Accelerated epigenetic age at birth and child emotional and behavioral development in early childhood: a meta-analysis of four prospective cohort studies in ECHO

Ashley Y. Song  
Johns Hopkins Bloomberg School of Public Health

Catherine M. Bulka  
University of South Florida

Sierra S. Niemiec  
University of Colorado Anschutz Medical Campus

Katerina Kechris  
University of Colorado Anschutz Medical Campus

Kristen E. Boyle  
University of Colorado Anschutz Medical Campus

Carmen J. Marsit  
Emory University

T. Michael O'Shea  
University of North Carolina School of Medicine

Rebecca C. Fry  
University of North Carolina at Chapel Hill

Kristen Lyall  
A.J. Drexel Autism Institute

M. Daniele Fallin  
Emory University

Heather E. Volk  
Johns Hopkins Bloomberg School of Public Health

Christine Ladd-Acosta  
Johns Hopkins Bloomberg School of Public Health

Short Report

Keywords: epigenetic age, biologic age, age acceleration, Child Behavior Checklist, emotional and behavioral traits, child neurodevelopment, DNA methylation
Abstract

Background: “Epigenetic clocks” have been developed to accurately predict chronologic gestational age and have been associated with child health outcomes in prior work.

Methods: We meta-analyzed results from four prospective U.S cohorts investigating the association between epigenetic age acceleration estimated using blood DNA methylation collected at birth and preschool age Childhood Behavior Checklist (CBCL) scores.

Results: Epigenetic aging was not significantly associated with CBCL total problem scores ($\beta = 0.33$, 95% CI: -0.95, 0.28) and DSM-oriented pervasive development problem scores ($\beta = -0.23$, 95% CI: -0.61, 0.15). No associations were observed for other DSM-oriented subscales.

Conclusions: These findings may relate to our sample size or sample characteristics; future work should address the role of epigenetic age in child health in other study populations.

Introduction

Over 8 million children ages 3–17 years in the US were diagnosed with a mental or behavioral health condition in 2018–2019 [1]. Studies have shown that early identification and intervention is associated with better long-term health outcomes among affected individuals [2]. However, most neurodevelopmental disorders are not diagnosed until children reach 3–5 years of age at the earliest and typically diagnoses occur in mid-childhood. There is a critical need for biomarkers at the earliest stages of life to improve the identification of children at risk.

In recent years, biologic aging as measured using “epigenetic clocks” or DNA methylation (DNAm)-based age estimators has been robustly associated with prospective overall mortality and disease-specific risk in adults [3, 4], showing promise as early biomarkers of prospective health outcomes. The difference between chronologic age and predicted epigenetic age captures the age discordance of the individual. Over the past few years, epigenetic clocks robustly predicting gestational age [5–7] have been developed and studied in association with newborn and child health outcomes. Despite gestational age being associated with developmental delays in children, only two European studies have examined the relationship between epigenetic aging and cognitive and behavioral traits during childhood [8, 9]. One study reported that epigenetic age deceleration was associated with increased internalizing problems and overall problems measured by the Child Behavior Checklist (CBCL) at age 4 in boys in a Finnish cohort [9], while the second study observed an increased odds of psychiatric problems with epigenetic age acceleration in a different cohort of older Finnish children [8]. To our knowledge, no studies have evaluated the relationship between epigenetic age at birth and child emotional and behavioral traits in early childhood in a U.S population nor have studies considered CBCL subscale scores that may be more relevant to Diagnostic and Statistical Manual of Mental Disorders (DSM) mental disorder diagnoses.
In this study, we conducted a meta-analysis investigating the association between epigenetic age acceleration at birth and later DSM-oriented CBCL subscales and total score in a large U.S population from four prospective cohorts participating in the National Institutes of Health Environmental Influences on Child Health Outcomes (ECHO) program.

**Methods**

Existing prospective cohorts in the United States participated the ECHO program were invited to participate in the current study if they had the following information before April 2021 when this study was conducted 1) DNA methylation data measured in cord or peripheral blood collected at birth from singleton infants 2) preschool age CBCL data measured. Four cohorts contributed to this meta-analysis of epigenetic age at birth and CBCL outcomes in early childhood: Early Autism Risk Longitudinal Investigation (EARLI), Markers of Autism Risk in Babies – Learning Early Signs (MARBLES), Extremely Low Gestational Age Newborns (ELGAN), and the Healthy Start Study. The EARLI and MARBLES studies are prospective studies of autism spectrum disorder (ASD) that utilizes a familial history design, i.e. it enrolled pregnant women who had a previous child with ASD, while the ELGAN study recruited women who gave birth before 28 weeks gestational age. Detailed methods for each cohort are provided in Supplementary Material.

Umbilical cord blood or peripheral blood biosamples were collected shortly after delivery using standardized protocol at each study sites. DNA was bisulfite converted using an EZ-96 DNA Each cohort measured DNA methylation using the Infinium HumanMethylation450 BeadChip array or Infinium HumanMethylationEPIC BeadChip array (Illumina, San Diego, CA) in cohort-specific laboratories. Quality control of samples was performed by each cohort and failed samples were excluded based on high detection p value, low sample DNA concentration, sample call rate, CpG-specific percentage of missing values, bisulfite conversion efficiency, sex verification, and other quality control metrics specific to cohorts. The data were then further processed for background correction and normalization in each cohort. Cell type proportions were estimated using a combined cord blood reference data set in each cohort. More details on quality control, normalization, and cell type proportions are provided in the Supplemental Materials.

Each cohort derived epigenetic age at birth according to a pre-specified analysis plan independently using algorithms published by Knight [6]. The Knight method used an elastic net approach with 10-fold cross-validation in the training set and identified a model consisting of 148 CpG sites that can estimate epigenetic gestational age. The estimation of epigenetic age using the Knight clock was performed using R statistical software with codes supplied by Knight et al. Due to the cohort-specific QC filters applied to DNA methylation data, some CpG sites required to compute the Knight clock may have been missing. If a required CpG for the clock computation was filtered out, methylation values were either imputed using simple random sampling imputation or substituted from the closest CpG sites (less than 2000 base pairs). Epigenetic age acceleration was then derived as the residual from a linear model of epigenetic age regressed on gestational age adjusting for cell-type proportions.
In each cohort, the primary caregiver of the child or an interviewer completed the CBCL 1.5-5 questionnaire at the age of 2–3 years. Respondents choose one of three responses to the 99 characteristics listed in the CBCL: “not true,” “somewhat or sometimes true,” and “very true or often true. The normalized T-scores, including total problem score and DSM-oriented subscale scores (attention-deficit hyperactivity disorder problems, anxiety problems, oppositional defiant problems, pervasive development problems), were calculated, and analyzed as the primary outcomes, as these are more clinically interpretable than raw scores.

Each cohort completed independent analysis to evaluate the relationship between epigenetic age acceleration at birth and CBCL outcomes in early childhood according to a pre-specified analysis plan. Descriptive analyses were conducted to examine the maternal and child characteristics and the distribution of epigenetic age/epigenetic age acceleration. Spearman correlation coefficients between chronologic gestational age, estimated epigenetic age, and epigenetic age acceleration/deceleration across Knight clock were calculated. Adjusted multivariable linear regression models were completed to evaluate the association between continuous epigenetic age acceleration at birth and continuous CBCL outcomes, adjusted for ancestry principal components (if available), maternal education, prenatal smoking exposure, maternal pre-pregnancy body mass index (BMI), and child sex. If ancestry principal components were not available, self-reported race/ethnicity was used. Missings across covariates in each cohort was low (<2%). Subjects with missing covariates were excluded from the cohort-specific analysis. Binary CBCL outcomes based on clinical relevant cut-offs were also examined individually for each of the CBCL outcomes (≥ 65 vs < 65 for total problem T-score and ≥ 60 vs < 60 for DSM-oriented subscales) using adjusted logistic regression models.

Random-effects meta-analyses were conducted using the results from cohort-specific analyses, which accounts for between- and within-study variance, allowing for the possibility that individual studies estimated different effect sizes. Forest plots were generated to display each study’s contribution to each summary estimate. Potential heterogeneity between studies was assessed using the Cochran Q statistic and $I^2$. Additionally, leave-one-out and cumulative meta-analyses were conducted to evaluate the contribution of each study to the results. All meta-analyses were implemented in R 3.6 (R Project for Statistical Computing).

Results

Characteristics of participating studies are shown in Table S1. Due to the inclusion of a preterm birth cohort and cohorts enriched for a family history of autism, the mean (SD) gestational age ranged from 26.2 (1.2) weeks to 39.5 (1.2). DNAm estimates of epigenetic age, derived using the Knight clock algorithm for blood samples, showed significant positive overall correlations with chronologic gestational age (Figure S1, ranged from 0.38 to 0.62, $P< 0.001$ for all cohorts).

The summary estimates of the association between epigenetic age acceleration at birth and CBCL total problem scores and DSM-oriented pervasive development problem scores were $-0.33$ (95% CI: -0.95,
0.28) and −0.23 (95% CI: -0.61, 0.15), respectively (Fig. 1), based on all four studies. Between-study heterogeneity among all studies was not statistically significant (P = 0.81). Effect estimates were similar in directions and magnitudes across the leave-one-out and cumulative meta-analyses. We did not observe any significant associations between epigenetic age acceleration at birth and other DSM-oriented subscale of CBCL scores (Figure S2) nor when stratified by child sex. Sensitivity analyses examining binary CBCL outcomes did not yield materially changed results from the primary meta-analysis results.

Discussion

In this meta-analysis, we observed that epigenetic gestational age acceleration was not associated with measures of child emotional and behavioral development overall or stratified by child sex in participants from 4 U.S cohorts. Our results show that DNAm based epigenetic age accurately predicted chronologic gestational age in all four birth cohorts with diverse race and ethnicity, suggesting that epigenetic clocks can be a valuable molecular biomarker in assessing biologic age at birth.

Previous work has suggested greater total or internalizing problems with decelerated epigenetic age at birth in boys but not in girls or in combined samples [9]. We did not observe significant associations overall or by child sex in the current meta-analyses. It is possible that we are not powered to detect smaller magnitude changes due to our relatively small sample size and sample characteristics. Our point estimates for total score, oppositional defiant problems, and pervasive development problems association in the full study population are in the same direction as the European study and consistently indicate an inverse relationship between epigenetic age acceleration and child emotional and behavior problems. Results from one of the familial enriched studies showed suggestive associations for DSM-oriented subscales, suggesting those with familial ASD risk may be impacted in multiple domains and biological aging as measured by epigenetic age may be relevant to child neurodevelopment. Due to our sample size, we were not permitted to run meta-analysis among familial enriched studies alone. To our knowledge, this is the first study examining the association between epigenetic age acceleration and emotional and behavioral development in early childhood in a non-European population. Our study included participants from four well-characterized prospective cohorts with diverse racial and ethnic backgrounds drawn from both high-risk and general populations. Whether mechanisms underlying these relationships differ by genetic background or by familial risk remains unclear. Understanding the impact of genetic factors on biologic age and its potential impact on child cognitive and behavioral development in general population stratified by low or high familial risk should be an area of future research.

Differences in socio-economic status (SES) between our study population and previous studies may contribute to the mixed findings. For example, the association between epigenetic aging and child emotional and behavioral development may be modified by SES-related factors such as neighborhood deprivation. It is also possible that nutrition or maternal history of depression or other parental mental health conditions might influence child neurodevelopment and be associated with epigenetic aging at birth, for which we were unable to control in this study. Future studies with larger sample size would benefit from incorporating social factors related to environmental exposures and maternal/paternal
mental health. In addition, we cannot rule out limited generalizability of our findings, which may be limited to those with increased familial likelihood of ASD or high risk of developmental delays.

Here, we examined cell-intrinsic aging by adjusting for proportions of estimated blood cell types, as age-related health changes may shift blood composition resulting in differences in measures of epigenetic aging. We did not explore second generation epigenetic clocks that incorporate lifestyle or clinical factors. Epigenetic clocks in tissues relevant to early life exposures and epigenetic clocks sensitive to health outcomes associated with clinical markers have been developed to accurately capture the biological process of aging and age-related epigenetic shift. Extrinsic or environmental factors could have influenced age-related shifts in cord blood cells and be related to child development. Future research utilizing these second generation epigenetic clocks will be informative.

In summary, in this meta-analysis including four cohorts participating in the ECHO program, epigenetic age acceleration at birth was not associated with child emotional and behavioral problems measured by CBCL scores in early childhood. However, the null results should be interpreted in the context of our study population, which includes two cohorts enriched for ASD and one preterm birth cohort. The heterogeneity of results across diverse cohorts suggest the need for further investigations to examine the associations with other standardized child cognitive and behavior developmental measures in both high risk and general populations and in larger samples that permit examination of sex differences and the additional role of environmental and genetic factors in these relationships.

Declarations

Ethnics approval and consent to participate

The institutional review boards (IRB) at organizations in each study site approved each study. All participants provided written informed consent.

Consent for publications

All authors have reviewed this manuscript and approved its publication.

Availability of data and materials

Data used for these analyses may be available through approval from individual study investigators and the Environmental Influences on Child Health Outcomes (ECHO) program.

Competing interests

The authors have no conflicts of interest relevant to this manuscript to disclose.

NIH Funding Acknowledgment and Disclaimer
The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Research reported in this publication was supported by the Environmental influences on Child Health Outcomes (ECHO) program, Office of the Director, National Institutes of Health, under Award Numbers U2COD023375 (Coordinating Center), U24OD023382 (Data Analysis Center), U24OD023319 with co-funding from the Office of Behavioral and Social Science Research (PRO Core), UH3OD023248 (Dabelea), UH3OD023342 (Lyall), UH3OD023348 (O'Shea).

Authors’ contribution

AS, HV, and CL-A conceptualized and designed the project. AS, CB, and SN performed the cohort-specific analysis. AS performed meta-analysis. AS, HV, and CL-A drafted the manuscript. CB, KK, KB, CM, MO, RF, KL, and DF provided the manuscript revisions and editing. All authors reviewed the results and approved the submitted version of the manuscript.

ECHO Collaborators Acknowledgment

The authors wish to thank our ECHO colleagues; the medical, nursing, and program staff; and the children and families participating in the ECHO cohorts. We also acknowledge the contribution of the following ECHO program collaborators:

ECHO Components—Coordinating Center: Duke Clinical Research Institute, Durham, North Carolina: Smith PB, Newby KL; Data Analysis Center: Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland: Jacobson LP; Research Triangle Institute, Durham, North Carolina: Catellier DJ; Person-Reported Outcomes Core: Northwestern University, Evanston, Illinois: Gershon R, Cella D.

ECHO Awardees and Cohorts—University of Colorado Denver, Denver, CO: Dabelea D; University of California, UC Davis Medical Center Mind Institute, Davis, CA: Schmidt R; Kaiser Permanente Northern California Division of Research: Croen L; Baystate Children's Hospital, Springfield, MA: Vaidya R; Beaumont Health Medical Center, Royal Oak, MI: Obeid R; Boston Children's Hospital, Boston, MA: Rollins C; East Carolina University Brody School of Medicine, Greenville, NC: Bear K; Helen DeVos Children's Hospital, Grand Rapids, MI: Pastymak S; Michigan State University College of Human Medicine, East Lansing, MI: Lenski, M; University of Chicago, Chicago IL: Msall M; University of Massachusetts Medical School, Worcester, MA: Frazier J; Wake Forest Baptist Health (Atrium Health), Winston Salem, NC: Washburn, L; Yale School of Medicine, New Haven, CT: Montgomery A.

References


**Supplementary Files**

Figures S1 and S2 and Table S1 are not available with this version

**Figures**
Figure 1

Results of random effects meta-analyses of the associations between epigenetic age acceleration at birth estimated by Knight clock and childhood behavior checklist A) total-T score B) pervasive development problems assessed in early childhood. Results of individual studies were from multivariable linear regression models, which adjusted for child sex, maternal education, maternal smoking during pregnancy, and maternal pre-pregnancy BMI status. Point estimates of individual study beta estimates are shown as black squares; area of squares indicates relative weights of individual studies; point and interval overall ratio estimates are horizontal apices of black diamonds.