Gestational DNA methylation age as a marker for fetal development and birth outcomes: findings from the Boston Birth Cohort

Anat Yaskolka Meir[1], Maria Jimena Gutierrez[2], Xiumei Hong[3], Guoying Wang[3], Xiaobin Wang[3,4]* and Liming Liang[1]*

Abstract

Background Gestational DNA methylation age (GAmAge) has been developed and validated in European ancestry samples. Its applicability to other ethnicities and associations with fetal stress and newborn phenotypes such as inflammation markers are still to be determined. This study aims to examine the applicability of GAmAge developed from cord blood samples of European decedents to a racially diverse birth cohort, and associations with newborn phenotypes.

Methods GAmAge based on 176 CpGs (Haftorn GAmAge) was calculated for 940 children from a US predominantly urban, low-income, multiethnic birth cohort. Cord blood DNA methylation was profiled by Illumina EPIC array. Newborn phenotypes included anthropometric measurements and, for a subset of newborns ($N = 194$), twenty-seven cord blood inflammatory markers (sandwich immunoassays).

Results GAmAge had a stronger correlation with GEAA in boys ($r = 0.89$, 95% confidence interval (CI) [0.87,0.91]) compared with girls ($r = 0.83$, 95% CI [0.80,0.86]), and was stronger among extremely preterm to very preterm babies ($r = 0.91$, 95% CI [0.81,0.96]), compared with moderate ($r = 0.48$, 95% CI [0.34,0.60]) and term babies ($r = 0.58$, 95% CI [0.53,0.63]). Among White newborns ($N = 51$), the correlation between GAmAge vs. GEAA was slightly stronger ($r = 0.89$, 95% CI [0.82,0.94]) compared with Black/African American newborns ($N = 668$; $r = 0.87$, 95% CI [0.85,0.89]) or Hispanic ($N = 221$; $r = 0.79$, 95% CI [0.74,0.84]). Adjusting for GEAA and sex, GAmAge was associated with anthropometric measurements, cord blood brain-derived neurotrophic factor (BDNF), and monocyte chemoattractant protein-1 (MCP-1) ($p < 0.05$ for all).

Conclusions GAmAge estimation is robust across different populations and racial/ethnic subgroups. GAmAge may be utilized as a proxy for GEAA and for assessing fetus development, indicated by inflammatory state and birth outcomes.

Keywords Epigenetic clock, Pediatrics, Inflammation, Birth weight

*Correspondence:
Xiaobin Wang
xwang82@jhu.edu
Liming Liang
lliang@hsph.harvard.edu
Full list of author information is available at the end of the article

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 2 of 12

Introduction

In recent years, gestational DNA methylation age (GAmAge) has emerged as a molecular measure of gestational age (GEAA) based on cord blood DNA methylation patterns [1–3]. Bohlin [1] and Knight [2] described GAmAge clocks based on the previous 27K and 450K methylation arrays, whereas the newest Haftorn clock [3] used methylation data from the EPIC methylation array. These cord blood epigenetic clocks were trained on data from European or mostly White population samples. The growing interest in GAmAge lies in its potential to better capture individual variation in biological maturity beyond the chronological GEAA and gain deeper insight into pathophysiology by defining acceleration vs. deceleration relative to chronological GEAA.

For a long time, GEAA at birth, as estimated either by the last menstrual period or early ultrasound or both, has been a major clinical measure of newborn maturity and predictor of neonatal and infant outcomes [4]. Preterm birth, determined by a GEAA cutoff of 37 weeks, is associated with an increased risk of neonatal and infant morbidity and mortality, in part due to immature organs/systems, including the immune system, making them more susceptible to infections and other health risks [5], with extremely preterm newborns (born before 28 weeks of gestation) having an even higher risk of infections [6]. We have previously shown that in a sample of 927 children, some cord blood inflammation markers were associated with preterm birth, with markers such as interleukin (IL)-2, IL-4, and IL-5 increased in preterm birth, while others, such as IL1-b and IL-18, decreased in preterm birth [7]. Inflammation markers measured in cord blood were also associated with DNA methylation; cord blood DNA methylation was associated with cord blood C-reactive protein (CRP) in a sample of 358 newborns from the EAGeR study [8]. The levels of cord blood inflammation-related intercellular adhesion molecule 1, CRP, IL-6, and serum amyloid A were associated with locus- and placenta-side-specific differential placenta DNA methylation [9].

The link between GAmAge, GEAA, and various inflammatory markers and anthropometric measures may be the first step in understanding how these are central to infant development. Furthermore, more research is needed to fully understand the implications of GAmAge, as estimated from the EPIC array by the Haftorn clock in European descents, and its potential uses, especially in minorities and other under-representative populations in capturing individual biological maturity variation.

In this analysis, we used data from a multiethnic, mostly Black/African American and Hispanic birth cohort with cord blood DNA methylation and inflammatory markers measured. We examined the utility of pre-trained GAmAge based on European samples as a predictor of GEAA among our multiethnic cohort and in different subgroups of sex, ethnicity/race, delivery week, and fetal growth. We further studied the association of GAmAge with immediate birth outcomes as anthropometric measurements (weight, length, head circumference) and inflammatory state (represented by 27 cord blood inflammation markers). All of the above associations were examined beyond the impact of GEAA to elucidate the independent effect of DNA methylation-based biological age as a predictor of health status.

Patients and methods

This study included 940 mother–newborns pairs from the Boston Birth Cohort (BBC; registered in ClinialTrial.gov NCT03228875), a US predominantly urban, low-income, Black and Hispanic population. The BBC was initiated in 1998 with rolling enrollment at the Boston Medical Center in Boston, MA, as detailed elsewhere [10]. In brief, mothers who delivered a singleton live birth at the Boston Medical Center were invited to participate 24–72 h after a vaginal delivery. The BBC is enriched by preterm (< 37 weeks of gestation) and low birth weight (< 2500 g) births by design of over-sampling preterm birth at enrollment. Pregnancies resulting from in vitro fertilization, multiple gestations (e.g., twins, triplets), fetal chromosomal abnormalities, major birth defects, or preterm birth due to maternal trauma were excluded. After mothers provided written informed consent, research assistants (RAs) administered a standardized questionnaire interview on maternal sociodemographic characteristics, lifestyle, including smoking and alcohol consumption, diet, and reproductive and medical history. Maternal and newborn clinical information, including birth outcomes, was abstracted from the medical records. The study protocol has received initial and annual approval from the Institutional Review Boards (IRBs) of Boston Medical Center and the Johns Hopkins Bloomberg School of Public Health.

Main covariates

Mother–child characteristics: For background characteristics and adjustment of the statistical models, we used the following data: maternal age at delivery, parity (nulliparous or multiparous), maternal education (below college or college and higher), maternal self-reported ethnicity/race (Black/African American, White, and Hispanic), maternal pre-pregnancy BMI, maternal diabetes (non, gestational diabetes or preexisting diabetes mellitus), method of delivery, child's sex (female versus male), maternal smoking, birth weight (gram; as continuous and binary with above/below 2500g defined as the low birth weight (LBW) [11]), birth length (cm), birth head

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 3 of 12

circumference (cm), and GEAA as continuous and binary with below week 37 defined as "preterm" [11]. Delivery week subgroups were defined as follows: extremely preterm (< 28 weeks), very preterm (28 to 32 weeks), moderate to late preterm (week 32 to 37), term (37 to 41 weeks), late term (41 to 42 weeks), and post-term (> 42 weeks) as per World Health Organization and American College of Obstetricians and Gynecologists definitions [12]). The estimation of GEAA was detailed before [13] and was performed using an established algorithm based on both the last menstrual period and the result of early ultrasound (< 20 weeks' gestation). Fetal growth groups were determined by the birth weight and gestational age as described before [14]: small for gestational age (SGA), appropriate for gestational age (AGA), and large for gestational age (LGA).

Inflammation markers: The BBC has available cytokines measured for 927 children. A subset of 194 children had both cord blood inflammation markers and DNA methylation measured. The comparison of background characteristics between the group with both cord blood DNA methylation and cord blood inflammation markers measured vs. cord blood DNA methylation only is presented in Additional file 1: Table S1. Twenty-seven inflammation markers were quantified using sandwich immunoassays as detailed before [7] and included: brain-derived neurotrophic factor (BDNF), granulocyte/macrophage colony-stimulating factor (GM-CSF); interferon-gamma (IFN-gamma); IL-1beta, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12, IL-17, and IL-18; monocyte chemoattractant protein1 (MCP-1); macrophage migration inhibitory factor (MIF); macrophage inflammatory protein 1 alpha (MIP-1alpha), MIP-1beta, matrix metalloproteinase 9 (MMP-9); neurotrophin 3 (NT-3) and NT-4; regulated on activation, normal T-cell expressed and secreted (RANTES); soluble IL-6 receptor alpha (sIL-6r alpha); soluble tumor necrosis factor receptor I (sTNF RI); transforming growth factor beta (TGF-beta); TNF-alpha and TNF-beta; and triggering receptor expressed on myeloid cells 1 (TREM-1).

DNA methylation profiling and calculation of gestational methylation age

DNA methylation profiling and quality control (QC) steps were previously detailed [15]. Briefly, cord blood was obtained post-delivery by a trained staff member. We used the following criteria to select cord blood samples eligible for DNA methylation analysis: 1. having enough DNA amounts (> 50 ng/ul); 2. having been followed; and 3. having available phenotypes (e.g., BMI, morbidity, etc.). Characteristics of mother–newborn pairs included versus excluded from the DNA methylation study were detailed elsewhere [15]. Genome-wide

DNA methylation of 963 samples and 21 replicates were analyzed in the University of Minnesota Genomics Center using the Illumina Infinium MethylationEpic BeadChip. Following sample-level QC, in which 23 samples were removed due to technical or quality issues, we performed a Noob preprocessing and obtained the beta values of > 865,000 CpG sites. No probes were removed to calculate the GAmAge by Haftorn et al. [3], who presented a GAmAge prediction model based on the EPIC array using 176 specific CpGs. All 176 CpGs were available and used for the current analysis (0% missing). The Haftorn GAmAge was calculated for all children with available DNA methylation data from cord blood samples using the "methylclock" R package [16]. Cell type (CD4 +, CD8 +, B cells, monocytes, granulocytes, natural killer cells, and nucleated red blood cells) was estimated using the "minfi" R package [17].

Statistical analysis

This study examined the association of GAmAge with GEAA across different subgroups, and the association of GAmAge with immediate birth outcomes (weight, height, and head circumference) and cord blood inflammation markers. Summary statistics were performed to compare newborns' demographic and clinical characteristics across GAmAge quintiles using the Chi-square test or Fisher exact test for categorical variables and ANOVA for continuous variables. To examine the p of trend of a continuous variable across quintiles, we used the "Kendall tau" correlation test. Pearson correlation test was used to examine the correlation between continuous variables. Linear models were used to examine the association with continuous dependent outcomes adjusting for GEAA and other covariates and to test interactions. The analyses adjusting for GEAA are the equivalent to analyzing associations with epigenetic gestational age acceleration calculated using the residual method. Out of the 940 children included in this analysis, 194 had measurements of inflammation markers. We imputed below the level of detection inflammation markers data using the level of detection divided by the square root of two. We applied log10 transformation on the inflammation markers to achieve normal distribution. For the association with the CpGs in the Haftorn prediction model, we applied quantile normalization on the CpGs generated for the prediction model. The Benjamini and Hochberg false discovery rate (FDR) [18] was applied to correct for multiple testing (inflammation markers and CpGs), with $FDR < 0.05$ denoting significance. All statistical analyses were performed using R (version 4.1; R Foundation for Statistical Computing).

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 4 of 12

Results

Population characteristics

Maternal and child characteristics across GAmAge quintiles are presented in Table 1. Mothers of children in the lowest GAmAge quintile tended to be older (tau = −0.07, $p = 0.0017$), with higher pre-pregnancy BMI (tau = −0.06, $p = 0.012$), a higher proportion of gestational (6.4%) and diabetes mellitus (8.0%).

GEAA and GAmAge: sex, ethnicity/race, and fetal development differences

GAmAge (275.7 ± 12.3 days) and GEAA (270.2 ± 17.3 days) were strongly correlated ($r = 0.86$, $p = 4.70e−281$; Fig. 1a). Within each sex strata, the correlations were stronger among boys ($N = 500$, $r = 0.89$, 95% confidence interval (CI) [0.87, 0.91], $p = 3.18e−171$) compared with girls ($N = 440$, $r = 0.83$, 95% CI [0.80, 0.86], $p = 2.51e−113$). We did not observe a GAmAge X sex interaction (p of interaction = 0.088).

We further examined the correlation between GAmAge and GEAA in ethnicity/race strata (Fig. 1b). Among children to White mothers ($N = 51$), we observed the strongest correlation between GAmAge and GEAA ($r = 0.89$, 95% CI [0.82,0.94], $p = 2.16e−18$), followed by Black/African American descendants ($N = 668$; $r = 0.87$, 95% CI [0.85, 0.89], $p = 2.09e−210$), and Hispanic descendants ($N = 221$; $r = 0.79$, 95% CI [0.74, 0.84], $p = 3.19e−49$). No GAmAge X ethnicity/race interactions were observed in a linear model replacing the correlation model, with GEAA as the dependent variable and GAmAge, ethnicity/race, and the interaction term of GAmAge X ethnicity/race as the independent variables.

Next, we examined subgroups of the delivery week (extremely preterm to very preterm, $N = 25$; moderate to late preterm, $N = 145$; the term ($N = 678$); late to post-term ($N = 92$); extreme preterm and late-term groups combined with the next/previous category due to small sample size). A stronger correlation of GAmAge and GEAA was observed among the extremely preterm to

Table 1 Prenatal and perinatal characteristics across GAmAge quintiles[1]

	Q1: 195 to 269 days; N = 188	Q2: 269 to 275 days; N = 188	Q3: 275 to 280 days; N = 188	Q4: 280 to 284 days; N = 188	Q5: 284 to 310 days; N = 188	P value[2]
Maternal age at delivery, years	28.9 (6.28)	29.5 (6.62)	28.4 (6.88)	26.8 (6,28)	27.8 (6.48)	**0.001**
Maternal pre-pregnancy BMI, kg/m^2	27.7 (6.77)	27.6 (7.04)	27.1 (6.48)	26.1 (6.03)	26.3 (5.80)	0.053
Parity						**0.036**
Nulliparous	85 (45.2%)	58 (30.9%)	82 (43.6%)	95 (50.5%)	99 (52.7%)	
Multiparous	103 (54.8%)	130 (69.1%)	106 (56.4%)	93 (49.5%)	89 (47.3%)	
Maternal ethnicity/race						0.079
Black/African American	133 (70.7%)	137 (72.9%)	125 (66.5%)	138 (73.4%)	135 (71.8%)	
White	19 (10.1%)	7 (3.7%)	10 (5.3%)	5 (2.7%)	10 (5.3%)	
Hispanic	36 (19.1%)	44 (23.4%)	53 (28.2%)	45 (23.9%)	43 (22.9%)	
Maternal diabetes[3]						**4.99e-04**
Non	160 (85.1%)	172 (91.5%)	174 (92.6%)	184 (97.9%)	177 (94.1%)	
Gestational diabetes	12 (6.4%)	10 (5.3%)	8 (4.3%)	0 (0%)	9 (4.8%)	
Pre-gestational diabetes	15 (8.0%)	6 (3.2%)	4 (2.1%)	2 (1.1%)	1 (0.5%)	
Maternal education						0.118
Below college	134 (71.3%)	127 (67.6%)	127 (67.6%)	111 (59.0%)	130 (69.1%)	
College and above	54 (28.7%)	61 (32.4%)	61 (32.4%)	77 (41.0%)	58 (30.9%)	
Delivery method[3]						0.189
C-section	71 (37.8%)	65 (34.6%)	52 (27.7%)	54 (28.7%)	56 (29.8%)	
Vaginal	117 (62.2%)	122 (64.9%)	136 (72.3%)	131 (69.7%)	130 (69.1%)	
Maternal smoking						**6.50e-04**
No	119 (63.3%)	135 (71.8%)	139 (73.9%)	155 (82.4%)	145 (77.1%)	
Yes	69 (36.7%)	53 (28.2%)	49 (26.1%)	33 (17.6%)	43 (22.9%)	
Baby's sex						
Female	83 (44.1%)	85 (45.2%)	101 (53.7%)	84 (44.7%)	87 (46.3%)	0.319
Male	105 (55.9%)	103 (54.8%)	87 (46.3%)	104 (55.3%)	101 (53.7%)	

[1] Mean (SD) for continuous measurements and number (percentage) for categorical

[2] P-value calculated as differences between groups (ANOVA or Chi-square test/Fisher's exact test). Bold values denote $p < 0.05$

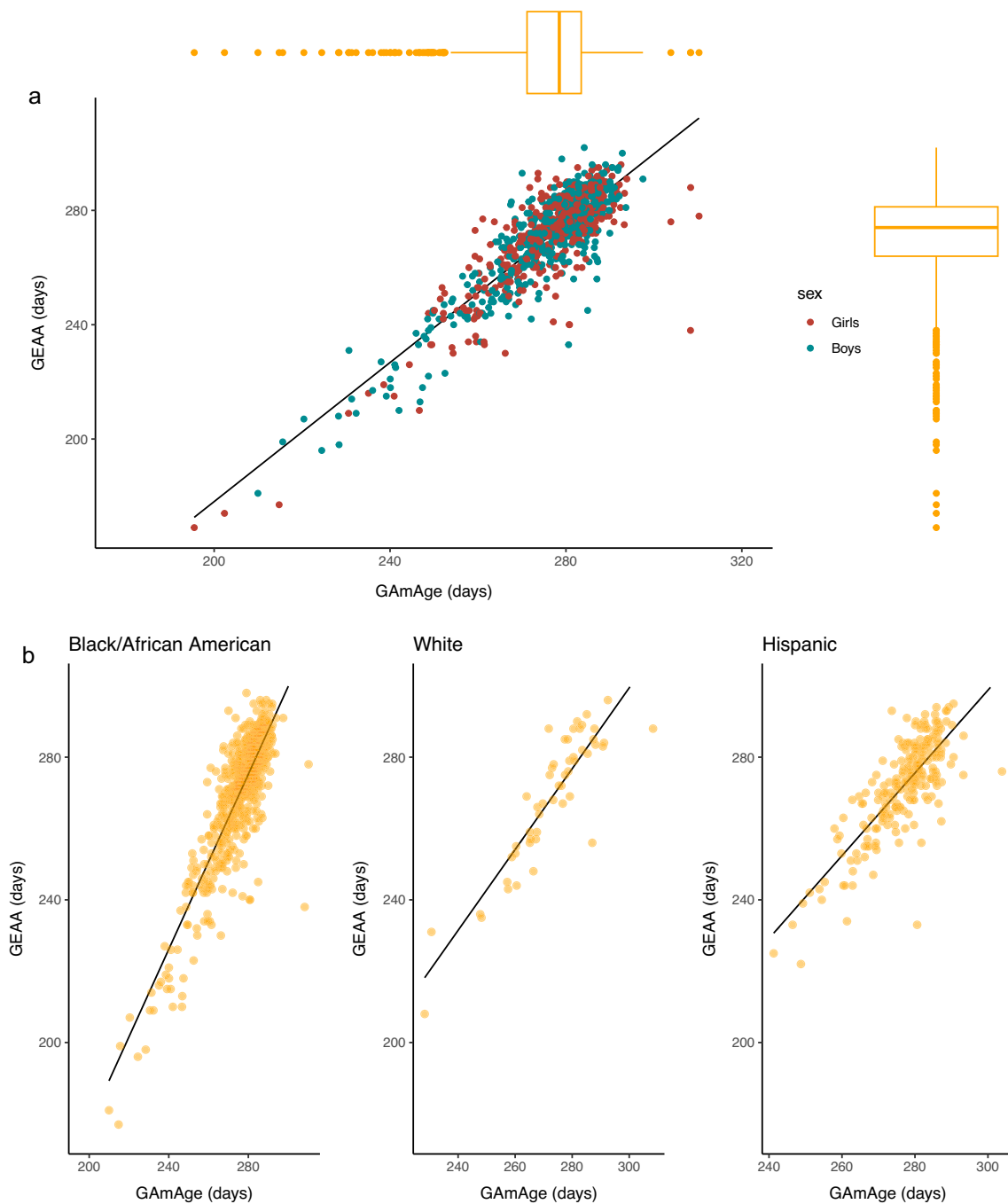[3] Data availability: Maternal diabetes and delivery method: $N = 934$

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 5 of 12

Fig. 1 **a**–**d** GAmAge and GEAA. **a** Correlation between GAmAge and GEAA; **b** correlation between GAmAge and GEAA stratified by ethnicity/race groups; **c** correlation between GAmAge and GEAA stratified by delivery week. Extremely preterm to very preterm delivery week ranged from 24.14 to 31.86 weeks; moderate to late preterm delivery week ranged from 32.14 to 36.83 weeks; term delivery week ranged from 37 to 41 weeks; late to post-term delivery week ranged from 41.14 to 43.14 weeks; **d** correlation between GAmAge and GEAA stratified by fetal growth groups. AGA, appropriate gestational age; GAmAge, gestational methylation age; GEAA, gestational age; LGA, large for gestational age; SGA, small for gestational age

very preterm ($r = 0.91$, 95% CI [0.81, 0.96], $p = 2.07\mathrm{e}{-}10$) compared with the other delivery week groups (moderate to late preterm: $r = 0.48$ 95% CI [0.34, 0.60], $p = 9.56\mathrm{e}{-}10$; term: $r = 0.58$ 95% CI [0.53, 0.63], $p = 2.67\mathrm{e}{-}63$). There was no correlation between GAmAge and gestational age in the late to the post-term group: $r = 0.09$, 95% CI [-0.12, 0.29], $p = 0.40$; Fig. 1c).

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110
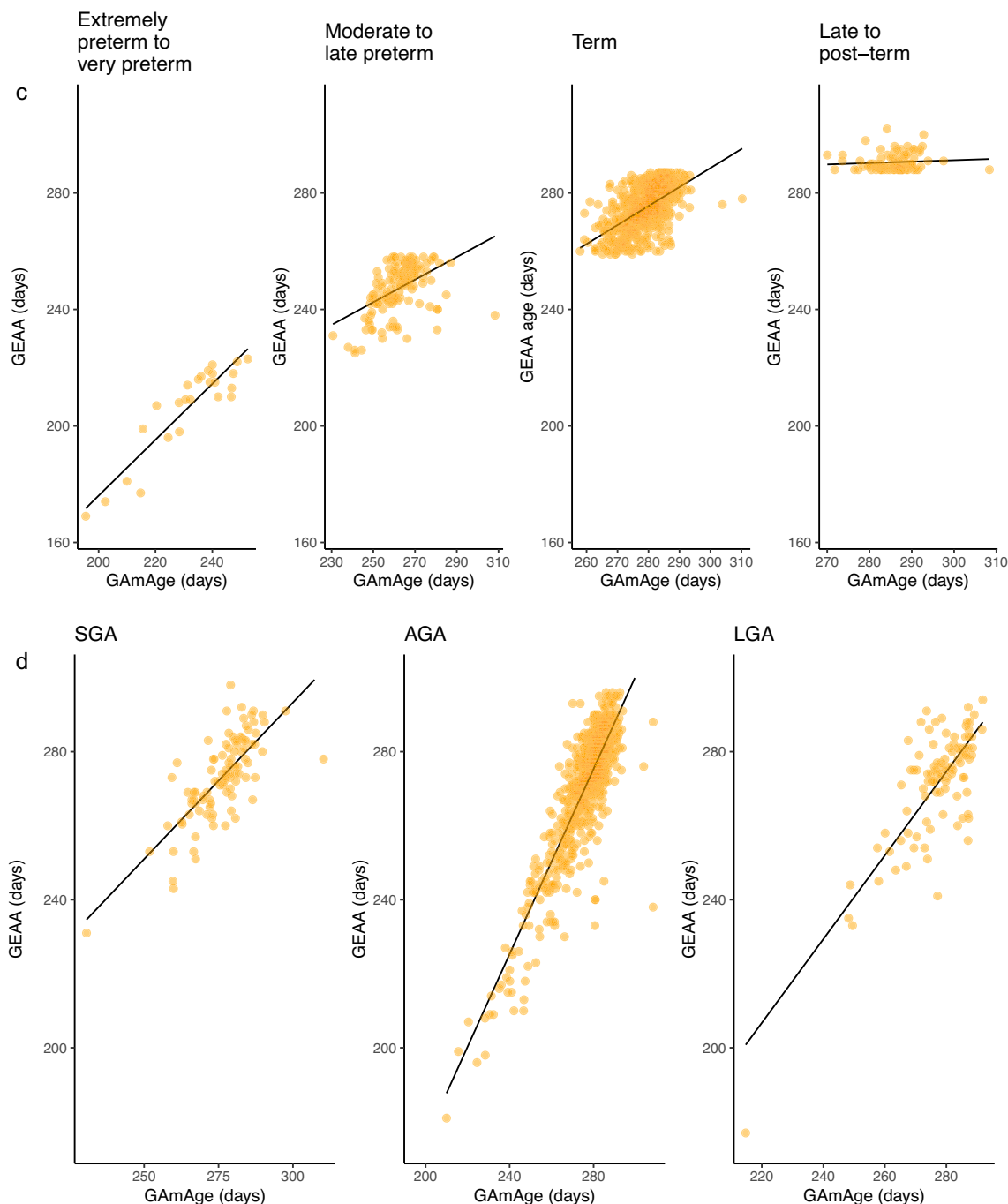
Page 6 of 12

Fig. 1 continued

While there were no differences in mean GAmAge between subgroups of fetal growth (SGA: 276.3 ± 10.1 days, AGA: 275.3 ± 12.8 days, LGA:276.1 ± 11.6 days; $p = 0.69$), the strongest correlation between GAmAge and GEAA (Fig. 1d) was observed within the AGA group ($N = 744$, $r = 0.88$, 95% CI [0.86, 0.90], $p = 6.87e{-}244$). The correlation between GAmAge and GEAA was $r = 0.79$, 95% CI [0.70, 0.85] and $r = 0.75$, 95% CI [0.65, 082] in the LGA ($N = 97$) and SGA ($N = 99$) groups, respectively ($p = 1.38e{-}21$ and $p = 4.40e{-}19$). A significant GAmAge X fetal growth interaction was observed for GAmAge X SGA (compared with AGA: p of interaction $= 4.58e{-}06$; compared with LGA: p of interaction $= 0.014$). The overlap of newborns between delivery week and fetal growth

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 7 of 12

groups is presented in Additional file 1: Table S2. For all delivery week groups, the majority of newborns were AGA.

GAmAge associations with anthropometric measurements

GAmAge was correlated with birth weight ($r = 0.65$, $p = 2.62e-113$), birth length ($r = 0.70$, $p = 1.27e-66$), and birth head circumference ($r = 0.60$ and $p = 2.35e-44$). Following GEAA and sex adjustment using multivariable linear models, the association between GAmAge and each of the anthropometric outcomes remained significant (Fig. 2). No GAmAge X sex interactions were observed in these models for birth weight (p of interaction $= 0.06$), birth length (p of interaction $= 0.14$), and birth length (p of interaction $= 0.82$). Sex stratification of the correlations and further GEAA adjustment are presented in Additional file 1: Figure S1, Additional file 1: Figure S2.

GAmAge associations with cord blood inflammation markers

Out of 27 inflammatory markers measured in cord blood, 9 were significantly correlated with GAmAge (all FDR < 0.05; Fig. 3, Additional file 1: Table S3). Adjustment for GEAA attenuated most of the observed associations; however, GAmAge remained significantly associated with BDNF and MCP-1 (beta $= 0.008$, $p = 0.04$ and beta $= -0.014$, $p = 0.006$, respectively). No sex X GAmAge interactions were observed for BDNF (p of interaction $= 0.30$) and MCP-1 (p of interaction $= 0.53$) multivariable models. Sex stratification of the correlation of GAmAge and either MCP-1 or BDNF is presented in Additional file 1: Figure S3.

Fetal growth stratification showed 8 significant correlations between the GAmAge and inflammatory markers (FDR < 0.05; Additional file 1: Table S4) in the AGA group ($N = 142$). After adjustment for GEAA, only MCP-1 remained significantly correlated with GAmAge (beta $= -0.019$, $p = 0.002$). Among LGA babies ($N = 26$) and SGA babies ($N = 26$), none of the correlations remained significant after correcting for multiple comparisons. Yet, the small sample size for these groups should be considered when examining these correlations.

Since the inflammation panel was restricted to 194 children, the delivery week groups were significantly smaller (extremely preterm to very preterm, $N = 4$; moderate to late preterm, $N = 55$; term, $N = 120$; late to post-term, $N = 15$). Thus, we examined the correlations between GAmAge and inflammation markers in strata of preterm yes/no, as defined in the "Patients and methods" section. In the preterm strata ($N = 59$), 4 inflammation markers were significantly correlated with GAmAge (Additional file 1: Table S5), but only sTNF-RI had FDR < 0.05. Among term babies ($N = 135$), only BDNF was correlated with GAmAge ($r = 0.212$, $p = 0.01$), but correction for multiple comparisons completely attenuated the correlation (FDR $= 0.362$).

GAmAge associations with candidate cord blood inflammation markers accounted for estimated cell type proportions

Next, we wanted to elucidate whether the associations of GAmAge with BDNF and MCP-1 relate to age or cell type (correlation matrix presented in Additional file 1: Figure S4). We used linear models adjusting for estimated cell type (CD4+, CD8+, B cells, monocytes, granulocytes, natural killer cells, and nucleated red blood cells) and GEAA. GAmAge remained significantly associated MCP-1 (beta $= -0.013$, $p = 0.02$), but not with BDNF (beta $= 0.006$, $p = 0.12$).

Sensitivity analysis

As a sensitivity analysis, we calculated GAmAge residuals by regressing GAmAge on GEAA. We then correlated the GAmAge residuals with the three anthropometric birth outcomes, MCP-1, and BDNF (Additional file 1:
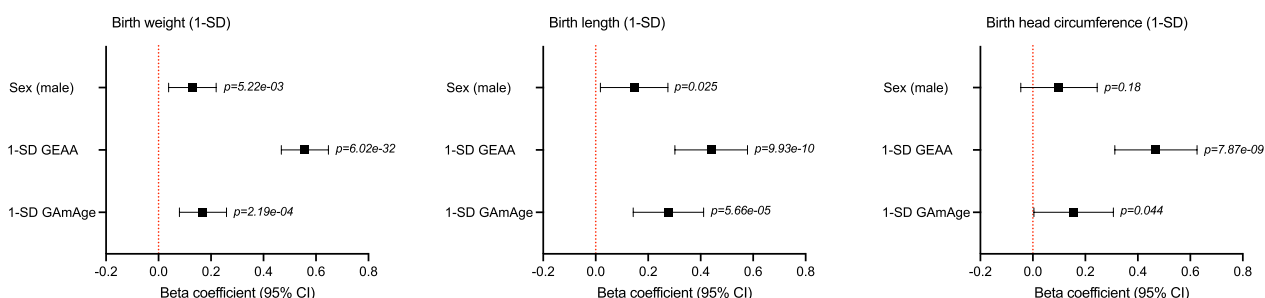

Fig. 2 Association between GAmAge and immediate birth anthropometric outcomes. Beta coefficients from a linear model adjusting for GEAA and sex. Data presented for 1-SD of GAmAge, GEAA, and the outcomes. GAmAge, gestational methylation age; GEAA, gestational age; SD, standard deviation
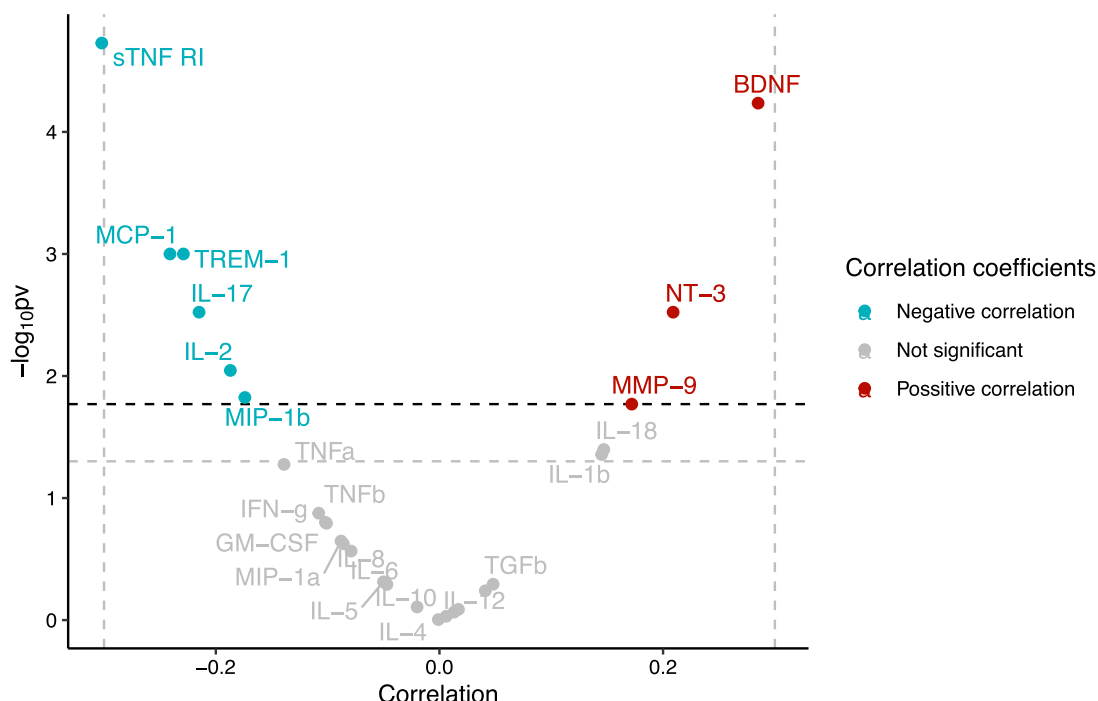
Yaskolka Meir *et al. Clinical Epigenetics* (2024) 16:110

Page 8 of 12

Fig. 3 Volcano plot for inflammation markers associations with GAmAge. Horizontal dashed gray line indicates $p < 0.05$. Horizontal dashed black line indicates FDR < 0.05

Figure S5). This allowed us to show acceleration/deacceleration, which was in accordance with the direction of the main analysis.

As a sensitivity analysis for the multivariable models, we added the following models to examine the association between GAmAge and anthropometric birth outcomes, MCP-1, and BDNF: model 1: adjusted for GEAA, sex, maternal pre-pregnancy BMI, maternal diabetes, and ethnicity/race and model 2 (model 1+cell type proportion). These models yielded similar results to the previous analysis for birth weight, birth length, and MCP-1, and attenuated for birth head circumference and BDNF (Additional file 1: Figure S6).

CpGs contributing to the association of GAmAge and candidate inflammation markers

Finally, we examine the 176 CpGs from the Haftorn GAmAge for the association with either BDNF or MCP-1 to see which CpG was driving the correlation observed above. We observed that 26 CpGs were correlated with BDNF (all FDR < 0.05; Fig. 4a) and 36 with MCP-1 (all FDR < 0.05; Fig. 4b). After adjusting for GEAA and cell type, three CpG remained significantly associated with BDNF (cg12434132: beta $= -1.38$, $p = 0.03$; cg04347477: beta $= -0.76$, $p = 0.04$; cg19640090: beta $= -1.65$, $p = 0.047$; Additional file 1: Table S6), and 12 CpGs remained significantly associated with MCP-1 (top CpG:

cg25975961, beta $= 1.46$, $p = 0.006$; Additional file 1: Table S7).

Discussion

Our analysis confirmed the applicability of a gestational age clock based on the Illumina EPIC array [3] to a mostly Black/African American and Hispanic population. In our study of 940 children, the GAmAge, calculated using cord blood DNA methylation, was directly correlated with GEAA, birth weight, head circumference, birth length, and BDNF levels and inversely correlated with MCP-1, varied by sex and fetal growth status in some of these associations.

Among boys, a stronger correlation between GAmAge and GEAA was observed compared with girls. Previous studies have identified sex differences in DNA methylation. A study that examined over 1000 whole-blood DNA methylation data using the Illumina EPIC array found 396 sex-associated differentially methylated CpG sites, with 74% of these sites showing elevated methylation in females compared with males [19]. Another study used 111 cord blood samples to detect sex differences in DNA methylation (Illumina 450K) and showed that about 3% of the CpGs were differentially methylated between girls and boys at birth, with 83% of the CpGs located on the autosomes being hypermethylated in girls [20]. These observations on cord blood were recently included in a
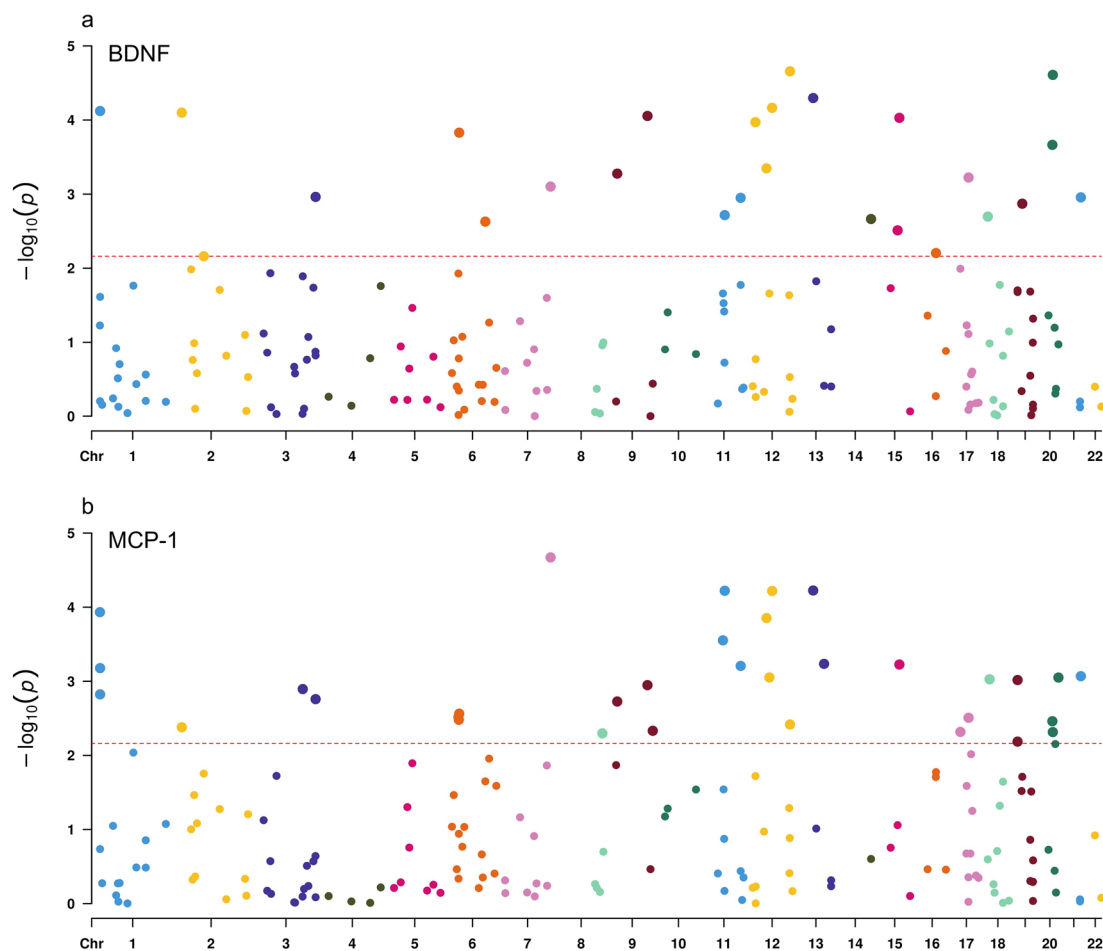
Fig. 4 **a–b** CpGs associated with candidate inflammation markers. **a** BDNF; **b** MCP-1. Dots above the red line are significant CpGs at FDR < 0.05

meta-analysis of 8438 newborns demonstrating sex differences in DNA methylation at 46,979 autosomal CpG sites [21]. Of the 176 CpGs in the EPIC GAmAge prediction formula, 33 overlapped with the above-published CpGs. Thus, the correlation between cord blood DNAm age estimated using this small set of age-related CpGs predicting GEAA may be sex-specific.

We further stratified the association between GAmAge and GEAA by subgroups that may indicate intrauterine growth: delivery week and fetal growth. Among different delivery week groups of preterm, term, or late-term birth, we found that the strongest correlation of GAmAge and GEAA was within the preterm groups, gradually weakening within the next two groups based on GEAA, and completely attenuated in the late-term group. This indicates that the gestational age clock may represent fetal development and fetal stress rather than gestational age. Another interpretation to these findings could be mathematically and not biologically, as the latter group has the narrowest window of GEAA, thus the least variability to correlate against. Exposure to maternal stress in utero

may increase risk for preterm birth [22]. The maternal stress may further lead to epigenetic modifications in the offspring. In a sample of 537 mother–child pairs (122 of which had high stress), high maternal stress was associated with higher offspring regional DNA methylation at the *MEST* gene, compared with infant of the non-stressed mothers [23]. Fetal DNA methylation may also reflect developmental age. This was demonstrated in a study of different human fetal tissues at different gestational age [24]. In that study, the results indicated that between the 9th and the 18th gestational weeks, there is a dynamic of DNA methylation remodeling expressed as gain and loss of methylation. This was consistent across different tissues. DNA methylation at week 22 of gestation were similar to the DNA methylation in adults. When examining the association between GAmAge and GEAA in fetal growth subgroups, we found that the strongest association was observed within the AGA group, reflecting weight appropriate for GEAA. This observation may reflect the epigenetic differences between these groups; higher global placental DNA methylation in a study of

1023 mothers was observed in LGA offspring compared to AGA [25]. Another study found over 500 differentially methylated genes between LGA and AGA in placental samples [26]. While we examined preterm and fetal growth groups separately, these might be linked and have a specific epigenetic signature. Thirty-eight SGA preterm babies (< 32 weeks) were compared with seventy five preterm AGA in a study examining differences in whole-blood DNA methylation of specific genes [27]. While the DNA methylation levels in the SGA group were not significantly different from the AGA group in this study, it may be due to the small sample size.

Ethnicity/race stratification showed that the correlation between GAmAge and GEAA among the White group was the strongest compared to the Black/African American and Hispanic groups. This may be expected as the EPIC GAmAge clock was trained on the European population. While ethnic differences may contribute to differences in DNA methylation [28], we have previously demonstrated that an adult's mAge methylation clock trained on the Chinese population performed well among the White population in the CENTRAL trial [29]. In this current study, all ethnic groups showed significant correlations between GAmAge and GEAA. However, there were significant differences in these observed correlations between the Hispanic and Black/African American ethnic groups. White and Hispanic and White and Black/African American ethnic groups did not differ in the correlation between GAmAge and GEAA. Further investigation should be conducted to elucidate whether these differences are due to the population the GAmAge clock was trained on or differences in methylation patterns in different ethnicities.

Our results confirmed the known correlations between GAmAge and immediate birth anthropometric outcomes and strengthened the potential use of epigenetic clocks to assess developmental features. In a study of 863 children, birth weight, length, and head circumference were all associated with age acceleration, the regressed GAmAge of age [30]. Another study employed the Haftorn clock in 688 mostly White children and found that GAmAge is associated with a developmental delay within the first 3 years of life in the same manner as clinical GEAA [31]. We also found that beyond GEAA or estimated cell type, the methylation-based age was inversely correlated with cord blood MCP-1 levels and directly associated with BDNF. Our findings are aligned with a previous study of 135 infants, where MCP-1 levels in cord blood were significantly higher in very premature neonates when compared with more mature neonates [32]. Additionally, the direct association between GAmAge and BDNF was previously demonstrated to reflect neural maturity in preterm infants and may be driven, at least partially, by GEAA [33]. Notably, prematurity-associated factors such as maternal infection or use of steroids for antenatal lung maturation could also affect the cord blood levels of BDNF and other soluble immune mediators and should be considered in the interpretation of these findings. Another factor that may explain the association of GAmAge and BDNF is the direct correlation of GAmAge with birth weight. Our study did not examine BDNF at birth and obesity but rather a correlation with GAmAge, associated with birth weight and long-term overweight or obesity. In previous work, circulating levels of BDNF in children were previously associated with BMI, as obese children tended to have higher levels of BDNF than lean children [34]. Yet, more studies should be conducted targeting the association between cord blood inflammation markers and obesity.

The 176 CpGs in the Haftorn GAmAge calculation were selected using penalized regression to predict GEAA. Using these GEAA-associated methylation sites, we searched for specific CpGs driving the association between GAmAge and the two candidate inflammation markers BDNF and MCP-1. Out of the 15 CpGs identified to be associated with the inflammation markers, some were previously identified as associated with adverse pregnancy outcomes [35] or newborn birth weight [36]. One cg site (cg24797865) found to be associated with MCP-1 in our analysis, is mapped to the immune gene integrin subunit beta 2 (*ITGB2*) which encodes the beta subunit of LFA-1, a protein which plays a crucial role in leukocyte migration, adhesion, and activation [37]. Since the current analysis cannot determine whether maternal conditions affected the inflammatory status of the newborn, we encouraged further studies to examine the interplay between adverse pregnancy outcomes and other maternal exposures, GAmAge, and inflammation status in the newborn.

There are some limitations to this study. First, the small sample size for some of the subgroups as the extremely preterm to very preterm. This may not only affect our power to detect significant effect size, but also to detect significant interactions. However, this is a secondary analysis and remains exploratory. Cord blood inflammation markers were measured for a small subset of participants. Still, we could detect associations that remained significant after adjustments, even with the given sample size. Our study results may be misinterpreted due to a residual confounding effect. To overcome this, we included both stratification and sensitivity analyses with further adjustments. Beyond its large sample size, this study's strengths are the novel associations described of GAmAge with inflammatory markers.

Yaskolka Meir *et al. Clinical Epigenetics*   (2024) 16:110

Page 11 of 12

Conclusions

The existing GAmAge algorithm from European descents predicts GEAA slightly better in White newborns compared to Black/African American or Hispanic newborns. Our findings support the call for inclusion of diverse populations in developing robust GAmAge algorithm across ethnic/racial groups. A large difference between GAmAge and GEAA might indicate development impact not captured by GEAA alone; therefore, it might capture underline biological stage for fetus development. Moreover, the GAmAge could be utilized as a proxy for fetus maturity beyond GEAA and used to assess immediate health outcomes. Further studies should explore the interplay between GEAA, DNA methylation, and immunological markers in newborns.

Abbreviations

AGA	Appropriate for gestational age
BBC	Boston Birth Cohort
BDNF	Brain-derived neurotrophic factor
CRP	C-reactive protein
FDR	False discovery rate
GAmAge	Gestational methylation age
GEAA	Gestational methylation age
IL	Interleukin
LGA	Large for gestational age
MCP-1	Monocyte chemoattractant protein-1
MIF	Macrophage migration inhibitory factor
MIP-1alpha	Macrophage inflammatory protein 1 alpha
MMP-9	Matrix metalloproteinase
NT	Neurotrophin
RANTES	Regulated on activation, normal T-cell expressed and secreted
SGA	Small for gestational age
sIL-6r alpha	Soluble IL-6 receptor alpha
sTNF RI	Soluble tumor
TGF	Transforming growth factor
TREM-1	Triggering receptor expressed on myeloid cells 1

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-024-01714-x.

Additional file 1.

Data availability statement

The data, data dictionary, and analytical programs for this manuscript are not currently available to the public. However, they can be made available upon reasonable request and after the review and approval of the institutional review board.

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all the study mothers. Institutional Review Boards of the Boston Medical Center and the Johns Hopkins Bloomberg School of Public Health approved the study. The study is also registered on ClinicalTrials.gov (NCT03228875).

Competing interest

The authors declare no competing interests.

Author details

[1]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA. [2]Division of Pediatric Allergy, Immunology and Rheumatology, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [3]Department of Population, Family and Reproductive Health, Center On Early Life Origins of Disease, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD 21205, USA. [4]Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, USA.

References

1. Bohlin J, Håberg SE, Magnus P, Reese SE, Gjessing HK, Magnus MC, et al. Prediction of gestational age based on genome-wide differentially methylated regions. Genome Biol. 2016;17:1–9.
2. Knight AK, Craig JM, Theda C, Baekvad-Hansen M, Bybjerg-Grauholm J, Hansen CS, et al. An epigenetic clock for gestational age at birth based on blood methylation data. Genome Biol. 2016;17:1–11.
3. Haftorn KL, Lee Y, Denault WRP, Page CM, Nustad HE, Lyle R, et al. An EPIC predictor of gestational age and its application to newborns conceived by assisted reproductive technologies. Clin Epigenetics. 2021;13:1–13.
4. Behrman RE, Butler AS. Measurement of fetal and infant maturity. US: Preterm Birth: Causes, Consequences, and Prevention. National Academies Press; 2007.
5. Humberg A, Fortmann I, Siller B, Kopp MV, Herting E, Göpel W, et al. Preterm birth and sustained inflammation: consequences for the neonate. Berlin: Semin Immunopathol. Springer; 2020. p. 451–68.
6. Clapp DW. Developmental regulation of the immune system. Semin Perinatol. 2006;30:69–72.
7. Matoba N, Yu N, Mestan K, Pearson C, Ortiz K, Porta N, et al. Differential patterns of 27 cord blood immune biomarkers across gestational age. Pediatrics [Internet]. 2009;123:1320–8.
8. Yeung EH, Guan W, Zeng X, Salas LA, Mumford SL, De Prado BP, et al. Cord blood DNA methylation reflects cord blood C-reactive protein levels but not maternal levels: a longitudinal study and meta-analysis. Clin Epigenetics [Internet]. 2020;12:1–10. https://doi.org/10.1186/s13148-020-00852-2.
9. van Otterdijk SD, Binder AM, Michels KB. Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy. Epigenetics [Internet]. 2017;12:875–85. https://doi.org/10.1080/15592294.2017.1361592.

10. Pearson C, Bartell T, Wang G, Hong X, Rusk SA, Fu L, et al. Boston Birth Cohort profile: rationale and study design. Precis Nutr. 2022;1: e00011.

11. World Health Organization. International Statistical Classification of Diseases and Related Health Problems: Alphabetical index. Geneva: World Health Organization; 2004.

12. Spong CY. Defining "term" pregnancy: recommendations from the Defining "Term" Pregnancy Workgroup. JAMA. 2013;309:2445–6.

13. Wang X, Zuckerman B, Pearson C, Kaufman G, Chen C, Wang G, et al. Maternal cigarette smoking, metabolic gene polymorphism, and infant birth weight. JAMA. 2002;287:195–202.

14. Wang H, Mueller N, Wang G, Hong X, Chen T, Ji Y, et al. The joint association of small for gestational age and nighttime sleep with blood pressure in childhood. Sci Rep. 2018;8:1–9.

15. Xu R, Hong X, Zhang B, Huang W, Hou W, Wang G, et al. DNA methylation mediates the effect of maternal smoking on offspring birthweight: a birth cohort study of multi-ethnic US mother–newborn pairs. Clin Epigenetics [Internet]. 2021;13:1–13. https://doi.org/10.1186/s13148-021-01032-6.

16. Pelegí-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. methylclock: a Bioconductor package to estimate DNA methylation age. Bioinformatics. 2021;37:1759–60.

17. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30:1363–9.

18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc: Ser B (Methodol). 1995;57:289–300.

19. Grant OA, Wang Y, Kumari M, Zabet NR, Schalkwyk L. Characterising sex differences of autosomal DNA methylation in whole blood using the Illumina EPIC array. Clin Epigenetics. 2022;14:1–16.

20. Yousefi P, Huen K, Davé V, Barcellos L, Eskenazi B, Holland N. Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. BMC Genomics. 2015;16:1–12.

21. Solomon O, Huen K, Yousefi P, Küpers LK, González JR, Suderman M, et al. Meta-analysis of epigenome-wide association studies in newborns and children show widespread sex differences in blood DNA methylation. Mutat Res Rev Mutat Res. 2022;789: 108415.

22. Lilliecreutz C, Larén J, Sydsjö G, Josefsson A. Effect of maternal stress during pregnancy on the risk for preterm birth. BMC Pregnancy Childbirth. 2016;16:1–8.

23. Vidal AC, Neelon SEB, Liu Y, Tuli AM, Fuemmeler BF, Hoyo C, et al. Maternal stress, preterm birth, and DNA methylation at imprint regulatory sequences in humans. Genet Epigenet. 2014;6:GEG-S18067.

24. Slieker RC, Roost MS, van Iperen L, Suchiman HED, Tobi EW, Carlotti F, et al. DNA methylation landscapes of human fetal development. PLoS Genet. 2015;11: e1005583.

25. Dwi Putra SE, Reichetzeder C, Hasan AA, Slowinski T, Chu C, Krämer BK, et al. Being born large for gestational age is associated with increased global placental DNA methylation. Sci Rep. 2020;10:1–10.

26. Chen P-Y, Chu A, Liao W-W, Rubbi L, Janzen C, Hsu F-M, et al. Prenatal growth patterns and birthweight are associated with differential DNA methylation and gene expression of cardiometabolic risk genes in human placentas: a discovery-based approach. Reprod Sci. 2018;25:523–39.

27. Tobi EW, Heijmans BT, Kremer D, Putter H, Delemarre-van de Waal HA, Finken MJJ, et al. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. Epigenetics. 2011;6:171–6.

28. Elliott HR, Burrows K, Min JL, Tillin T, Mason D, Wright J, et al. Characterisation of ethnic differences in DNA methylation between UK-resident South Asians and Europeans. Clin Epigenetics [Internet]. 2022;14:1–17. https://doi.org/10.1186/s13148-022-01351-2.

29. Yaskolka Meir A, Keller M, Bernhart SH, Rinott E, Tsaban G, Zelicha H, et al. Lifestyle weight-loss intervention may attenuate methylation aging: the CENTRAL MRI randomized controlled trial. Clin Epigenetics. 2021;13:1–10.

30. Khouja JN, Simpkin AJ, O'Keeffe LM, Wade KH, Houtepen LC, Relton CL, et al. Epigenetic gestational age acceleration: a prospective cohort study investigating associations with familial, sociodemographic and birth characteristics. Clin Epigenetics. 2018;10:1–9.

31. Polinski KJ, Robinson SL, Putnick DL, Guan W, Gleason JL, Mumford SL, et al. Epigenetic gestational age and the relationship with developmental milestones in early childhood. Hum Mol Genet. 2023;32:1565–74.

32. Otsubo Y, Hashimoto K, Kanbe T, Sumi M, Moriuchi H. Association of cord blood chemokines and other biomarkers with neonatal complications following intrauterine inflammation. PLoS ONE. 2017;12: e0175082.

33. Chouthai NS, Sampers J, Desai N, Smith GM. Changes in neurotrophin levels in umbilical cord blood from infants with different gestational ages and clinical conditions. Pediatr Res. 2003;53:965–9.

34. Gutierrez PTV, Delgado GG, Renteria CT, Orozco ER, Coronado OG. Obesity and overweight influence BDNF serum levels in the pediatric population. Metab Clin Exp. 2020;104: 154105.

35. Antoun E, Kitaba NT, Titcombe P, Dalrymple KV, Garratt ES, Barton SJ, et al. Maternal dysglycaemia, changes in the infant's epigenome modified with a diet and physical activity intervention in pregnancy: Secondary analysis of a randomised control trial. PLoS Med. 2020;17: e1003229.

36. Küpers LK, Monnereau C, Sharp GC, Yousefi P, Salas LA, Ghantous A, et al. Meta-analysis of epigenome-wide association studies in neonates reveals widespread differential DNA methylation associated with birthweight. Nat Commun. 2019;10:1–11.

37. Walling BL, Kim M. LFA-1 in T cell migration and differentiation. Front Immunol. 2018;9:952.

Publisher's Note