

UC Berkeley

UC Berkeley Previously Published Works

Title

Prenatal and birth associations of epigenetic gestational age acceleration in the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) cohort

Permalink

<https://escholarship.org/uc/item/7rz8c2tv>

Journal

Epigenetics, 17(13)

ISSN

1559-2294

Authors

Daredia, Saher

Huen, Karen

Van Der Laan, Lars

et al.

Publication Date

2022-12-09

DOI



10.1080/15592294.2022.2102846

Peer reviewed

RESEARCH PAPER



Prenatal and birth associations of epigenetic gestational age acceleration in the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) cohort

Saher Daredia ^a, Karen Huen^{b,c}, Lars Van Der Laan^b, Philip A. Collender^b, Jamaji C. Nwanaji-Enwerem^b, Kim Harley^{c,d}, Julianna Deardorff^{c,d}, Brenda Eskenazi^{a,b,c}, Nina Holland^{b,c}, and Andres Cardenas ^{a,b,c,e}

^aDivision of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, CA, USA; ^bDivision of Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA, USA; ^cCenter for Environmental Research and Children's Health (CERCH), School of Public Health, University of California, Berkeley, CA, USA; ^dDivision of Community Health Sciences, School of Public Health, University of California, Berkeley, CA, USA; ^eCenter for Computational Biology, University of California, Berkeley, CA, USA

ABSTRACT

Gestational age (GA) is an important determinant of child health and disease risk. Two epigenetic GA clocks have been developed using DNA methylation (DNAm) patterns in cord blood. We investigate the accuracy of GA clocks and determinants of epigenetic GA acceleration (GAA), a biomarker of biological ageing. We hypothesize that prenatal and birth characteristics are associated with altered GAA, thereby disrupting foetal biological ageing.

We examined 372 mother-child pairs from the Center for the Health Assessment of Mothers and Children of Salinas study of primarily Latino farmworkers in California. Chronological GA was robustly correlated with epigenetic GA (DNAm GA) estimated by the Knight ($r = 0.48$, $p < 2.2 \times 10^{-16}$) and Bohlin clocks ($r = 0.67$, $p < 2.2 \times 10^{-16}$) using the Illumina 450K array in cord blood samples collected at birth. GA clock performance was robust, though slightly lower, using DNAm profiles from the Illumina EPIC array in a smaller subsample (Knight: $r = 0.39$, $p < 3.5 \times 10^{-5}$; Bohlin: $r = 0.60$, $p < 7.7 \times 10^{-12}$). After adjusting for confounders, high maternal serum triglyceride levels (Bohlin: $\beta = -0.01$ days per mg/dL, $p = 0.03$), high maternal serum lipid levels (Bohlin: $\beta = -4.31 \times 10^{-3}$ days per mg/dL, $p = 0.04$), preterm delivery (Bohlin: $\beta = -4.03$ days, $p = 9.64 \times 10^{-4}$), greater maternal parity (Knight: $\beta = -4.07$ days, $p = 0.01$; Bohlin: $\beta = -2.43$ days, $p = 0.01$), and male infant sex (Knight: $\beta = -3.15$ days, $p = 3.10 \times 10^{-3}$) were associated with decreased GAA.

Prenatal and birth characteristics affect GAA in newborns. Understanding factors that accelerate or delay biological ageing at birth may identify early-life targets for disease prevention and improve ageing across the life-course. Future research should test the impact of altered GAA on the long-term burden of age-related diseases.

ARTICLE HISTORY

Received 8 February 2022
Revised 23 Jun 2022
Accepted 11 July 2022



KEYWORDS

Epigenetic clocks; epigenetic age; DNA methylation age; cord blood; gestational age; age acceleration

Background

Ageing is a complex process that has been associated with physiological changes at molecular, cellular, and tissue levels [1]. While the causes of ageing are not fully understood, it is established that both genetic and environmental factors have an impact on this process [1]. Over the past decade, a number of studies have shown that epigenetic modifications can serve as biomarkers of ageing [2,3]. Differences and changes in methylation of cytosine nucleotides, one of the most widely studied epigenetic marks, have been strongly correlated with longevity and age-related diseases, such as cancer and cardiovascular disease [4,5].

Because the epigenome is influenced by the environment over time, epigenetic marks have the potential to predict 'biological aging' of an individual's cells, tissues, and organ systems as opposed to their chronological age, which is based on birthdate [6]. In recent years, epigenetic clocks have been established as highly accurate DNA methylation-based biomarkers of biological ageing, outperforming traditional markers such as telomere length [7]. Several epigenetic clocks have been developed, each trained using methylation levels of different sets of individual cytosines, or CpG sites in the genome, to predict the biological age of a specific DNA source [2]. The divergence between epigenetic age (DNAm age) and chronological age,

CONTACT Andres Cardenas  andres.cardenas@berkeley.edu  Division of Epidemiology and Biostatistics, School of Public Health, University of California, 2121 Berkeley Way, Room 5302, Berkeley, CA 94720, USA

known as epigenetic age acceleration (EAA), serves as an indicator of biological ageing associated with all-cause mortality and morbidities like cancer and obesity [2].

Infants might be particularly sensitive to biological ageing at birth. For example, differences in gestational age, as small as 1 week, can significantly impact neonatal morbidity and mortality, as well as long-term outcomes related to cognitive ability and respiratory health [8–10]. Yet, studies of epigenetic age acceleration in infants are limited. Adult studies have shown accelerated epigenetic ageing to be associated with poor health outcomes [2]. However, due to the importance of proper foetal development and maturity, decelerated epigenetic ageing at birth may be as detrimental to later disease burden as age acceleration at birth.

Two epigenetic clocks developed by Knight et al. [11] and Bohlin et al. [12] were constructed to predict gestational age at delivery using DNA methylation profiles of CpG sites measured from foetal umbilical cord blood and newborn blood spots. Using methylation levels from 148 CpG sites analysed on the Illumina Infinium HumanMethylation27 and HumanMethylation450 (450K) arrays, Knight's clock reported a high predictive accuracy between DNAm GA and chronological GA ($r = 0.91$) [11]. Using a more reduced model with 96 CpG sites analysed on the 450K array, Bohlin's clock also resulted in estimates that were positively correlated with chronological GA ($r^2 = 0.66$) [12]. Bohlin and Knight's clocks only have two CpG sites in common (cg07816074, cg16536918), suggesting that they may offer a complementary means of estimating epigenetic gestational age.

Both Knight and Bohlin's clocks were initially trained and developed with DNA methylation data from the Illumina 450K array. Until 2016, the Illumina 450K array was the most widely used method to characterize DNA methylation levels, measuring around 450,000 CpG sites in humans. The newer Illumina Infinium HumanMethylationEPIC 850K array (EPIC) improves coverage of regulatory elements by measuring methylation levels of over 850,000 CpG sites, including over 90% of the sites analysed with the 450K and an additional 413,743 CpGs [13]. A cord blood-based gestational age clock using the EPIC DNA methylation array was recently

developed [14]. In this study, we will focus on the Knight and Bohlin clocks due to their widespread use and because of the greater availability of 450K array DNAm profiles in our cohort.

While adult epigenetic clocks have been validated with the EPIC array [15,16], the impact of the newer EPIC array on the estimation of Bohlin and Knight's DNAm GA has been evaluated by few other studies thus far [17,18]. Some studies have explored the associations between pregnancy and birth outcomes with GAA [19–23], but findings have been inconsistent. For example, one study reported female infant sex to be associated with increased GAA [20], while another found an association between male infant sex and increased GAA [21]. Furthermore, these studies have solely used DNA methylation profiles from the 450K array and applied them to predominantly white, European cohorts [19–23]. There is a clear need for epigenetic studies that include diverse populations and validate the EPIC DNA methylation array with existing cord blood-based gestational age clocks.

We aimed to evaluate the performance of the Knight and Bohlin gestational age clocks using the Illumina 450K and EPIC arrays by the correlation between chronological GA and DNAm GA in infants. We also aimed to determine the associations of prenatal maternal factors and birth outcomes with altered GAA using data from Latino mothers and children participating in the CHAMACOS birth cohort. We hypothesize that prenatal maternal and newborn birth characteristics influence epigenetic gestational age acceleration, thereby disrupting foetal programming and ageing.

Methods

Study population

Between October 1999 and October 2000, the Center for the Health Assessment of Mothers and Children of Salinas (CHAMACOS) study recruited 601 pregnant women from farmworker communities of Salinas Valley in California. At enrollment, women were ≤ 20 weeks gestation, English- or Spanish-speaking, Medicare eligible, planning to deliver at the county hospital, and

attending prenatal care visits at one of the six local community clinics or hospitals. Of 601 initial enrollees of the cohort, 526 were followed to delivery of live, singleton newborns in 2000–2001 [24,25]. For this study, we included mother-child pairs with available DNA methylation data who provided samples and consent for genomic analysis. Mothers were interviewed near the end of the 1st and 2nd trimesters of pregnancy and shortly after delivery and provided biological samples at these time points. Study activities were conducted by well-trained, bilingual, and bicultural study staff [24,25].

A subset of children in the CHAMACOS cohort had DNAm measured from cord blood specimens using the 450K array (N = 381). Of these 381 participants, a subset also had cord blood specimens analysed by the EPIC array (N = 109). Nine participants of N = 381 were excluded due to missing reported sex or mismatch between reported sex and predicted sex derived from X chromosome intensities. These samples were deemed of low quality based on array intensity. One participant of N = 109 was excluded due to mismatch between reported and predicted sex results. Based on this exclusion criterion, the final sample sizes were 372 children for the epigenetic data derived from the 450K array and 108 children for the epigenetic data generated from the EPIC array. N = 359 (96.5%) mothers of the 372 children with 450K array data identified as Latina. Specifically, N = 355 (95.4%) identified as being Mexican, Mexican-American, or Mexican-Indian. All N = 108 (100%) mothers of the 108 children with EPIC array data identified as Latina, with N = 107 (99.1%) being Mexican, Mexican-American, or Mexican-Indian.

Umbilical cord blood DNA methylation

Cord blood specimens were collected at the time of delivery, and DNA was extracted from the banked non-heparinized umbilical cord using QIAamp DNA Blood Maxi Kits (Qiagen, Valencia, CA) according to the manufacturer's protocol with some modifications, as previously described [26]. DNA aliquots of 1 μ g were bisulphite converted using Zymo Bisulphite

Conversion Kits (Zymo Research, Orange, CA). After whole-genome amplification, enzymatic fragmentation, and purification, DNA was applied to the Illumina Infinium HumanMethylation450 and HumanMethylationEPIC BeadChips (Illumina, San Diego, CA) according to the Illumina methylation protocol [27,28]. Genome-wide DNAm levels were measured at 485,512 CpG sites on the 450K array. The EPIC array measured DNA methylation at 866,836 CpG sites, including 90% of the sites used by the 450K array [29]. These arrays calculate the relative methylation (β) for each CpG site as the ratio of methylated-probe signal to total (methylated + unmethylated) fluorescent signal intensity, ranging from 0 (no methylation) to 1 (complete methylation).

Quality control steps included the use of repeats and randomization of samples across chips and plates [30]. Data processing was completed using the *minfi* R package [31] and involved removal of poor-quality samples, filtering of probes containing common SNPs, cross-reactive probes, or low-quality probes (detection p-values >0.01), background correction, dye bias adjustment using the 'noob' method, and quantile normalization, as described elsewhere [29,31]. Proportions of seven cell types (CD8⁺ T cells, CD4⁺ T cells, natural killer cells, B cells, monocytes, granulocytes, and nucleated red blood cells) were estimated from cord blood methylation profiles, using the method of Bakulski et al. [32].

DNAm GA prediction and gestational age acceleration

DNAm GA was calculated using the method and model published by Knight et al., which uses 148 CpG sites to predict GA from cord blood methylation profiles [11]. DNAm GA was additionally calculated using the Bohlin et al. model, which is based on the methylation profiles of 96 CpG sites using the *GAPrediction* Bioconductor package [12,33]. The 450K and EPIC array data in our samples measured methylation at 148 (100%) and 141 (95.3%) of the 148 CpG sites used in the Knight model and 93 (96.9%) and 88 (91.7%) of the 96 CpGs used in the Bohlin et al. model, respectively, after quality control. The missing values for three CpG sites in the

450K array (N = 372 samples per CpG) and 15 CpG sites in the EPIC array (N = 108 samples per CpG) were imputed with the k-nearest-neighbours algorithm as implemented by the *impute.knn* function in the *impute* R package [34]. This imputation approach is commonly used with DNA microarray data and has been shown to be more robust and sensitive than other methods like singular value decomposition and row averages [35]. The first multi-tissue epigenetic clock developed by Horvath uses the k-nearest-neighbours algorithm for imputation [36]. No samples were missing values for the other CpG sites. GAA was calculated as the residuals from a linear model of DNAm GA regressed on chronological GA as a proxy for biological ageing.

Maternal characteristics

Data on maternal characteristics primarily came from medical records and maternal interviews conducted during pregnancy at around 13 and 26 weeks gestation and shortly after delivery. Clinical variables of interest were maternal age at delivery, pre-pregnancy body mass index (BMI), pregnancy weight gain, and serum cholesterol, triglyceride, and lipid levels.

Pre-pregnancy weight was obtained from prenatal medical records. If pre-pregnancy weight was missing from medical records, weight at the first prenatal visit (≤ 13 weeks gestation) was used. If women received prenatal care after 13 weeks gestation, regression models were used to impute pre-pregnancy weight based on weight at all prenatal visits. Pregnancy weight gain was computed as the difference between maternal weight at delivery and pre-pregnancy weight. Maternal weight at delivery was abstracted from medical records. If weight at delivery was missing, weight at last prenatal care visit was used if it occurred within 3 weeks of delivery.

Maternal lipid profiles, including total lipids, triglycerides, and cholesterol (mg/dL), were measured in blood collected at 26 weeks gestation. If a serum sample was not collected at 26 weeks, samples from the time of delivery were used. The serum concentration of total lipids was calculated with concentrations for triglycerides and cholesterol, using a mathematical formula described by Phillips et al. [37].

Other covariates obtained from prenatal study interviews included maternal smoking during

pregnancy (no, yes), alcohol consumption during pregnancy (no, yes), baseline marital status (married/living as married, separated, divorced, single), baseline maternal education status (≤ 6 th grade, 7–12th grade, \geq high school graduate), maternal parity (0, 1, 2+), and maternal occupation during pregnancy (did not work, field work, other agricultural work, other work only). Field work included activities such as harvesting crops and thinning or weeding fields. Agricultural work was defined as tasks other than those which involved direct field work, such as working in a packing shed, nursery/greenhouse, or as a truck driver in the fields. Maternal adversity was characterized in terms of the number of maternal adverse childhood experiences (ACEs) reported by the mother at the 18-year follow-up visit (0, 1–2, 3+). Type of delivery (vaginal, caesarean section), and preterm delivery (< 37 weeks, ≥ 37 weeks) were abstracted by a registered nurse from infant and maternal medical records or delivery logs.

Infant characteristics

Chronological GA in days was obtained by asking mothers the date of their last menstrual period (LMP). If LMP was unknown (N = 56), chronological GA largely based on ultrasound methods was abstracted from medical records by a registered nurse [38]. The data from the medical records was systematically biased down since it was available to the nearest completed week. To account for this, chronological GA from medical records was adjusted by adding a half week before converting to days. Other factors of interest included infant sex and birth weight, which were also abstracted by a registered nurse from infant and maternal medical records or delivery logs.

Statistical analysis

We described our study sample using means and proportions for the characteristics of interest among the entire sample with available 450K array data and the subsample with EPIC array data. We tested the accuracy of epigenetic ageing biomarkers derived from two DNA methylation

technologies (450K and EPIC arrays) via their empirical correlation with chronological GA using Pearson correlation coefficients. The median absolute error (MAE) for each gestational age clock was calculated as the median of the absolute difference between estimated DNAm GA and chronological GA as previously done for the adult epigenetic clocks [39]. We used scatter plots with fitted regression lines and 95% confidence intervals (95% CIs) to depict the accuracy of the clocks across DNA methylation arrays. We also tested the inter-clock agreement by determining the correlation between Knight and Bohlin clock estimates derived from the same DNA methylation array. Similar analysis was also performed to determine inter-array agreement of estimates from the same epigenetic clock, but different DNA methylation arrays.

Due to a larger sample size, associations between maternal and infant characteristics with GAA were determined for the 372 mother-child pairs with DNA methylation profiles analysed with the 450K array. We evaluated the impact of adjusting for potential confounders selected *a priori*, including maternal age at delivery, smoking during pregnancy, maternal education, maternal marital status, and infant sex. Maternal education, marital status, and income as a proxy for prenatal socioeconomic status [40] and prenatal smoking [41] have been associated with differential DNA methylation patterns in cord blood samples. Maternal age at delivery [42] and infant sex [20,21] have also been shown to be related to altered rates of gestational ageing in cord blood samples.

Ageing is associated with changes in immune-related components such as white blood cell composition. CD8⁺ T cells and plasma B cells have been shown to be correlated with extrinsic epigenetic age acceleration (EEAA), a type of epigenetic age acceleration that incorporates immune system ageing [43]. A second multivariate linear regression model for each maternal child characteristic was further adjusted for estimated cell-type proportions (CD8⁺ T cells, CD4⁺ T cells, natural killer cells, B cells, monocytes, granulocytes, and nucleated red blood cells). Adjusting for maternal and infant covariates as well as cell-type proportions allowed us to evaluate associations with GAA independent of changes in immune cell

composition. In this model, GAA is not confounded by changes in blood cell counts, and thus measures ‘pure’ epigenetic ageing effects, or intrinsic epigenetic age acceleration (IEAA) [43]. To examine associations between estimated cell-type proportions, the clocks, and their respective measures of GAA, we also tested their empirical correlations across estimated nucleated cells from the cord blood samples.

We report effect sizes as regression coefficients and 95% CIs and consider statistical significance at $p < 0.05$. We also used the Benjamini–Hochberg method to correct for multiple hypotheses for GAA derived from each epigenetic clock. To improve the comparability of effect sizes, we converted the units of maternal serum biomarkers from mg/dL to g/L in Forest plots. All statistical analyses were performed using RStudio Version 4.1.2 (RStudio, Boston, MA, USA).

Results

Study sample characteristics

Table 1 describes the clinical and socioeconomic characteristics of the 372 mother-child pairs whose DNA methylation profiles were obtained by analysing foetal cord blood samples with the Illumina 450K array. Of the 372 participants, 189 (50.81%) were female and 183 (49.19%) were male. The mean chronological GA at birth for this sample was 39.46 (standard deviation [SD] 1.54) weeks and 6.99% were born preterm. The mean age of mothers at delivery was 25.73 (SD 5.00) years. Mean pre-pregnancy BMI among mothers was 26.78 (SD 5.10) kg/m² with over 50% of mothers being overweight or obese. About 6.45% of women reported smoking tobacco during pregnancy and 22.04% reported consuming any alcohol during pregnancy. Table S1 describes the characteristics of a subset of 108 mother-child pairs who had their foetal cord blood DNA samples also analysed with the EPIC array. This subsample of 108 mother-child pairs had similar clinical and socioeconomic characteristics.

Epigenetic clock performance

We observed positive Pearson correlations between chronological gestational age and the estimated DNAm GA from the Knight clock using

Table 1. Characteristics of participants in the CHAMACOS birth cohort with Illumina 450K DNA methylation profiles from umbilical cord blood samples (N = 372).

	n (%) or Mean \pm SD
Pre-pregnancy body mass index (BMI), kg/m²	26.78 \pm 5.10
Underweight (<18.5)	3 (0.81%)
Normal (18.5 to <25)	152 (40.86%)
Overweight (25 to <30)	140 (37.63%)
Obese (\geq 30)	77 (20.70%)
Pregnancy weight gain, pounds	30.42 \pm 12.35
Maternal serum cholesterol, mg/dL	224.84 \pm 39.93
Maternal serum triglyceride, mg/dL	251.07 \pm 98.99
Total maternal serum lipid, mg/dL	823.75 \pm 157.29
Smoking during pregnancy	
No	346 (93.01%)
Yes	24 (6.45%)
Missing	2 (0.54%)
Alcohol use during pregnancy	
No	279 (75.00%)
Yes	82 (22.04%)
Missing	11 (2.96%)
Maternal age at delivery, years	25.73 \pm 5.00
18–24 years	175 (47.04%)
25–29 years	119 (31.99%)
30–34 years	57 (15.32%)
35–45 years	21 (5.65%)
Type of delivery	
Vaginal	295 (79.30%)
Caesarean	77 (20.70%)
Preterm delivery	
<37 weeks	26 (6.99%)
\geq 37 weeks	346 (93.01%)
Baseline maternal marital status	
Married/Living as married	293 (78.76%)
Single	55 (14.78%)
Separated/Divorced	22 (5.91%)
Missing	2 (0.54%)
Baseline maternal education status	
\leq 6th grade	157 (42.20%)
7–12th grade	138 (37.10%)
\geq High school graduate	77 (20.70%)
Baseline maternal occupation	
Did not work	126 (33.87%)
Agricultural field work	102 (27.42%)
Other agricultural work	46 (12.37%)
Other work only	88 (23.66%)
Missing	10 (2.69%)
Maternal adverse childhood experiences (ACEs)	
0	101 (27.15%)
1–2	49 (13.17%)
3+	54 (14.52%)
Missing	168 (45.16%)
Maternal parity	
0	134 (36.02%)
1	107 (28.76%)
2+	131 (35.22%)
Chronological gestational age, weeks	39.46 \pm 1.54
Infant sex	
Female	189 (50.81%)
Male	183 (49.19%)
Birth weight, grams	3453.25 \pm 487.57

both the 450K ($r = 0.48$, $p < 2.2 \times 10^{-16}$) and EPIC ($r = 0.39$, $p = 3.5 \times 10^{-5}$) methylation profiles (Figure 1a, b). The median absolute difference between Knight's DNAm GA and chronological

GA ('median absolute error') in test samples was 5.89 days among the subjects analysed with the 450K array and 8.99 days among those with the EPIC array. A positive Pearson correlation was also observed between chronological GA and DNAm GA as calculated with the Bohlin clock using the 450K ($r = 0.67$, $p < 2.2 \times 10^{-16}$) and EPIC ($r = 0.60$, $p = 7.7 \times 10^{-12}$) methylation measurements (Figure 1c, d). The median absolute errors between Bohlin's DNAm GA and chronological GA were 4.89 days and 4.95 days when this clock was applied to the 450K and EPIC samples, respectively.

Comparison of DNAm GA estimates from the Bohlin and Knight clocks

Figure 2 shows the relationship between the DNAm GA estimates from the two clocks across the same methylation array. As estimated with the 450K DNA methylation array, DNAm GA from the Knight clock was strongly correlated with DNAm GA from the Bohlin clock ($r = 0.81$, $p < 2.2 \times 10^{-16}$) (Figure 2a). Similarly, in the subsample with EPIC array measurements, DNAm GA from the Knight clock was strongly correlated with DNAm GA from the Bohlin clock ($r = 0.71$, $p < 2.2 \times 10^{-16}$) (Figure 2b). GAA derived from both clocks was also found to be highly correlated ($r = 0.76$, $p < 2.2 \times 10^{-16}$) (Figure 5).

Comparison of DNAm GA estimates from 450K and EPIC methylation array

Figure 3 compares DNAm GA estimates from the two methylation arrays across the same gestational age clock from the same child. Using the Knight clock, DNAm GA estimates from the EPIC array were strongly correlated with those from the 450K array ($r = 0.77$, $p < 2.2 \times 10^{-16}$) (Figure 3a). The correlation across arrays using the Bohlin clock was stronger with less deviation ($r = 0.98$, $p < 2.2 \times 10^{-16}$) (Figure 3b). Figure S2 shows a strong, positive correlation between average estimated DNA methylation levels (β) measured by the 450K vs. EPIC array from CHAMACOS cord blood samples ($r = 0.9963$, $p < 2.2 \times 10^{-16}$). Each point represents one of the 224

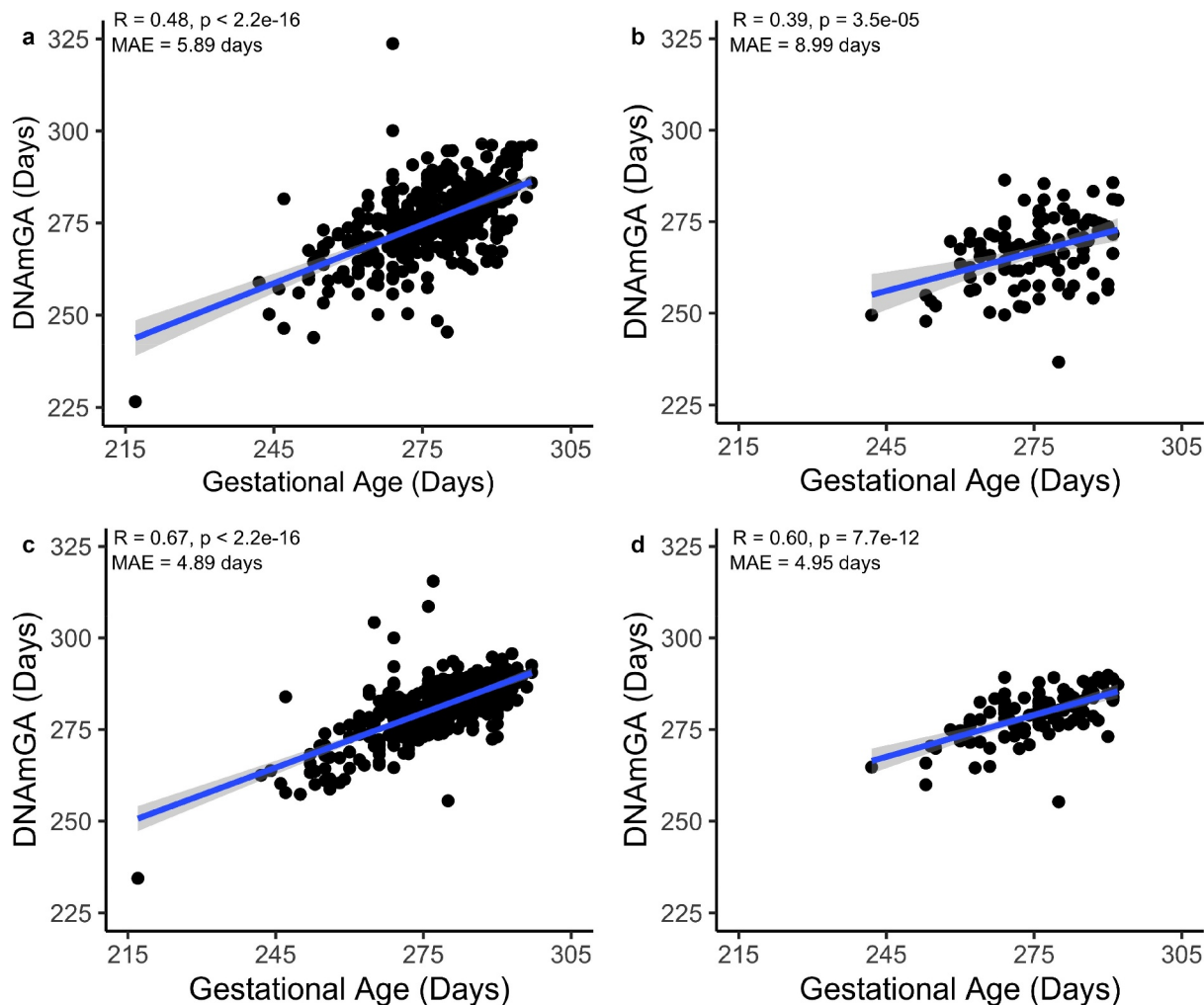


Figure 1. Performance of the DNAm GA Clock Predictions in the CHAMACOS Cohort Using Illumina 450K ($N = 372$) and EPIC ($N = 108$) DNAm Arrays. Pearson's correlation and median absolute error (MAE) between chronological gestational age and DNAm GA calculated using (a) Knight's clock and the 450K methylation data, (b) Knight's clock and the EPIC methylation data, (c) Bohlin's clock and the 450K methylation data, and (d) Bohlin's clock and the EPIC methylation data.

non-missing CpG sites used by the Knight and Bohlin epigenetic clocks.

Adjusted associations of maternal and infant characteristics with GAA

Table 2 shows the results from linear regression models testing associations between GAA and individual maternal and infant characteristics among the 372 mother-child pairs with DNA methylation profiles from the Illumina 450K array. These associations were adjusted for maternal age at delivery, smoking during pregnancy, maternal education, marital status, and infant sex, selected *a priori* as potential confounders. GAA was defined as the residual from regressing

DNAm GA estimates from the Knight and Bohlin clocks on chronological gestational age. Thus, GAA is independent of chronological gestational age. Positive or increased GAA indicates that epigenetic age, or DNAm GA, is greater than chronological age. Negative or decreased GAA indicates that DNAm GA is less than chronological age.

After covariate adjustment, higher levels of maternal serum triglycerides were significantly associated with decreased GAA (Bohlin: $\beta = -0.01$ days per mg/dL increase, 95% CI: $-0.01, -8.95 \times 10^{-4}$). Higher levels of total maternal serum lipids were also significantly associated with decreased GAA (Bohlin: $\beta = -4.31 \times 10^{-3}$ days per mg/dL increase, 95%

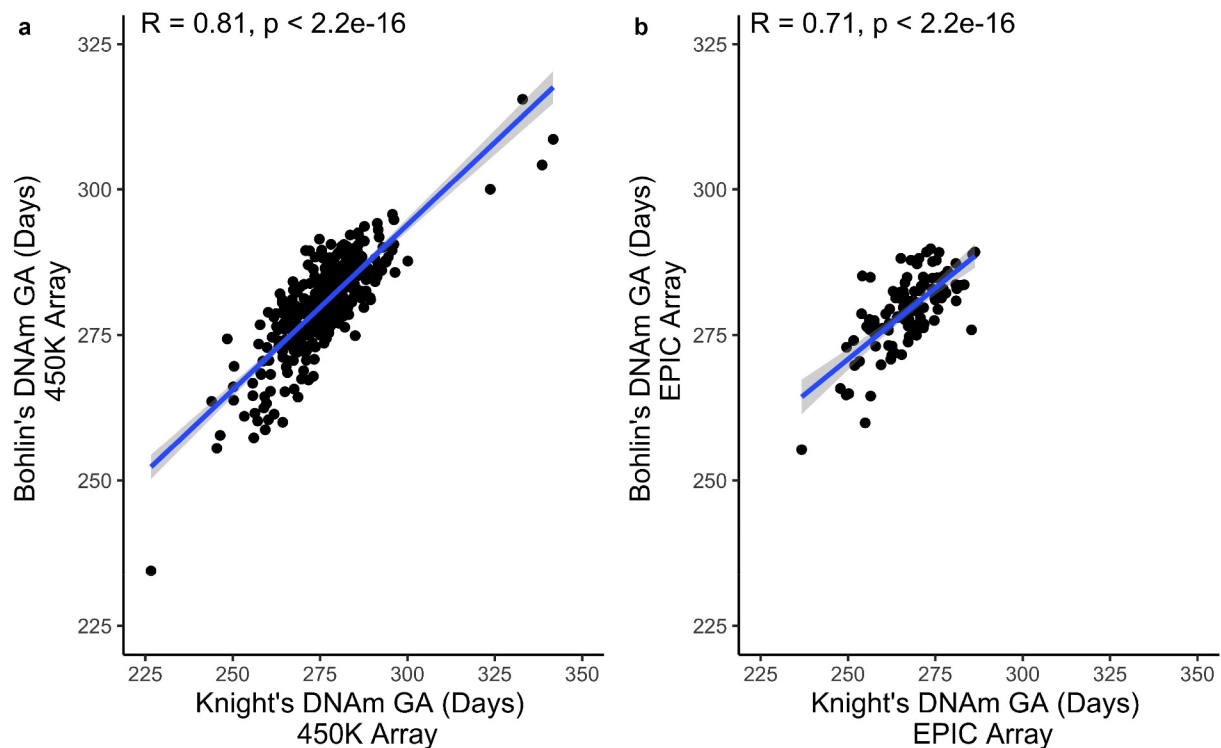


Figure 2. Comparison of DNAm GA Predictions by the Knight vs. Bohlin Clock Across the Same DNAm Array. Pearson's correlation between (a) Knight DNAm GA and Bohlin DNAm GA determined from the 450K DNA methylation data ($N = 372$) and (b) Knight DNAm GA and Bohlin DNAm GA determined from the EPIC DNA methylation data ($N = 108$).

CI: -8.82×10^{-3} , -1.12×10^{-4}). Preterm delivery (<37 weeks) was associated with decreased GAA (Bohlin: $\beta = -4.03$ days, 95% CI: -6.41 , -1.65). Compared to infants of mothers without any prior deliveries (parity = 0), infants of mothers who had two or more prior deliveries had decreased GAA (Knight: $\beta = -4.07$ days, 95% CI: -6.96 , -1.19 ; Bohlin: $\beta = -2.43$ days, 95% CI: -4.15 , -0.72). Male infants had decreased GAA compared to females (Knight: $\beta = -3.15$ days, 95% CI: -5.22 , -1.07). After adjusting for multiple comparisons, preterm delivery remained significantly associated with decreased GAA. [Figure 4](#) depicts estimates and 95% CIs for these associations after covariate adjustment.

To determine if associations between the prenatal and birth environment with GAA might be influenced by age-related changes in immune cell-type composition, we further adjusted models by estimated nucleated cell-type proportions in cord blood (CD8+, CD4+, natural killer cells, B-cells, monocytes, granulocytes, and nucleated red blood cells), as well as the maternal and

child covariates from the initial adjusted model. After additional adjustment of cell-type proportions, greater maternal serum triglyceride levels (Bohlin: $\beta = -0.01$, 95% CI: -0.01 , -2.03×10^{-5}), greater maternal parity (Knight: $\beta = -3.50$ days, 95% CI: -6.21 , -0.80 ; Bohlin: $\beta = -2.00$ days, 95% CI: -3.64 , -0.37) and male infant sex (Knight: $\beta = -2.75$ days, 95% CI: -4.69 , -0.81) remained associated with decreased GAA ([Table S2](#), [Figure S1](#)).

We further tested correlations between immune cell-type proportions and GAA as derived from the Bohlin and Knight clocks using the DNA methylation profiles from the 450K array. GAA estimated from the Bohlin and Knight clocks were both significantly correlated with lower proportions of B cells (Bohlin GAA: $r = -0.34$, Knight GAA: $r = -0.41$) and greater proportions of monocytes (Bohlin GAA: $r = 0.28$, Knight GAA: $r = 0.24$), and granulocytes (Bohlin GAA: $r = 0.20$, Knight GAA: $r = 0.27$) in the cord blood samples ([Figure 5](#)). CD4 T-cell proportions were significantly negatively correlated with GAA derived from the Bohlin clock ($r = -0.15$). Natural

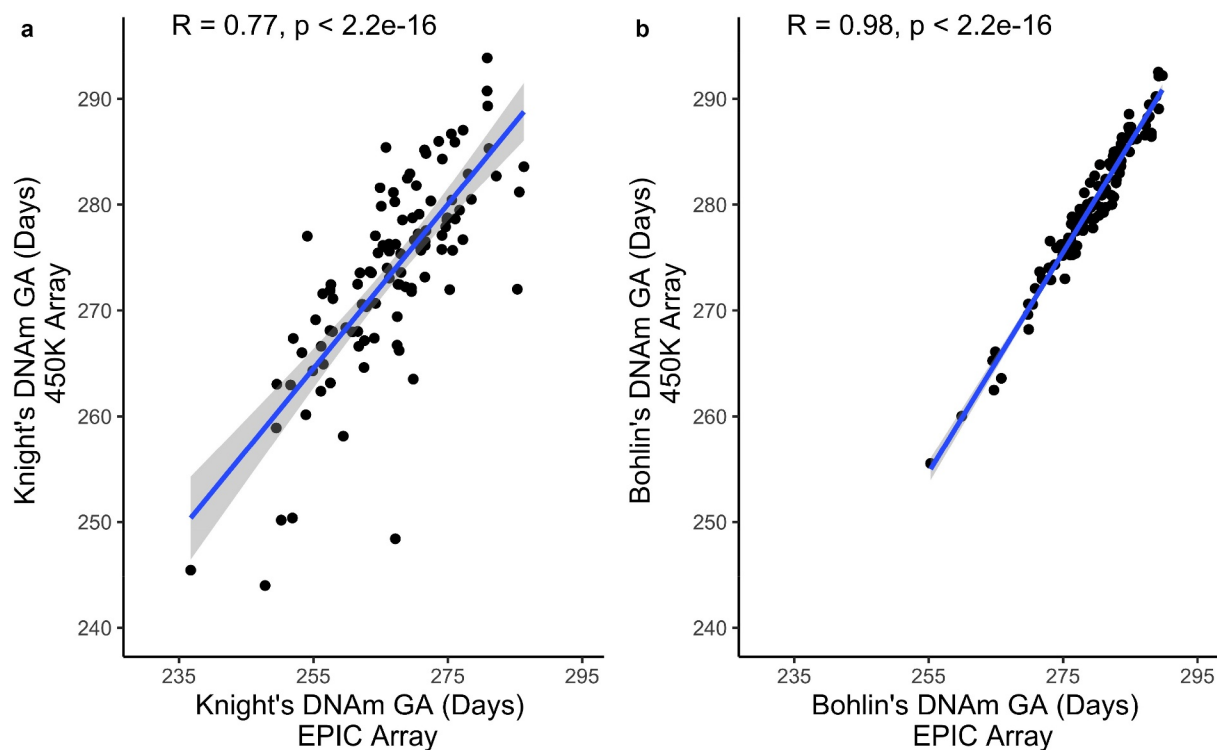


Figure 3. Comparison of DNAm GA Predictions by the Illumina 450K vs. EPIC DNAm array across the same epigenetic clock (N=108). Pearson's correlation between (a) Knight's DNAm GA from the EPIC array and Knight's DNAm GA from the 450K array (b) Bohlin's DNAm GA from the EPIC array and Bohlin's DNAm GA from the 450K array.

killer cell (NK) and nucleated red blood cell (nRBC) proportions were not correlated with GAA for either clock (Figure 5).

Discussion

We observed that several maternal and newborn parameters, some known to reflect an adverse intrauterine environment, were associated with changes in GAA, a measure of biological ageing at birth. Higher levels of maternal serum triglycerides and total serum lipids during pregnancy, pre-term delivery, greater maternal parity, and male infant sex were associated with decreased GAA. After adjustment for estimated cell-type proportions, some of these covariates remained negatively associated with GAA, which suggests that these exposures influence intrinsic rates of ageing. There was no consistent evidence of associations of GAA with pre-pregnancy BMI, pregnancy weight gain, maternal age at delivery, alcohol consumption and smoking during pregnancy, type of delivery, maternal adverse childhood experiences,

or maternal socioeconomic status (*i.e.*, marital status, education, occupation).

Our study presented novel findings on the association between maternal serum triglyceride levels, total maternal serum lipid levels, preterm delivery, parity, and infant sex with GAA. Although some of these factors associated with rates of epigenetic ageing are known to influence chronological gestational age, our results now provide evidence that, beyond disrupting time of delivery, these factors also influence the biological age of the newborns. GAA was derived from DNA methylation levels in cord blood samples. Identifying exposures associated with GAA makes it clear that biological ageing is a phenomenon that starts as early as conception, influencing the progression of health outcomes later in life, a paradigm in line with the Developmental Origins of Health and Disease (DOHaD) hypothesis [44].

In adults, accelerated epigenetic ageing has been shown to be predictive of poor health outcomes [2,4,5]. However, in the intrauterine environment, gestational age acceleration and gestational age deceleration can both be

Table 2. Adjusted associations between gestational age acceleration estimated by the Knight and Bohlin epigenetic clocks with prenatal and birth characteristics in the CHAMACOS cohort (N = 372)^{†,‡}.

	Knight Clock 450K DNAm Array			Bohlin Clock 450K DNAm Array		
	N	β (95% CI)	P-Value [†]	β (95% CI)	P-Value [†]	BH Adjusted P-Value [‡]
Pre-pregnancy BMI, kg/m²	369					
Ref = Underweight/Normal (<25)						
Overweight (25 to <30)		1.44 (-0.91, 3.78)	0.23	0.26 (-1.14, 1.67)	0.71	0.96
Obese (30 or higher)		-2.77 (-5.65, 0.11)	0.06	-1.20 (-2.92, 0.53)	0.17	0.68
Pregnancy weight gain, pounds	368	0.03 (-0.06, 0.12)	0.48	-2.24x10 ⁻³ (-0.05, 0.05)	0.93	0.96
Maternal serum cholesterol, mg/dL	306	-2.49x10 ⁻³ (-0.03, 0.03)	0.86	-8.37x10 ⁻³ (-0.03, 8.82 x 10 ⁻³)	0.34	0.90
Maternal serum triglyceride, mg/dL	306	-0.01 (-0.2, 1.96 x 10 ⁻³)	0.11	-0.01 (-0.01, -8.95x10 ⁻⁴)	0.03	0.28
Total maternal serum lipid, mg/dL	306	-4.02x10 ⁻³ (-0.01, 3.04 x 10 ⁻³)	0.26	-4.31x10 ⁻³ (-8.82x10 ⁻³ , -1.12x10 ⁻⁴)	0.04	0.28
Smoking during pregnancy (Ref = No)	369					
Yes		-0.01 (-4.34, 4.32)	0.99	-0.87 (-3.45, 1.71)	0.51	0.95
Alcohol use during pregnancy (Ref = No)	360					
Yes	369	-1.57 (-4.17, 1.03)	0.23	-1.28 (-2.81, 0.26)	0.10	0.56
Maternal age at delivery, years	369					
Ref = 18–24 years						
25–29 years		-1.53 (-3.93, 0.87)	0.21	-0.41 (-1.84, 1.02)	0.58	0.96
30–34 years		-0.05 (-3.14, 3.04)	0.97	-4.37x10 ⁻³ (-1.85, 1.84)	0.99	0.99
35–45 years		-0.03 (-4.68, 4.62)	0.99	0.41 (-2.36, 3.18)	0.77	0.96
Type of delivery (Ref = Caesarean)	369					
Vaginal						
Preterm delivery (Ref = ≥37 weeks)	369	-0.78 (-3.36, 1.80)	0.55	-0.27 (-1.81, 1.26)	0.73	0.96
<37 weeks	369	-3.17 (-7.23, 0.88)	0.12	-4.03 (-6.41, -1.65)	9.64x10 ⁻⁴	0.03
Baseline maternal marital status (Ref = Married/Living as married)	369					
Single		2.13 (-0.88, 5.13)	0.16	1.24 (-0.55, 3.03)	0.17	0.68
Separated/Divorced		2.14 (-2.29, 6.56)	0.34	0.49 (-2.14, 3.13)	0.71	0.96
Baseline maternal education status (Ref = ≤ 6th grade)	369					
7–12th grade		1.12 (-1.23, 3.47)	0.35	0.79 (-0.60, 2.19)	0.26	0.81
≥ High school graduate		-0.54 (-3.41, 2.32)	0.71	-0.72 (-2.43, 0.98)	0.40	0.90
Baseline maternal occupation (Ref = Did not work)	361					
Field work		0.58 (-2.17, 3.32)	0.68	0.67 (-0.95, 2.29)	0.42	0.90
Other agricultural work		-0.20 (-3.69, 3.29)	0.91	0.24 (-1.82, 2.30)	0.82	0.96
Other work only		1.47 (-1.47, 4.41)	0.33	0.74 (-1.00, 2.47)	0.40	0.90
Maternal ACEs (Ref = 0)	203					
1–2		-1.04 (-4.72, 2.65)	0.58	-0.17 (-2.25, 1.91)	0.87	0.96
3+		-2.31 (-5.92, 1.30)	0.21	-0.77 (-2.80, 1.27)	0.46	0.92
Maternal parity (Ref = 0)	369					
1		-0.67 (-3.32, 1.98)	0.62	-0.42 (-2.00, 1.15)	0.60	0.96
2+		-4.07 (-6.96, -1.19)	0.01	-2.43 (-4.15, -0.72)	0.01	0.14
Infant sex (Ref = Female)	369					
Male		-3.15 (-5.22, -1.07)	3.10x10 ⁻³	0.14 (-1.10, 1.37)	0.82	0.96
Birth weight, grams	369					
Male		1.58x10 ⁻³ (-6.06x10 ⁻⁴ , 3.76 x 10 ⁻³)	0.16	8.29x10 ⁻⁴ (-4.70x10 ⁻⁴ , 2.13 x 10 ⁻³)	0.21	0.74

[†] Adjusted for maternal age at delivery, smoking during pregnancy, maternal education, marital status, and infant sex

[‡] Adjusted for multiple comparisons using the Benjamini-Hochberg method
 GAA, gestational age acceleration; BH, Benjamini-Hochberg; CI, confidence interval; BMI, body mass index; ACE, adverse childhood experience

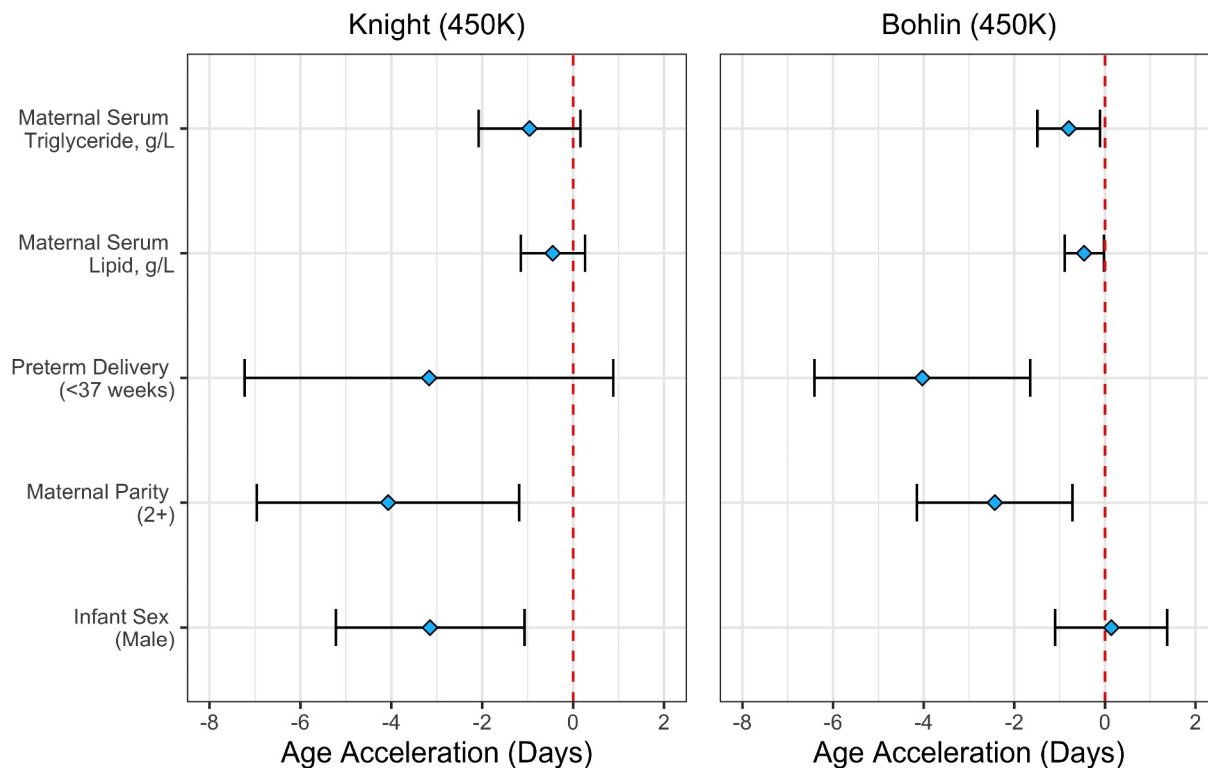


Figure 4. Adjusted Associations between Gestational Age Acceleration and Prenatal and Birth Characteristics in the CHAMACOS Cohort ($N = 372$)[†]. Estimates (β) and 95% CIs for characteristics significantly associated with gestational age acceleration after adjusting for maternal age at delivery, maternal smoking during pregnancy, maternal education, marital status, and infant sex across the two clocks. [†]Units for maternal serum triglyceride and maternal serum lipid levels converted from standard mg/dL to d/L for greater comparability of effect sizes in plot.

considered risk factors for later disease burden. Increased GAA has been associated with higher birth weight, length, and head circumference [21,23], which have been shown to contribute to age-related diseases such as obesity and increased mortality [45]. Decreased GAA has been linked to prenatal environmental adversity and lower developmental maturity [22,46,47] – risk factors for age-related diseases like cardiovascular disease and non-insulin dependent diabetes [48,49]. Longitudinal analyses should be designed to follow the impact of altered GAA on disease onset and progression later in life. We hypothesized that deaccelerated or accelerated ageing at birth and in childhood are both risk factors for health and development. Our hypothesis is that birth and paediatric epigenetic ageing might follow a specific tempo and disruptions (i.e., acceleration or deacceleration) are both detrimental.

Consistent with findings from Simpkin et al. [50] and Dieckmann et al. [17], we also demonstrate that the Bohlin clock predicted GA with

a higher accuracy compared to the Knight clock using DNAm data from both the 450K and EPIC arrays. The Bohlin clock also resulted in predictions that had a lower median absolute error compared to the Knight clock. Even though only two CpGs (cg07816074, cg16536918) overlap between the Knight and Bohlin epigenetic clocks, we observed strong correlations ($r > 0.7$) between DNAm GA estimates from both clocks across the 450K and EPIC arrays. Consistently, the residual GAA estimates as determined by both clocks were found to be highly correlated, supporting that both clocks are measuring biological gestational age by leveraging different changes that occur in the human genome from conception to birth. This suggests that while both clocks reflect foetal development timing, they might differ in their biological insight by utilizing methylation of different sets of genes. Indeed, associations with sex were only captured by Knight's GAA and not Bohlin's GAA. Our results suggest that the greater accuracy of the Bohlin clock makes it more sensitive to detect

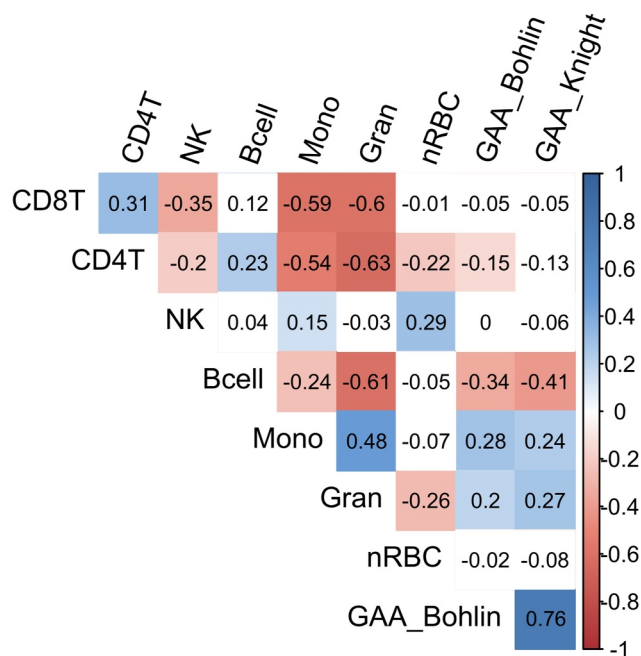


Figure 5. Correlations Coefficients Between Estimated Cell-Type Proportions and Gestational Age Acceleration Estimated by the Knight and Bohlin Clocks (N = 372). Pearson's correlations between cell-type proportions estimated from DNA methylation profiles and gestational age acceleration calculated by the Bohlin and Knight models using the 450K array (blank/white cell = non-significant correlation $p > 0.05$).

associations, including those related to the external environment (i.e., triglycerides and lipids), while the Knight clock might be able to capture biological influences such as parity and sex. However, this needs further confirmation in other cohorts.

Another important observation in our data is that DNAm GA estimates from both clocks were more accurate when derived from the 450K DNAm data as opposed to the EPIC data. As seen by the strong correlation in average CpG DNAm levels estimated by the 450K and EPIC arrays, this difference in accuracy is not driven by some CpG sites. Instead, it may arise because both clocks were initially developed and trained with DNAm data from the 450K arrays, which measures fewer and different CpG sites than the newer EPIC array. Poorer accuracy using the EPIC DNAm array can also be attributed to a smaller overall sample size and more missing CpG sites in our EPIC DNAm profile. Seven CpGs (out of 148 CpGs) used in the Knight model and eight CpGs (out of 96 CpGs) used in the Bohlin model were missing. In contrast, none

of the 148 CpGs used in the Knight model and three CpGs (out of 96 CpGs) used in the Bohlin model were missing in the 450K data after quality control. Compared to the Knight clock, the Bohlin clock demonstrated a higher accuracy using data from the EPIC array as shown by its stronger correlations between DNAm GA derived from EPIC data with chronological GA as well as between DNAm GA estimates from both array types. Correlations between DNAm GA and GA may provide insight into the accuracy of these clocks in predicting chronological gestational age. However, greater predictive accuracy does not necessarily translate to a better measure of biological age, which often deviates from chronological age.

Heterogeneous tissues, like cord blood, can vary in immune cell-type proportions over time and might be responsive to environmental conditions. We observed significant relationships between certain cell-type proportions and GAA estimates from both the Knight and Bohlin clocks. GAA residuals from both clocks had a significant moderate negative correlation with B cell proportions and significant positive correlations with monocyte and granulocyte proportions. GAA from the Bohlin clock was also negatively correlated with CD4T proportions. These findings highlight that, prior to adjustment, GAA from the Knight and Bohlin clocks might partially reflect and depend on cell-type proportions for a given sample and inherently measure extrinsic epigenetic age acceleration, a measure of ageing that captures age-related changes in immune cell composition [43]. However, our findings related to maternal parity and infant sex remained significant after adjustment for cell-type composition, demonstrating an influence on intrinsic ageing processes independent of the immune system. Going forward, we suggest that investigators test both unadjusted and cell-type composition adjusted associations to address changes that might be related or independent of immune system changes. This approach will help test the influence of exposure on different aspects of ageing to help target interventions and biological pathways.

Our finding regarding male sex and decreased GAA was only supported by another study by Girchenko et al. [20]. In contrast, Khouja et al.

reported that male infants had greater GAA [21]. Similarly, Horvath et al. found that adult men experienced higher rates of epigenetic ageing [51], but this study did not include newborns, which could influence findings. According to Knight et al., the predictive accuracy of their clock is not dependent on neonatal sex, so further investigation is needed to corroborate the relationship of infant sex with differences in GAA derived from the Knight clock [11]. We did not observe some previously reported associations in our cohort. For example, Girchenko et al. found a higher maternal age (>40 years at delivery) to be associated with increased GAA among newborns [20]. Furthermore, Khouja et al. reported that increased maternal pre-pregnancy BMI and vaginal delivery were associated with increased GAA – a finding that we also did not observe in our population [21]. We also did not find increased GAA to be associated with increased maternal BMI, maternal smoking during pregnancy, caesarean section delivery [19], and increased birth weight and length [23], as reported by Simpkin et al. and Bright et al., respectively.

The main strength of our study was its use of a well-characterized, primarily Latino birth cohort to extend studies of early life influences on epigenetic age acceleration. In addition to increasing diversity and representation within epigenetic studies, our study filled important gaps in knowledge by incorporating two relatively new epigenetic clocks that were specifically developed using umbilical cord blood and designed for newborns. By validating the performance of these clocks with both the traditional 450K and newer EPIC arrays, we hope to pave the way for future epigenetic ageing research on prenatal and early life exposures and the programming of ageing. The prospective nature of the study allowed us to assess exposures that occurred before or during the prenatal period and their association with GAA at birth. Our definition of GAA as the residual of DNAm GA regressed on chronological GA as opposed to the raw difference between these two measures allowed us to entirely remove the effect of chronological GA on DNAm GA in our analysis. This analytical approach has been taken with adult epigenetic clocks to reduce the residual influence of chronological age from inferences.

There are also several study limitations to consider. The chronological GA data are of different granularity with accuracy to the closest day for those with LMP estimates available and to the nearest completed week for those with ultrasound-based estimates from medical records. However, this is unlikely to be systematic and any exposure misclassification would be non-differential biasing results towards the null. The ethnic homogeneity of our sample may preclude generalizations to other groups. Our study design is primarily observational and draws upon cohort data, but some of our analyses are cross-sectional (i.e., maternal age at delivery, type of delivery, preterm delivery, birth weight, parity), which means we cannot establish causality but can only identify correlates of gestational age acceleration. While we adjusted for multiple comparisons, we also presented estimates and 95% CIs without this adjustment due to our limited sample size and interest in describing the properties of various epigenetic clocks in our cohort. We reported effect sizes and 95% CIs as measures of precision to allow for comparability to future studies.

Conclusions

Our study provides important evidence that the prenatal environment and neonatal birth parameters significantly impact epigenetic age acceleration at birth. Understanding the factors that affect biological ageing in newborns can help to identify early life influences of age-related conditions and new targets for public health interventions.

Acknowledgments

We gratefully acknowledge CHAMACOS participants and staff.

Disclosure statement

The authors declare they have no actual or potential competing financial interests.

Authors' contributions

SD analysed and visualized the data and drafted the manuscript. K Huen assisted with the design of the data analysis, performed laboratory epigenetics analyses, and edited the

manuscript. LVDL and PAC helped with epigenetic data analysis and data management. JCN edited the manuscript. K Harley, JD, BE, and NH conceived the original study, supervised initial data and sample collection, and edited the manuscript. AC conceptualized the project, developed the data analysis plan, was involved in developing study methodology and funding acquisition, and edited the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

Datasets generated and analysed during the current study are available from the corresponding authors with appropriate permission from the CHAMACOS study team and investigators upon reasonable request and institutional review board approval.

Ethics approval and consent to participate

The University of California, Berkeley Committee for the Protection of Human Subjects approved all study activities. Written informed consent was obtained from the CHAMACOS mothers.

Funding

This work was supported by the National Institutes of Health (R03AG067064; R01ES031259; R01ES026994; 1R01ES021369; P01ES009605; R01ES017054; R01DA035300; R24ES028529 and R01MD016595) and by the Environmental Protection Agency (R82670901; RD83171001; and RD83451301);US Environmental Protection Agency [R82670901, RD83171001, RD83451301].

ORCID

Saher Daredia  <http://orcid.org/0000-0002-7906-4087>

Andres Cardenas  <http://orcid.org/0000-0003-2284-3298>

References

- [1] Yu M, Hazelton WD, Luebeck GE, et al. Epigenetic aging: more than just a clock when it comes to cancer. *Cancer Res.* 2020 Feb 1;80(3):367–374.
- [2] Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet.* 2018 Jun;19(6):371–384.
- [3] Pal S, Tyler JK. Epigenetics and aging. *Sci Adv.* 2016 Jul 1;2(7):e1600584.
- [4] Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene.* 2002 Aug;21(35):5427–5440.
- [5] Lorenzen JM, Martino F, Thum T. Epigenetic modifications in cardiovascular disease. *Basic Res Cardiol.* 2012 Jan 11;107(2):245.
- [6] Bell CG, Lowe R, Adams PD, et al. DNA methylation aging clocks: challenges and recommendations. *Genome Biol.* 2019 Nov 25;20(1):249.
- [7] Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY).* 2019 Jan 21;11(2):303–327.
- [8] Engle WA. Morbidity and mortality in late preterm and early term newborns: a continuum. *Clin Perinatol.* 2011 Sep 1;38(3):493–516.
- [9] Yang S, Platt RW, Kramer MS. Variation in child cognitive ability by week of gestation among healthy term births. *Am J Epidemiol.* 2010 Feb 15;171(4):399–406.
- [10] Hansen AK, Wisborg K, Uldbjerg N, et al. Risk of respiratory morbidity in term infants delivered by elective caesarean section: cohort study. *BMJ.* 2008 Jan 10;336(7635):85–87.
- [11] Knight AK, Craig JM, Theda C, et al. An epigenetic clock for gestational age at birth based on blood methylation data. *Genome Biol.* 2016 Oct 7;17(1):206.
- [12] Bohlin J, Häberg SE, Magnus P, et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biol.* 2016 Oct 7;17(1):207.
- [13] Pidsley R, Zotenko E, Peters TJ, et al. Critical evaluation of the illumina methylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol.* 2016 Oct 7;17(1):208.
- [14] Haftorn KL, Lee Y, Denault WRP, et al. An EPIC predictor of gestational age and its application to newborns conceived by assisted reproductive technologies. *Clin Epigenetics.* 2021 Apr 19;13(1):82.
- [15] Dhingra R, Kwee LC, Diaz-Sanchez D, et al. Evaluating DNA methylation age on the illumina MethylationEPIC bead chip. *PLoS ONE.* 2019 Apr 19;14(4):e0207834.
- [16] McEwen LM, Jones MJ, Lin DTS, et al. Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. *Clin Epigenetics.* 2018 Oct 16;10(1):123.
- [17] Dieckmann L, Lahti-Pulkkinen M, Kvist T, et al. Characteristics of epigenetic aging across gestational and perinatal tissues. *Clin Epigenetics.* 2021 Apr 29;13(1):97.
- [18] Chen L, Wagner CL, Dong Y, et al. Effects of maternal vitamin D3 supplementation on offspring epigenetic clock of gestational age at birth: a post-hoc analysis of a randomized controlled trial. *Epigenetics.* 2020 Aug 2;15(8):830–840.
- [19] Simpkin AJ, Hemani G, Suderman M, et al. Prenatal and early life influences on epigenetic age in children: a study of mother–offspring pairs from two cohort studies. *Hum Mol Genet.* 2016 Jan 1;25(1):191–201.
- [20] Girchenko P, Lahti J, Czamara D, et al. Associations between maternal risk factors of adverse pregnancy and

- birth outcomes and the offspring epigenetic clock of gestational age at birth. *Clin Epigenetics*. 2017 May 8;9(1):49.
- [21] Khouja JN, Simpkin AJ, O’Keeffe LM, et al. Epigenetic gestational age acceleration: a prospective cohort study investigating associations with familial, sociodemographic and birth characteristics. *Clin Epigenetics*. 2018 Jun 27;10(1):86.
- [22] Suarez A, Lahti J, Czamara D, et al. The epigenetic clock at birth: associations with maternal antenatal depression and child psychiatric problems. *J Am Acad Child Adolesc Psychiatry*. 2018 May;57(5):321–328.e2.
- [23] Bright HD, Howe LD, Khouja JN, et al. Epigenetic gestational age and trajectories of weight and height during childhood: a prospective cohort study. *Clin Epigenetics*. 2019 Dec 16;11(1):194.
- [24] CHAMACOS Study | CERCH [Internet]. [cited 2021 Jun 5]. Available from: <https://cerch.berkeley.edu/research-programs/chamacos-study>
- [25] Eskenazi B, Bradman A, Gladstone EA, et al. CHAMACOS, a longitudinal birth cohort study: lessons from the fields. *J Child Health*. 2003 Jan 1;1(1):3–27.
- [26] Holland N, Furlong C, Bastaki M, et al. Paraoxonase polymorphisms, haplotypes, and enzyme activity in Latino mothers and newborns. *Environ Health Perspect*. 2006 Jul;114(7):985–991.
- [27] Bibikova M, Barnes B, Tsan C, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011 Oct;98(4):288–295.
- [28] Sandoval J, Heyn H, Moran S, et al. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics*. 2011 Jun 1;6(6):692–702.
- [29] Solomon O, MacIsaac J, Quach H, et al. Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics*. 2018 Jun 3;13(6):655–664.
- [30] Yousefi P, Huen K, Davé V, et al. Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. *BMC Genomics*. 2015 Nov 9;16(1):911.
- [31] Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinforma Oxf Engl*. 2014 May 15;30(10):1363–1369.
- [32] Bakulski KM, Feinberg JI, Andrews SV, et al. DNA methylation of cord blood cell types: applications for mixed cell birth studies. *Epigenetics*. 2016 May 3;11(5):354–362.
- [33] Bohlin J. GAprediction: prediction of gestational age with Illumina HumanMethylation450 data [Internet]. Bioconductor version: Release (3.14); 2022 [cited 2022 Apr 17]. Available from: <https://bioconductor.org/packages/GAprediction/>
- [34] Hastie T, Tibshirani R, Narasimhan B, et al. impute: impute: imputation for microarray data [Internet]. Bioconductor version: Release (3.14); 2022 [cited 2022 Apr 17]. Available from: <https://bioconductor.org/packages/impute/>
- [35] Troyanskaya O, Cantor M, Sherlock G, et al. Missing value estimation methods for DNA microarrays. *Bioinformatics*. 2001 Jun 1;17(6):520–525.
- [36] DNAm age: frequently asked questions [Internet]. [cited 2022 Apr 17]. Available from: <https://horvath.genetics.ucla.edu/html/dnamage/faq.htm>
- [37] Phillips DL, Pirkle JL, Burse VW, et al. Chlorinated hydrocarbon levels in human serum: effects of fasting and feeding. *Arch Environ Contam Toxicol*. 1989 Jul 1;18(4):495–500.
- [38] Eskenazi B, Harley K, Bradman A, et al. Association of in utero organophosphate pesticide exposure and fetal growth and length of gestation in an agricultural population. *Environ Health Perspect*. 2004 Jul;112(10):1116–1124.
- [39] Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115.
- [40] Laubach ZM, Perng W, Cardenas A, et al. Socioeconomic status and DNA methylation from birth through mid-childhood: a prospective study in project viva. *Epigenomics*. 2019;11(12):1413–1427.
- [41] Joubert BR, Håberg SE, Nilsen RM, et al. 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environ Health Perspect*. 2012 Oct;120(10):1425–1431.
- [42] Markunas CA, Wilcox AJ, Xu Z, et al. Maternal age at delivery is associated with an epigenetic signature in both newborns and adults. *PLoS ONE*. 2016 Jul 6;11(7):e0156361.
- [43] Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson’s disease patients. *Aging (Albany NY)*. 2015 Dec 9;7(12):1130–1142.
- [44] Heindel JJ, Vandenberg LN. Developmental origins of health and disease: a paradigm for understanding disease etiology and prevention. *Curr Opin Pediatr*. 2015 Apr;27(2):248–253.
- [45] Williams CB, Mackenzie KC, Gahagan S. The effect of maternal obesity on the offspring. *Clin Obstet Gynecol*. 2014 Sep;57(3):508–515.
- [46] Knight AK, Smith AK, Conneely KN, et al. Relationship between epigenetic maturity and respiratory morbidity in preterm infants. *J Pediatr*. 2018 Jul;198:168–173.e2.
- [47] Palma-Gudiel H, Eixarch E, Crispi F, et al. Prenatal adverse environment is associated with epigenetic age deceleration at birth and hypomethylation at the hypoxia-responsive EP300 gene. *Clin Epigenetics*. 2019 May 9;11(1):73.
- [48] Barker DJ, Godfrey KM, Fall C, et al. Relation of birth weight and childhood respiratory infection to adult lung function and death from chronic obstructive airways disease. *BMJ*. 1991 Sep 21;303(6804):671–675.

- [49] Visentin S, Grumolato F, Nardelli GB, et al. Early origins of adult disease: low birth weight and vascular remodeling. *Atherosclerosis*. 2014 Dec;237(2):391–399.
- [50] Simpkin AJ, Suderman M, Howe LD. Epigenetic clocks for gestational age: statistical and study design considerations. *Clin Epigenetics*. 2017 Sep 15;9(1):100.
- [51] Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol*. 2016 Aug 11;17(1):171.