: 1–10.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

50 Kiyonaka S, Nakajima H, Takada Y, Hida Y, Yoshioka T, Hagiwara A *et al.* Physical and functional interaction of the active zone protein CAST/ERC2 and the β-subunit of the voltage-dependent Ca2+ channel. *J Biochem* 2012; **152**: 149–159.

51 Curtis D, Vine AE, McQuillin A, Bass NJ, Pereira A, Kandaswamy R *et al.* Case-case genome-wide association analysis shows markers differentially associated with schizophrenia and bipolar disorder and implicates calcium channel genes. *Psychiatr Genet* 2011; **21**: 1–4.

52 Hatzimanolis A, Bhatnagar P, Moes A, Wang R, Roussos P, Bitsios P *et al.* Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood. *Am J Med Genet Part B Neuropsychiatr Genet* 2015; **168**: 392–401.

bioRxiv preprint doi: https://doi.org/10.1101/806844. this version posted October 16, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. It is made available under a CC-BY 4.0 International license.

53 Epstein JN, Delbello MP, Adler CM, Altaye M, Kramer M, Mills NP *et al.* Differential patterns of brain activation over time in adolescents with and without attention deficit hyperactivity disorder (ADHD) during performance of a sustained attention task. *Neuropediatrics* 2009; **40**: 1–5.

54 Monden Y, Dan H, Nagashima M, Dan I, Tsuzuki D, Kyutoku Y *et al.* Right prefrontal activation as a neuro-functional biomarker for monitoring acute effects of methylphenidate in ADHD children: An fNIRS study. *NeuroImage Clin* 2012; **1**: 131–140.

55 Franke B, Neale BM, Faraone S V. Genome-wide association studies in ADHD. *Hum Genet* 2009; **126**: 13–50.

56 Klein M, Walters RK, Demontis D, Stein JL, Hibar DP, Adams HH *et al.* Genetic markers of ADHD-related variations in intracranial volume. *Am J Psychiatry* 2019; **176**: 228–238.

57 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**: 680–696.

58 Gustavson K, Ystrom E, Stoltenberg C, Susser E, Magnus P, Knudsen GP *et al.* Smoking in Pregnancy and Child ADHD. 2017; **139**. doi:10.1542/peds.2016-2509.

Tables

Table 1: Cohort characteristics of birth methylation EWAS

						Standardized regression coefficients			BACON estimates		
Cohort	Ancestry/ Ethnicity	n	Methylation Age	ADHD Age	Instrument	33%	50%	66%	λ	Inflation	Bias
Birth EWAS											
ALSPAC	European	714	0	8, 11, 14, 15	DAWBA	-0.21	0.25	0.89	1.60	1.10	0.37
GENR	European	1191	0	6,8,10	CBCL (6,10), Conners (8)	-0.48	0.01	0.53	1.51	1.20	0.05
INMA	European	325	0	7,9	Conners (7), CBCL (9)	-1.37	-0.40	0.43	0.80	0.87	-0.19
NEST	Black	55	0	5	BASC	-3.50	-0.03	3.63	1.16	1.10	0.00
NEST	White	56	0	5	BASC	-2.54	-0.09	2.36	0.80	0.92	-0.01
PREDO	European	136	0	5	Conners	-1.55	-0.25	1.20	1.45	0.95	0.21
META	-	2477	-	-	-	-0.37	0.02	0.42	1.86	1.10	0.01
School-age EWAS											
ALSPAC	European	651	7	8	DAWBA	-0.61	-0.10	0.54	1.09	1.00	-0.08
GENR	European	395	10	10	CBCL	-0.93	-0.00	0.98	1.00	0.97	-0.01
GLAKU	European	215	12	12	CBCL	-0.79	0.31	1.50	0.92	0.96	0.13
HELIX	European	1034	8	8	CBCL	-0.26	0.47	1.40	1.11	0.98	0.28
HELIX	Pakistani	79	7	7	CBCL	-1.66	1.86	5.48	0.98	0.96	0.26
Meta	-	2374	-	-	-	-0.24	0.14	0.62	0.96	0.92	0.14

n Number of participants

33%, 50%, 66% Quartiles of regression coefficient distribution

λ Inflation of p-values

Inflation Inflation of p-values due to suspected bias

Bias Trend toward negative/positive distribution of regression coefficients due to suspected bias

Table 2: EWAS Results

				Birth methylation					School-age methylation				
CpG	Gene	Chr	Position	$n_{studies}$	n	B	SE	p	$n_{studies}$	n	B	SE	p
cg25520701	CREB5	7	28800657	6	2450	-3.53	0.60	4.95E-09	5	2279	-0.13	1.09	0.94
cg24838839	Intergenic	5	61031569	6	2468	-4.15	1.79	3.95E-08	5	2287	1.52	1.38	0.33
cg22997238	Intergenic	7	36014218	6	2465	-1.63	0.30	8.81E-08	5	2291	-0.06	0.47	0.94
cg21600027	Intergenic	4	124443502	6	2464	-3.04	0.81	2.64E-08	5	2281	0.98	0.89	0.33
cg17876201	ZBTB38	3	141139991	6	2457	-4.41	1.20	7.58E-09	4	2066	0.56	1.32	0.73
cg11251614	PPIL1	6	36839846	6	2451	-3.43	0.68	3.89E-08	5	2276	0.77	1.52	0.68
cg09762907	TRERF1	6	42290256	6	2460	-2.11	0.39	8.76E-08	5	2284	-0.55	0.64	0.46
cg09158638	Intergenic	16	62309996	6	2470	-2.55	1.40	1.89E-08	5	2270	-0.33	1.04	0.80
cg01271805	ERC2	3	55694954	6	2469	-2.86	1.71	5.24E-08	5	2289	0.28	0.73	0.76

Chr Chromosome
$n_{studies}$ Number of studies
n Number of participants
B Regression coefficient
SE Standard error

Figure 1

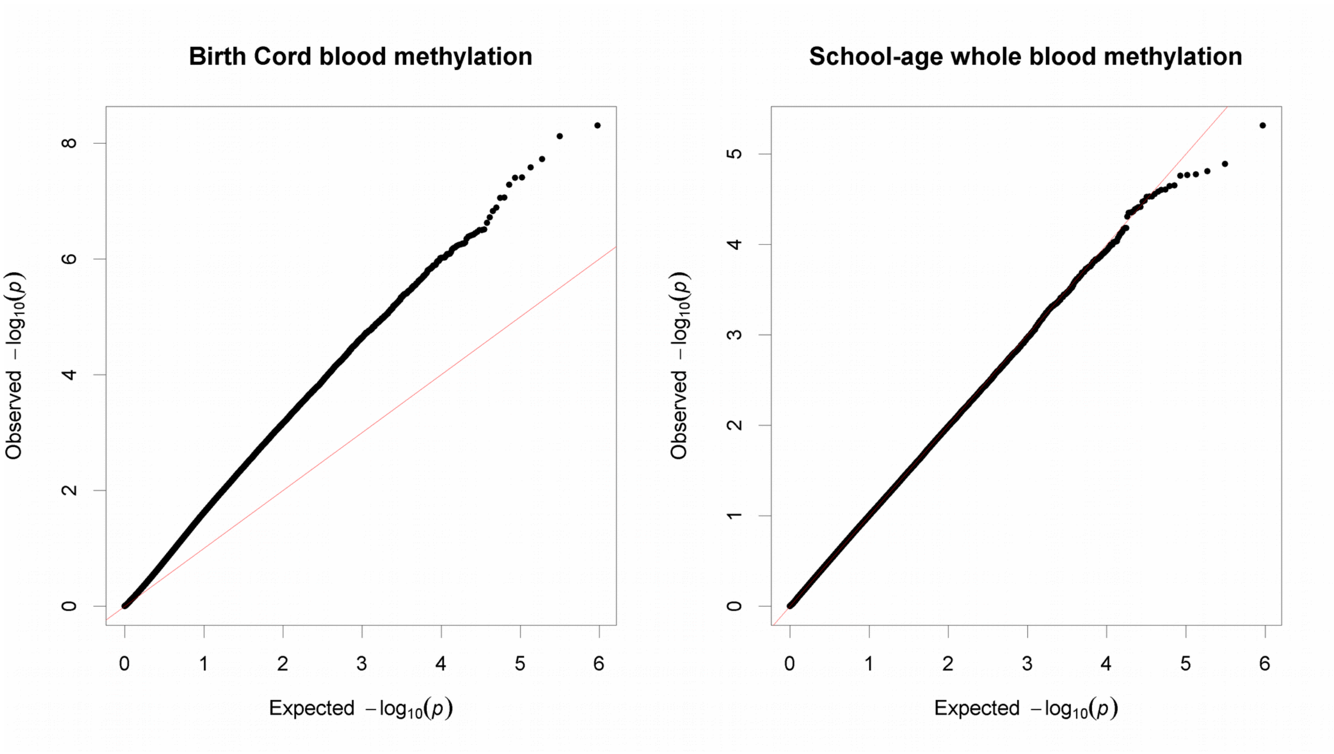


Fig. 1: Quantile-quantile plot of observed −log10 p-values in the cord blood and school-age EWAS vs expected −log10 p-values under assumption of chance findings only. The diagonal line represents the distribution of the expected p values under the null. Points above the diagonal indicate p-values which are lower than expected.

Figure 2

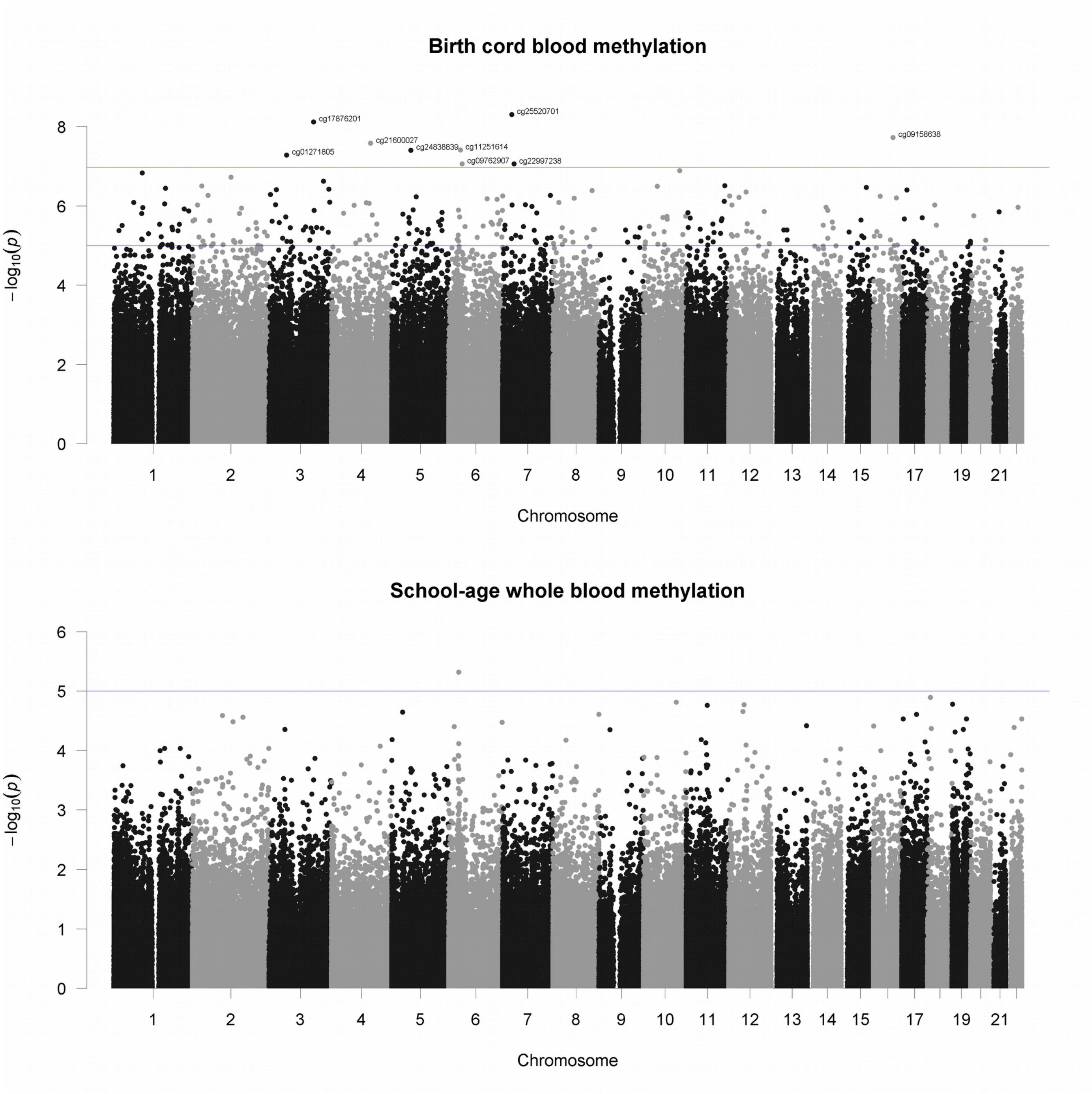
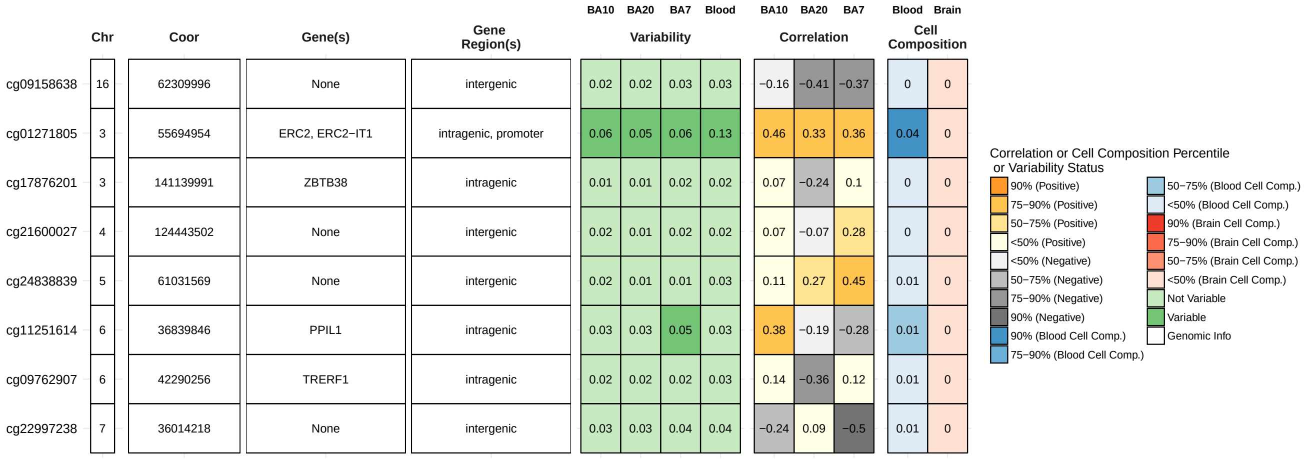


Fig. 2.: Manhattan plot of −log10 p values vs SNP position (basepair and chromosome). Red line indicates genome-wide significance and blue line suggestive level

Figure 3



	Chr	Coor	Gene(s)	Gene Region(s)	Variability				Correlation			Cell Composition	
					BA10	BA20	BA7	Blood	BA10	BA20	BA7	Blood	Brain
cg09158638	16	62309996	None	intergenic	0.02	0.02	0.03	0.03	−0.16	−0.41	−0.37	0	0
cg01271805	3	55694954	ERC2, ERC2−IT1	intragenic, promoter	0.06	0.05	0.06	0.13	0.46	0.33	0.36	0.04	0
cg17876201	3	141139991	ZBTB38	intragenic	0.01	0.01	0.02	0.02	0.07	−0.24	0.1	0	0
cg21600027	4	124443502	None	intergenic	0.02	0.01	0.02	0.02	0.07	−0.07	0.28	0	0
cg24838839	5	61031569	None	intergenic	0.02	0.01	0.01	0.03	0.11	0.27	0.45	0.01	0
cg11251614	6	36839846	PPIL1	intragenic	0.03	0.03	0.05	0.03	0.38	−0.19	−0.28	0.01	0
cg09762907	6	42290256	TRERF1	intragenic	0.02	0.02	0.02	0.03	0.14	−0.36	0.12	0.01	0
cg22997238	7	36014218	None	intergenic	0.03	0.03	0.04	0.04	−0.24	0.09	−0.5	0.01	0

Correlation or Cell Composition Percentile or Variability Status

- 90% (Positive)
- 75−90% (Positive)
- 50−75% (Positive)
- <50% (Positive)
- <50% (Negative)
- 50−75% (Negative)
- 75−90% (Negative)
- 90% (Negative)
- 90% (Blood Cell Comp.)
- 75−90% (Blood Cell Comp.)
- 50−75% (Blood Cell Comp.)
- <50% (Blood Cell Comp.)
- 90% (Brain Cell Comp.)
- 75−90% (Brain Cell Comp.)
- 50−75% (Brain Cell Comp.)
- <50% (Brain Cell Comp.)
- Not Variable
- Variable
- Genomic Info

Fig. 3.: Lookup of brain-blood correlations and variability of genome-wide significant CpG sites in the BECon database