

RESEARCH

Open Access



Maximizing insights from longitudinal epigenetic age data: simulations, applications, and practical guidance

Anna Großbach^{1,2*}, Matthew J. Suderman³, Anke Hüls^{4,5,6}, Alexandre A. Lussier^{7,8,9}, Andrew D. A. C. Smith¹⁰, Esther Walton¹¹, Erin C. Dunn^{7,8,9,12} and Andrew J. Simpkin^{1,2}

Abstract

Background Epigenetic age (EA) is an age estimate, developed using DNA methylation (DNAm) states of selected CpG sites across the genome. Although EA and chronological age are highly correlated, EA may not increase uniformly with time. Departures, known as epigenetic age acceleration (EAA), are common and have been linked to various traits and future disease risk. Limited by available data, most studies investigating these relationships have been cross-sectional, using a single EA measurement. However, the recent growth in longitudinal DNAm studies has led to analyses of associations with EA over time. These studies differ in (1) their choice of model; (2) the primary outcome (EA vs. EAA); and (3) in their use of chronological age or age-independent time variables to account for the temporal dynamic. We evaluated the robustness of each approach using simulations and tested our results in two real-world examples, using biological sex and birthweight as predictors of longitudinal EA.

Results Our simulations showed most accurate effect sizes in a linear mixed model or generalized estimating equation, using chronological age as the time variable. The use of EA versus EAA as an outcome did not strongly impact estimates. Applying the optimal model in real-world data uncovered advanced GrimAge in individuals assigned male at birth that decelerates over time.

Conclusion Our results can serve as a guide for forthcoming longitudinal EA studies, aiding in methodological decisions that may determine whether an association is accurately estimated, overestimated, or potentially overlooked.

Keywords Epigenetic age, Longitudinal studies, ALSPAC, Accelerated aging, DNA methylation

*Correspondence:

Anna Großbach
anna.grossbach@universityofgalway.ie

¹ School of Mathematical and Statistical Sciences, University of Galway, Galway, Ireland

² The SFI Centre for Research Training in Genomics Data Science, Galway, Ireland

³ MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

⁴ Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁵ Gangarosa Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁶ Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁷ Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

⁸ Department of Psychiatry, Harvard Medical School, Boston, MA, USA

⁹ Stanley Center for Psychiatric Research, The Broad Institute of Harvard and MIT, Cambridge, MA, USA

¹⁰ Mathematics and Statistics Research Group, University of the West of England, Bristol, UK

¹¹ Department of Psychology, University of Bath, Bath, UK

¹² Department of Sociology, College of Liberal Arts, Purdue University, West Lafayette, IN, USA



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Chronological age, the passage of time since birth, does not fully capture an individual's state or pace of biological aging [1]. Genetics, along with various environments, behaviors, and diseases faced throughout the life course appear to be potent causes of these disparities. While the exact biochemical mechanisms mediating these effects remain largely unknown, there is emerging evidence of a role for epigenetics, a biological process that can induce changes in gene expression without changing the underlying DNA sequence [2]. The adaptable nature of epigenetic modifications is therefore of utmost interest when evaluating the effects that exposures have on health and lifespan.

DNA methylation (DNAm) is the most studied and easiest-to-measure epigenetic modification. DNAm most commonly occurs at cytosine nucleotides that are followed by guanine, known as CpG sites. A decade ago, Horvath [3] as well as Hannum and colleagues [4] introduced algorithms, also known as epigenetic clocks, that identified a set of CpG sites whose methylation state can be used to accurately estimate chronological age. The resulting measure is called epigenetic age (EA). Additional epigenetic clocks have since been published; some aim to best estimate chronological age [5], while others focus on health and mortality [6–10]. Over the past decade, the scientific community has thoroughly investigated EA-related associations, largely indicating relations between advanced EA and adverse health outcomes [11–14].

In most of those studies, EA was investigated in a cross-sectional setting. However, the increasing accessibility of longitudinal epigenetic cohort data [15] has created growing interest in studying EA over time [16–44]. Unfortunately, there are many disparities in modeling strategies across these studies. Our study investigated whether these disparities impacted findings to an extent that might lead to false conclusions, through significantly inflating effects or leaving true associations unnoticed. We evaluated the robustness of methods using simulations. To test our results in real-world data, we applied the same methods in two examples from the Avon Longitudinal Study of Parents and Children (ALSPAC) [45, 46] involving biological sex and birthweight as predictors of longitudinal EA. Our aim was to provide readers with practical guidance in modeling choices that fit their data and maximize insights in epidemiological relations.

Methods

Modeling longitudinal epigenetic age

Numerous approaches exist to model effects that exposures have on longitudinal outcomes like EA. Approaches

differ in (1) the choice of model, (2) outcome, and (3) time variables, as well as (4) the number of repeated measures included in those methods.

The three most frequently applied models are linear mixed effect models (LME) [16–20, 23, 25, 30, 31, 33–36, 38, 41], generalized estimating equations (GEE) [21, 26, 29, 43], and Δ aging [22, 27, 37, 39, 40, 42, 47, 48]. LME models and GEE both analyze repeated measures, like tracking changes in EA over time in an individual's life, and model differences in mean trends between groups, such as those exposed to a factor compared to those who were not. In both methods, the fixed effect (β_2), which stays consistent throughout all measures, as well as the interaction effect (β_4), which accumulates over time, are typically modeled as:

$$\begin{aligned} Outcome_{ij} = & \beta_{0i} + \beta_1 TimeVariable_{ij} + \beta_2 Exposure_i + \beta_3 Sex_i \\ & + \beta_4 (TimeVariable_{ij} \times Exposure_i) + \epsilon_{ij} \end{aligned} \quad (1)$$

where $Outcome_{ij}$ is either EA or epigenetic age acceleration (EAA, the residual resulting from regressing EA on chronological age) measured in individual i , at timepoint j . To accommodate longitudinal changes, both models account for time. Time variables commonly used, summarized in Fig. 1, range from chronological age at the time of measurement to age-independent variables, including the duration in days or years between measurements, numerical ranks (e.g., 1, 2, 3) or factorized values (e.g., F07, F09, F15).

Another method to analyze variations in temporal changes is a two-step approach involving “ Δ aging” (“delta aging”). The Δ aging method is limited to studies with two repeated measures, as it quantifies the difference between measurements, often referred to as the “ Δ aging” score, and then compares these scores between groups, for example using linear regression. Firstly, Δ aging is typically calculated as the difference between a follow-up and a baseline measure of either EA or EAA, with or without adjustment for the duration of time between measurements:

$$\begin{aligned} \Delta Aging &= Outcome_2 - Outcome_1 \\ \Delta Aging (\text{age adjusted}) &= \frac{Outcome_2 - Outcome_1}{Age_2 - Age_1} \end{aligned} \quad (2)$$

where $Outcome_1$ and Age_1 represent EAA or EA and age at the initial measure, while $Outcome_2$ and Age_2 correspond to the follow-up measure. Secondly, models such as linear regression can be used to compare trends in Δ aging between the different groups:

$$\Delta Aging_i = \beta_0 + \beta_1 Exposure_i + \beta_2 Sex_i + \epsilon_i \quad (3)$$

Overview: Time Variables

Terminology	Meaning	Example in ALSPAC		
Timefactor	Categorical measurement rank	“F07”	“F09”	“F15”
Timepoint	Numerical measurement rank	1	2	3
Years between measures	Average number of years between measurements within cohort	0	2	8
Age	Chronological Age at measurement	7.3	9.8	15.2

Fig. 1 Overview of common time variables included in longitudinal epigenetic age studies and their application in ARIES cohort data. Age variables in our example should be interpreted as a single study participant’s age at measurement, since chronological age differs between individuals

where $\Delta Aging_i$ is the above-described difference score between two measures for individual i .

Study population

This study used longitudinal DNAm data generated as part of the Avon Longitudinal Study of Parents and Children (ALSPAC) [45, 46]. Initially, ALSPAC recruited 14,541 pregnant women, resident in Avon, UK, with expected dates of delivery between April 1991 and December 1992. Of the initial 14,541 pregnancies, 14,062 resulted in live births and 13,988 children were alive at 1 year of age. To bolster the initial sample, 1000 additional children were included after the initial participants were approximately 7 years old, increasing the total sample of data collected after age 7–14,901. As part of the Accessible Resource for Integrated Epigenomic Studies (ARIES) [15], a sub-sample of ALSPAC mother–child pairs have undergone genome-wide DNAm analysis. DNAm wet-lab and pre-processing analyses were performed at the University of Bristol as part of the ARIES project [15, 49]. Here, we included up to three within-person DNAm measures, drawn from blood, at age 7, 9, and 15 or 17. A summary of all variables used in our analyses is presented in Table 1. Out of 1,162 individuals, 942 have DNAm measured for at least two of these time points, and 178 individuals have DNAm measured at all three.

Simulation study

To better understand the extent to which methodological choices influence the robustness of results, we compared commonly used methods (introduced under *Modeling Longitudinal Epigenetic Age*). We conducted a series of simulations ($n=1000$), in which we manipulated longitudinal EA data from the ALSPAC cohort, applied all

Table 1 Characteristics of the ARIES sample across all measurement waves.

	Measurement wave	Mean / n	SD	Min	Max	
Sample size	F07	969				
	F09	361				
	F15	958				
Female	Birth	49%				
Age	F07	7.45	0.15	7.08	9.08	
	F09	9.84	0.27	9.33	11.17	
	F15	17.1	1.05	14.58	19.33	
Birthweight [kg]	Birth	3.49	0.49	1.49	5.14	
Epigenetic age	Horvath	F07	8.92	2.92	2.45	29.84
		F09	9.53	3.41	−0.01	29.56
		F15	17.71	4.53	4.28	35.29
PC Horvath	F07	13.00	2.63	7.30	25.1	
	F15	24.5	4.15	11.2	40.7	
GrimAge	F07	12.44	3.04	3.21	23.77	
	F09	12.55	3.48	2.48	30.12	
	F15	20.72	3.28	8.28	34.84	
PC GrimAge	F07	26.5	2.37	18.4	33.3	
	F15	34.9	2.66	26.9	43.2	

models introduced above, and compared the accuracy of effect estimates across methods (Fig. 2).

Simulations were based on longitudinal DNAm data from ARIES [15], condensed into EA measures. To compare EA derived from conceptually different epigenetic clocks, we investigated EA calculations from the Horvath clock [3], GrimAge [9], and their principal component (PC) versions [50] in separate simulation cycles. In our binary exposure simulations, 100 participants were randomly selected as “exposed” ($n=918$ “unexposed”).

Overview: Simulation Study

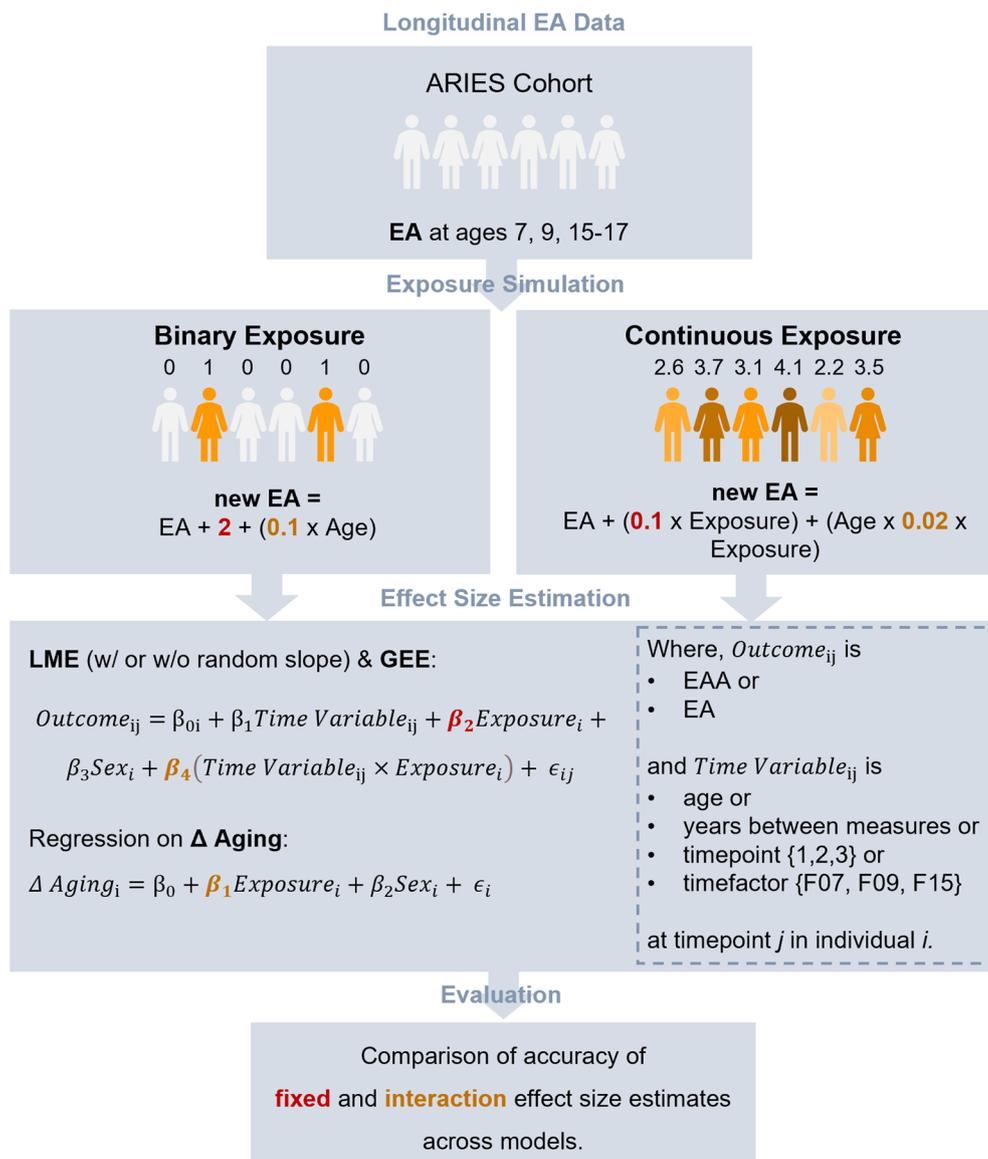


Fig. 2 Overview of simulation study. Simulations were based on longitudinal ARIES cohort data [15] available at ages 7, 9, and 15–17. Epigenetic age (EA) was calculated using the Horvath clock [3] or GrimAge [9] in separate simulations. The original EA measure was then altered based on a simulated exposure. In each binary exposure simulation, a random $n = 100$ individuals had their original EA increased by 2 years (fixed effect), which accumulated by 0.1 year of EA per year of life (interaction effect). In each continuous exposure, all individuals were assigned a value ($N(3.5, 0.5^2)$) which impacted their original EA by 0.1 years, times the level of exposure (fixed effect), and caused an interaction between the exposure and age by 0.02 (interaction effect). In the next step of our simulation, a series of methods was applied to model the simulated effects. Models are linear mixed effect models (LME), generalized estimating equations (GEE), and regression on difference between two epigenetic age (EA) measures (Δ aging). Outcome variables included are epigenetic age acceleration (EAA, residual from regressing EA on age), or EA itself. We ran $n = 1000$ simulations for each exposure type (binary vs. continuous), epigenetic clock (Horvath vs. GrimAge2), and different scopes of data (two measures: Age 7 and 15–17 vs. three measures: Age 7, 9, and 15–17)

To create an association between the exposure and the average outcome, EA was increased by a fixed effect of 2 years in those 100 “exposed” individuals (β_2). Additionally, to create an association between the exposure and change in the average outcome over time, an interaction with chronological age by 0.1 years EA per year of age was included (β_4). In our continuous exposure simulations, all participants were randomly assigned a value ($N(3.5, 0.5^2)$), which affected EA with a fixed effect coefficient of 0.1 (β_2) and an interaction coefficient of 0.02 (β_4).

We then iterated ($n=1000$) through all (i) models, (ii) outcomes and (iii) time variables discussed above, including (iv) different numbers of repeated measures (two timepoints vs. three timepoints). First, we evaluated three models (i): LME models (with and without random slope term), GEE, and regression on Δ aging. Second, we assessed two outcomes within these models (ii): EA an EAA. Third, we investigated four time variables (iii): age, years between measures, numerical ranks {1, 2, 3}, and factorized values {F07, F09, F15}. Δ aging was calculated with and without adjusting for the time between initial and follow-up measure. Fourth, we evaluated the impact of repeated measures (iv) on model accuracy by fitting LME models and GEE with two or three timepoints. Δ aging was limited to only two timepoints.

To compare robustness across methods and variables, we extracted fixed (β_2) and interaction (β_4) effect estimates from all models and evaluated how accurate they met the simulated exposure effect. We measured each model’s performance by comparing whether the 95% confidence interval (CI) contained the simulated effect size. Models resulting in CI that were fully above or below the simulated effect were labeled as “inflated” or “deflated,” while models resulting in CI that contained the simulated effect were defined as performed best.

Real-world example

To apply our simulation results in a real-world example, we used sex and birthweight as accessible biological parameters and examples of binary and continuous predictors of longitudinal EA. Longitudinal EAA or EA at ages 7 and 15–17, derived from the Horvath clock [3], GrimAge [9], and their PC versions [50], were modeled as the outcome. We applied all models and time variables discussed above and included sex and cell type proportions as covariates. Cell counts were estimated using the Houseman algorithm [51] applied to ALSPAC DNAm data with a peripheral blood reference [52].

In LME models and GEE, the fixed (β_2) and interactive (Formula 4: β_9 , Formula 5: β_{10}) effects were estimated as:

Binary exposure (Biological Sex):

$$\begin{aligned} Outcome_{ij} = & \beta_{0i} + \beta_1 TimeVariable_{ij} + \beta_2 Sex_i + \beta_3 NK_{ij} \\ & + \beta_4 Granu_{ij} + \beta_5 Mono_{ij} + \beta_6 CD4T_{ij} + \beta_7 CD8T_{ij} \\ & + \beta_8 Bcell_{ij} + \beta_9 (TimeVariable_{ij} \times Sex_i) + \epsilon_{ij} \end{aligned} \quad (4)$$

Continuous exposure (Birthweight):

$$\begin{aligned} Outcome_{ij} = & \beta_{0i} + \beta_1 TimeVariable_{ij} + \beta_2 Birthweight_i \\ & + \beta_3 Sex_i + \beta_4 NK_{ij} + \beta_5 Granu_{ij} + \beta_6 Mono_{ij} \\ & + \beta_7 CD4T_{ij} + \beta_8 CD8T_{ij} + \beta_9 Bcell_{ij} \\ & + \beta_{10} (TimeVariable_{ij} \times Birthweight_i) + \epsilon_{ij} \end{aligned} \quad (5)$$

where $Outcome_{ij}$ is EA or EAA, measured in individual i , at timepoint j .

Different trends (β_1) in Δ aging (with and without adjusting for time between measures) across groups were modeled using linear regression:

Binary exposure (Biological Sex):

$$\Delta Aging_i = \beta_0 + \beta_1 Sex_i + \epsilon_i \quad (6)$$

Continuous exposure (Birthweight):

$$\Delta Aging_i = \beta_0 + \beta_1 Birthweight_i + \beta_2 Sex_i + \epsilon_i \quad (7)$$

where $\Delta Aging_i$ is the difference score between two measures for individual i .

Results

Simulation study

Figure 3 and Table 2 summarize the effect estimates for a binary exposure for all models and time variables considered, using a subset of two within-person EA measures derived from the Horvath clock. Additional tables, summarizing the complete range of effect estimates, including the continuous exposure and models involving three within-person measures of EA derived from the Horvath clock as well as GrimAge, can be found in the supplemental material (Supplement Figs.1–24 and Tables 1–6). Across all approaches, the choice of time variable had the most substantial impact on the effect estimate bias, which stayed consistent across models (LME models, GEE and Δ aging) and outcome variables (EA, EAA). Including chronological age in the model gave most accurate estimates, while other age-independent time variables led to inflated results.

For both fixed and interaction effects alike, we obtained similar estimates across LME models and GEE, holding the time variable constant. Models that included timepoint as their time variable resulted in a

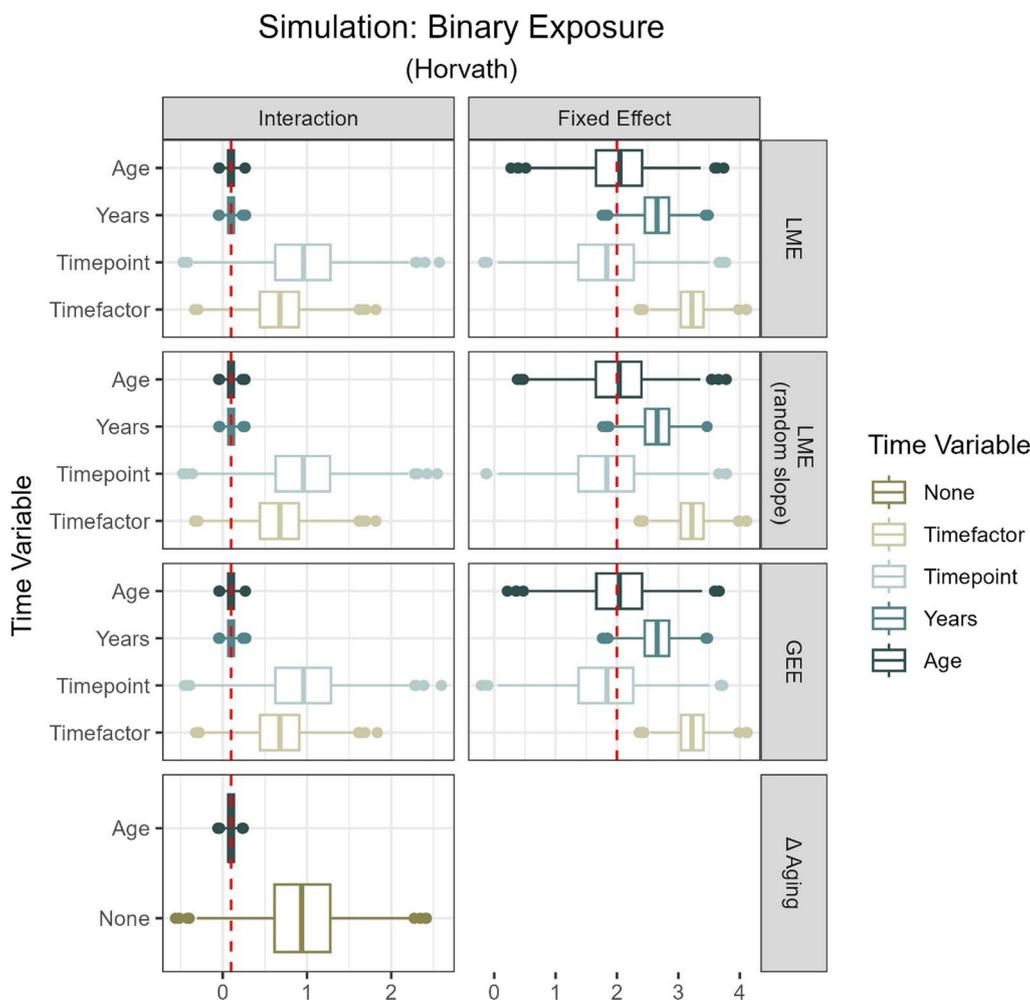


Fig. 3 Summary of $n = 1,000$ simulations from ARIES data (two measurements, age 7 and 15 or 17) [15]. Rows show the distribution (median, 25th and 75th percentile, outliers) of effect size estimates derived from different models and time variables included in those models, respectively. The two columns differentiate between estimates of the interaction term as well as the fixed effect. Simulated effect sizes are marked in red (interaction = 0.1; fixed effect = 2.0). Time variables are chronological age (Age), years between measures (Years), number of measure (Timepoint, i.e., 1, 2), factorized measure (Timefactor, i.e., F07, F15). Models are linear mixed effect models (LME), generalized estimating equations (GEE), and regression on difference between two epigenetic age (EA) measures (Δ aging). All models contained Horvath clock derived EA as outcome [3]

slightly higher proportion of deflated fixed effect estimates than models including age, when using two within-person measures (LME model with random slope, binary simulation: timepoint 5% deflated, age 3% deflated; Supplement Fig. 9). The majority of estimates from models accounting for years between measures or categorical time overestimated the simulated fixed effect (LME model with random slope, binary simulation: > 60% inflated; Supplemental Fig. 9). All estimates from models incorporating time as a categorical variable were inflated.

Regression on Δ aging showed high precision and accuracy for estimating the interaction coefficients, when adjusted for years between measures. Similar

performance in approximating the age-exposure interaction was achieved using LME models and GEE accounting for time through either chronological age or years between measures. Including either categorical time or numerical timepoint led to overestimated effects across all models, resulting in up to 52% inflated interaction estimates (LME model with random slope, binary exposure: timepoint, Supplemental Fig. 9). Models based on three repeated within-person measures resulted in exclusively deflated interaction estimates, when including categorical time as the time variable.

Table 2 Summary of $n = 1000$ simulations from ARIES data (two measurements, age 7 and 15 or 17) [15]

Simulated effect size:		0.1	2.0
Model	Time variable	Interaction effect Estimate (95% CI)	Fixed effect Estimate (95% CI)
LME	Age	0.1 (0, 0.19)	2.02 (0.73, 3.31)
	Years	0.1 (0, 0.2)	2.65 (1.86, 3.44)
	Timepoint	0.94 (0, 1.89)	1.81 (0.28, 3.34)
	Timefactor	0.67 (0, 1.34)	3.22 (2.65, 3.79)
LME (random slope)	Age	0.1 (0.01, 0.18)	2.02 (1, 3.04)
	Years	0.1 (0, 0.2)	2.65 (2.09, 3.22)
	Timepoint	0.94 (0.1, 1.78)	1.81 (0.58, 3.04)
	Timefactor	0.67 (0, 1.34)	3.22 (2.65, 3.79)
GEE	Age	0.1 (-0.01, 0.2)	2.02 (0.87, 3.17)
	Years	0.1 (-0.01, 0.21)	2.65 (2.07, 3.24)
	Timepoint	0.94 (-0.1, 1.99)	1.81 (0.42, 3.19)
	Timefactor	0.67 (-0.07, 1.41)	3.22 (2.7, 3.75)
Δ aging	Age	0.1 (0, 0.2)	
	None	0.94 (-0.02, 1.9)	

The effect size of the simulated binary exposure was 2.0, with an interaction effect of 0.1. Average fixed effect and interaction estimates and 95% confidence intervals (CI) from all models are displayed as columns. Rows represent time variables included in the respective model. Models are linear mixed effect models (LME), generalized estimating equations (GEE), and regression on difference between two epigenetic age (EA) measures (Δ aging). Time variables are chronological age (Age), years between measures (Years), number of measure (Timepoint, i.e., 1, 2), and factorized measure (Timefactor, i.e., F07, F15). All models contained Horvath clock derived EA as outcome [3].

Real-world examples

Supplemental Tables 7–10 as well as Supplemental Figs. 25–28 show the complete range of effect size estimates for both exposures on longitudinal EA, across models, clocks, and time variables. Again, the choice of time variable had the most substantial impact on effect estimates, which stayed consistent across models.

Figure 4A and Supplemental Table 7 show the fixed effect estimates of biological sex on longitudinal EA between the ages 7 and 15–17 years across methods. Different model choices led to very similar estimates of fixed effects, so we compared the choice of time variable within the LME model (random slope) using Horvath EA as the outcome. Individuals assigned male at birth had 0.12 years lower average EA compared to individuals assigned female at birth (95% CI -0.72, 0.47) using chronological age as the time variable, which based on our simulation study is an unbiased estimate. With the age-derived result as a reference, using either years between measures or categorical time resulted in estimates three times larger and in the opposite direction (males were on average 0.36 [95% CI 0.05, 0.66] years

older). The use of timepoint led to estimates almost three times larger in the same direction (males were on average 0.36 [95% CI -1.07, 0.36] years younger). While point estimates differed using GrimAge (likely due to the different training outcome measures of these clocks), the choice of time variable had a similar effect as seen when using the Horvath clock. In both Horvath and GrimAge EA measure-based models, including age or timepoint as a time variable led to much wider confidence intervals compared to using categorical time or time between measures.

Interaction effects between sex and age, leading to an accumulating positive or negative effect on EA over time, are shown in the left column of Fig. 4A and in Supplemental Table 8. The choice of model or outcome led to approximately equal estimates holding the time variable constant. Once again, while effect sizes differed profoundly between clocks, the effect of different time variables was similar using Horvath or GrimAge. Using LME models with age as the time variable, EA in individuals assigned male at birth increased by an extra 0.07 years per year of age, compared to those of individuals assigned female at birth (95% CI 0.02, 0.12). Using either years as the time variable, or Δ aging, led to the same point estimate. However, using either timepoint or categorical time inflated the interaction effect drastically (0.72, 95% CI 0.23, 1.21).

The fixed effect estimates of birthweight on longitudinal EA between the age 7 and 15–17 years of age are shown in the right column of Fig. 4B and in Supplemental Table 9. Due to similar estimates across models and outcome choices, we again compared results for different time variables within LME models (random slope) using Horvath EA as the outcome. Using chronological age as the time variable showed that an increase in birthweight of 1 kg is associated on average with an additional 1.08 years of EA (95% CI: 0.48, 1.69). With the age-derived estimate as a reference (assuming it is unbiased as per our simulation results), using either years between measures or categorical time as the time variable led to an estimate 2.5 times smaller (0.43 [95% CI 0.11, 0.74] years average increase per 1 kg). The use of timepoint increased estimates by 25% (1.3 [95% CI 0.56, 2.03] years average increase per kg). Models using GrimAge measures led to opposite effect estimates, none of which showed statistical significance. Trends across time variables were similar between Horvath and GrimAge EA measure-based models, leading to much wider confidence intervals when including age or timepoint as time variable compared to using categorical time or time between measures.

Estimates of interaction effects between birthweight and age are shown in the left column of Fig. 4B and in

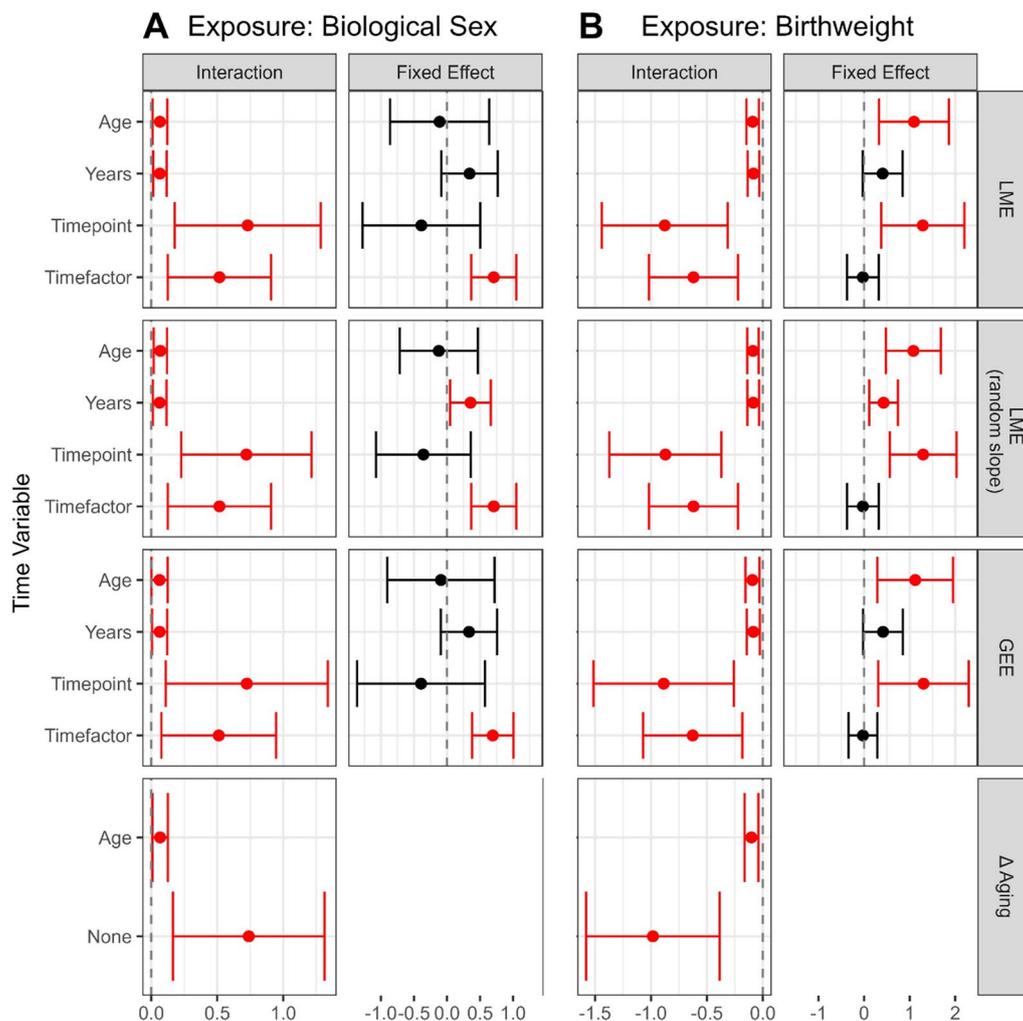


Fig. 4 Age interaction and fixed effect estimates of the effect of biological sex (A) and birthweight (B) on EA over time. Included were two within-person measures (age 7 and 15 or 17) from the ARIES cohort [15]. Rows contain effect size point estimates as well as 95% confidence intervals derived from different models and time variables included in those models, respectively. Effect estimates are based on individuals assigned female at birth as reference. Significant estimates are marked in red ($p < 0.05$). Models are linear mixed effect models (LME), generalized estimating equations (GEE), and regression on difference between two epigenetic age (EA) measures (Δ aging). Time variables are chronological age (Age), years between measures (Years), number of measure (Timepoint, i.e., 1, 2), and factorized measure (Timefactor, i.e., F07, F15). Models contained EA measures derived from the Horvath clock [3]. All models were corrected for cell-type proportion, while models estimating the effect of birthweight additionally account for biological sex

Supplementary Table 10. Again, the selection of either the model or the outcome led to similar estimates when keeping the time variable constant. Using LME models with age as the time variable, EA decreased on average by 0.09 years per year of age for each kg increase in birthweight (95% CI $-0.14, -0.04$). Using years between measures, or age-adjusted Δ aging, as the time variable led to a similar point estimate. Regressing on non-age-adjusted Δ aging or including timepoint or categorical time in the model inflated the interaction effect by a factor of 10 ($-0.87, 95\% \text{ CI } -1.37, -0.37$).

Discussion

The findings of this study present valuable insight into the intricacies of modeling associations with longitudinal epigenetic age, emphasizing the critical influence of different exposure and time variables on the robustness of effect estimates and conclusions.

Simulation and real-world examples

All simulation and real-world analyses showed consistent estimates across models and outcome variables, while the choice of time variable significantly impacted

the accuracy and precision of effect estimates. In simulations, including chronological age as the time variable in an LME model or GEE led to the highest number of correct fixed effect and interaction effect estimates. Hence, models accounting for chronological age also produced the most robust results in our real-world analysis of biological sex and birthweight. When estimating interaction effects, i.e., the accumulating effect that an exposure has on the pace of longitudinal EA, models including timepoint or categorical time had the highest number of incorrect effect size estimates (inflated or deflated) in simulations. We observed a difference of similar magnitude in interaction effect estimates in both real-world analyses compared to estimates from models using chronological age. Including years between measures or categorical time led to >60% inflated fixed effect estimates and narrow confidence intervals in our binary exposure simulation. The resulting bias led to what appears like a false positive finding in the real-world analysis of sex.

Application examples: epidemiological conclusion

The two application examples not only lend support to our simulation findings, but also yield several novel epidemiological findings. First, our study suggests that on average, individuals assigned male at birth have higher GrimAge than individuals assigned female at birth (2.83 additional years of EA, 95% CI 2.41, 3.26). However, individuals assigned male at birth showed a decelerated rate of GrimAge over the time observed. More specifically, individuals assigned male at birth showed 0.08 years EA decrease per year of life on average (95% CI 0.05–0.11 decrease per year), compared to individuals assigned female at birth. Both effects remained consistent and comparable in magnitude when using EA estimates from the PC GrimAge clock. Biological aging studies focusing on other molecular biomarkers to investigate sex-specific differences support the lower baseline aging in women [53, 54]. Although cross-sectional studies across different age groups, ancestries, and epigenetic clocks have identified an association between biological sex and EA [34, 55, 56], no prior study, to our knowledge, has shown that the effect changes longitudinally. While an age interaction effect was detected with Horvath's clock, it did not remain consistent when using its PC version. The differences in results between these clocks are likely due to their conceptual distinctions: the Horvath clock is designed to best predict chronological age, whereas the GrimAge clock is trained to predict health outcomes and lifespan. The variations in effect direction and magnitude observed across different clocks in our analysis highlight the importance of carefully selecting a clock that is

appropriate for the specific study design and population before conducting the analysis [57].

Second, our results suggest a positive association between birthweight and longitudinal EA based on Horvath EA measures (1.08 additional years of EA per increase in kg birthweight, 95% CI 0.48, 1.69), as well as a negative birthweight/age interaction effect over time (–0.09 years decrease in EA per year per increase in kg birthweight, 95% CI –0.14, –0.04). These results indicate that children born with higher birthweight have higher EA on average, while their pace of EA appears to slow down over the period measured, compared to children born with lower birthweight. However, the pattern does not remain consistent when using EA estimates from the PC Horvath clock. Results from models using EA estimates from PC GrimAge indicate lower baseline EA for individuals with higher birthweight, which aligns with other studies that have examined the relationship between birthweight and EA, and generally report that lower birthweight is associated with higher EA [58–61]. However, Simpkin et al. [34] found that birthweight was positively associated with EA at age 7, but negatively associated at age 17. Most studies analyzed data cross-sectionally, included dichotomized birthweight rather than continuous, and were based on EA measures in adult cohorts. Furthermore, associations were predominantly identified in individuals assigned male at birth [59, 60] or in male-dominated cohorts [58]. Future large-scale studies are needed to clarify longitudinal relationships and explore the effect birthweight has on EA throughout the life course.

Recommendations for future studies

As we recognize the importance of methodological choices, this section offers recommendations and guidance for researchers embarking on similar investigative paths. We highly recommend the use of LME models or GEE, including chronological age as the time variable, for studies working with repeated EA measures. These approaches improve precision and accuracy of fixed and interaction effect estimates. Alternatively, research evaluating interaction effects based on only two within-person measures can yield similar validity by using linear regression on age-adjusted Δ aging. We acknowledge that due to limited data collection or access, it is not always possible to implement the best-possible model. In cases where chronological age is not accessible and the cohort under study was measured in synchronized waves, we recommend the use of numerical timepoint to get more accurate fixed effect estimates. The effect size might appear slightly attenuated compared to models including age but is less susceptible to false positives. In studies aimed at exploring interaction effects, it is advisable

to opt for years between measures instead of timepoint in the absence of chronological age. Factorized categorical time should be avoided due to its potential to introduce bias in fixed effect and interaction effect estimates, especially in studies incorporating more than two within-person measurements. When EA itself is not the preferred outcome, we strongly recommend using EAA, the residual from regressing epigenetic age on chronological age, instead of difference scores (EA minus chronological age). Residualizing outcome variables is a well-established practice in epigenetics epidemiology [62–64] to capture part of the EA estimate that is not associated with age, making it more informative about factors beyond age that may impact biological aging. Although the underlying interpretation of results remains consistent, the presentation differs. An effect on longitudinal EA indicates how a factor influences the overall biological age estimate over time, while an effect on longitudinal EAA reflects how this factor impacts the deviation between age and EA over time.

Strengths and limitations

The main strength of our study lies in its comprehensive evaluation of longitudinal EA models, thereby offering valuable insight to direct future research toward more reliable results. The increasing accessibility of repeated DNAm measures and growing interest in comprehending the effects of exposures on EA and its subsequent influence on health have shown the necessity for guidelines to address the issue robustly. A further strength of our work is the incorporation of both simulation and application in two real-world examples, which support the credibility and applicability of our results. One limitation worth noting is that the real-world examples were modeled using EA measured at only two time points, which might have impacted the comprehensiveness of our epidemiological findings. However, our simulation based on three repeated EA measures suggests results from models including age as time variable are consistent across different scopes of data. It also reflects the reality faced by many researchers, as most studies typically have access to only two measures. Second, we limited our methodological evaluation to common models and time variables found in recent literature [16–44] and excluded rare and simplistic approaches. While there are certainly various other methods available, we assume that our study has addressed the applications most pertinent to most epidemiological studies. We furthermore limited our study to four epigenetic clocks, instead of evaluating all available options. The primary goal of this work is to explore the discrepancies that arise from using different methods to analyze associations with epigenetic age, independent of

the specific clock employed. Our selection includes both first- and second-generation clocks, ranging from those designed to predict chronological age to those predicting mortality risk, as well as their principal component versions, which are increasingly favored in longitudinal studies. The choice of an epigenetic clock is generally context-dependent, influenced by study design and available data [57]. Our work aims to provide methodological guidance for researchers after they have selected the appropriate clock for their study.

Conclusions

In conclusion, our study presents a comprehensive evaluation of various methods utilized in modeling exposure effects on EA over time. Through a combination of simulation and real-world analyses, we have demonstrated that the methodological decisions made in longitudinal EA modeling significantly impact the reliability of effect estimates. Findings highlight that LME models or GEE, using chronological age as the time variable, are the optimal approach. Moreover, recognizing the constraints faced by some studies regarding data availability, we have provided practical recommendations to accommodate such limitations. Our thorough assessment serves as a valuable resource for guiding future epidemiological epigenetic aging research endeavors. By optimizing methodological approaches based on the insights from our study, researchers can enhance the depth and accuracy of their investigations, ultimately advancing our understanding of the complex interplay between exposures and epigenetic aging processes.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-024-01784-x>.

Additional file 1.

Acknowledgements

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

Author contributions

Conception and design were done by all authors. Data analysis was done by AG, MJS, and AJS. Revision of the article was done by all authors. All authors read and approved the final manuscript.

Funding

The UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University of Bristol provide core support for ALSPAC. This publication is the work of the authors, who will serve as guarantors for the contents of this paper. A comprehensive list of grants funding is available on the ALSPAC website: <http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>. This research was funded by Science Foundation Ireland through the SFI Centre for Research Training in Genomics Data Science under

Grant number 18/CRT/6214. It was also supported in part by the EU's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant H2020-MSCA-COFUND-2019-945385. This work was supported by the National Institute of Mental Health of the National Institutes of Health (Grant Number R01MH113930 awarded to ECD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. EW is funded by the European Union's Horizon 2020/Europe Research and Innovation Programme (EarlyCause Grant Agreement No 848158) and by UK Research and Innovation (UKRI) under the UK government's Horizon Europe/ERC Frontier Research Guarantee (BrainHealth, Grant Number EP/Y015037/1). AAL is supported by a fellowship from the Canadian Institute of Health Research.

Availability of data and materials

ALSPAC data are available on request at <http://www.bristol.ac.uk/alspac/researchers/access/>. Details of all the data are available through a fully searchable data dictionary and variable search tool on the ALSPAC study website: <http://www.bristol.ac.uk/alspac/researchers/our-data/>.

Declarations

Ethics approval and consent to participate

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 27 May 2024 Accepted: 15 November 2024

Published online: 20 December 2024

References

- Elliott ML, Caspi A, Houts RM, Ambler A, Broadbent JM, Hancox RJ, et al. Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. *Nat Aging*. 2021;1:295–308.
- Aristizabal MJ, Anreiter I, Halldorsdottir T, Odgers CL, McDade TW, Goldenberg A, et al. Biological embedding of experience: a primer on epigenetics. *Proc Natl Acad Sci*. 2020;117:23261–9.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013. <https://doi.org/10.1186/gb-2013-14-10-r115>.
- Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada SV, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49:359–67.
- Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging (Albany NY)*. 2018;10:1758–75.
- Belsky DW, Caspi A, Corcoran DL, Sugden K, Poulton R, Arseneault L, et al. DunedinPACE, A DNA methylation biomarker of the pace of aging. *Elife*. 2022;11:1–26.
- Belsky DW, Caspi A, Arseneault L, Baccarelli A, Corcoran D, Gao X, et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. *Elife*. 2020;9:1–56.
- Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10:573.
- Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)*. 2019;11:303.
- McEwen LM, O'Donnell KJ, McGill MG, Edgar RD, Jones MJ, Maclsaac JL, et al. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proc Natl Acad Sci U S A*. 2020;117:23329–35.
- Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol*. 2015;16:1–12.
- Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, Montgomery GW, et al. Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome Med*. 2019;11:1–11.
- Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)*. 2016;8:1844–65.
- Marini S, Davis KA, Soare TW, Zhu Y, Suderman MJ, Simpkin AJ, et al. Adversity exposure during sensitive periods predicts accelerated epigenetic aging in children. *Psychoneuroendocrinology*. 2020;113:104484.
- Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H, et al. Data resource profile: accessible resource for integrated epigenomic studies (ARIES). *Int J Epidemiol*. 2015;44:1181–90.
- Rentscher KE, Bethea TN, Zhai W, Small BJ, Zhou X, Ahles TA, et al. Epigenetic aging in older breast cancer survivors and non-cancer controls: preliminary findings from the Thinking and Living with Cancer (TLC) Study. *Cancer*. 2023;129:2741–53.
- Schoepf IC, Esteban-Cantos A, Thorball CW, Rodés B, Reiss P, Rodríguez-Centeno J, et al. Epigenetic ageing accelerates before antiretroviral therapy and decelerates after viral suppression in people with HIV in Switzerland: a longitudinal study over 17 years. *Lancet Healthy Longev*. 2023;4:e211–8.
- Kho M, Wang YZ, Chaar D, Zhao W, Ratliff SM, Mosley TH, et al. Accelerated DNA methylation age and medication use among African Americans. *Aging (Albany NY)*. 2021;13:14604–29.
- Nwanaji-Enwerem JC, Van Der LL, Kogut K, Eskenazi B, Holland N, Dearnorff J, et al. Maternal adverse childhood experiences before pregnancy are associated with epigenetic aging changes in their children. *Aging (Albany NY)*. 2021;13:25653.
- Shiau S, Brummel SS, Kennedy EM, Hermetz K, Spector SA, Williams PL, et al. Longitudinal changes in epigenetic age in youth with perinatally acquired HIV and youth who are perinatally HIV-exposed uninfected. *AIDS*. 2021;35:811–9.
- Xiao C, Beitleer JJ, Peng G, Levine ME, Conneely KN, Zhao H, et al. Epigenetic age acceleration, fatigue, and inflammation in patients undergoing radiation therapy for head and neck cancer: a longitudinal study. *Cancer*. 2021;127:3361–71.
- Breen M, Nwanaji-Enwerem JC, Karrasch S, Flexeder C, Schulz H, Waldenberger M, et al. Accelerated epigenetic aging as a risk factor for chronic obstructive pulmonary disease and decreased lung function in two prospective cohort studies. *Aging (Albany NY)*. 2020;12:16539–54.
- Nwanaji-Enwerem JC, Nwanaji-Enwerem U, Van Der Laan L, Galazka JM, Redeker NS, Cardenas A. A longitudinal epigenetic aging and leukocyte analysis of simulated space travel: the mars-500 mission. *Cell Rep*. 2020;33:108406.
- Sehl ME, Rickabaugh TM, Shih R, Martinez-Maza O, Horvath S, Ramirez CM, et al. The effects of anti-retroviral therapy on epigenetic age acceleration observed in HIV-1-infected adults. *Pathog Immun*. 2020;5:291.
- Marioni RE, Suderman M, Chen BH, Horvath S, Bandinelli S, Morris T, et al. Tracking the epigenetic clock across the human life course: a meta-analysis of longitudinal cohort data. *J Gerontol Ser A*. 2019;74:57–61.
- Nannini DR, Joyce BT, Zheng Y, Gao T, Liu L, Yoon G, et al. Epigenetic age acceleration and metabolic syndrome in the coronary artery risk development in young adults study. *Clin Epigenetics*. 2019;11:1–9.
- Wolf EJ, Logue MW, Morrison FG, Wilcox ES, Stone A, Schichman SA, et al. Posttraumatic psychopathology and the pace of the epigenetic clock: a longitudinal investigation. *Psychol Med*. 2019;49:791–800.
- Sehl ME, Breen EC, Shih R, Chen L, Wang R, Horvath S, et al. Increased rate of epigenetic aging in men living with HIV prior to treatment. *Front Genet*. 2022;12:796547.
- Binder AM, Corvalan C, Mericq V, Pereira A, Santos JL, Horvath S, et al. Faster ticking rate of the epigenetic clock is associated with faster pubertal development in girls. *Epigenetics*. 2018;13:85–94.
- Degerman S, Josefsson M, Nordin Adolffsson A, Wennstedt S, Landfors M, Haider Z, et al. Maintained memory in aging is associated with young epigenetic age. *Neurobiol Aging*. 2017;55:167–71.

31. Grant CD, Jafari N, Hou L, Li Y, Stewart JD, Zhang G, et al. A longitudinal study of DNA methylation as a potential mediator of age-related diabetes risk. *GeroScience*. 2017;39:475–89.
32. Sehl ME, Henry JE, Storniolo AM, Ganz PA, Horvath S. DNA methylation age is elevated in breast tissue of healthy women. *Breast Cancer Res Treat*. 2017;164:209–19.
33. Simpkin AJ, Howe LD, Tilling K, Gaunt TR, Lyttleton O, McArdle WL, et al. The epigenetic clock and physical development during childhood and adolescence: longitudinal analysis from a UK birth cohort. *Int J Epidemiol*. 2017;46:549–58.
34. Simpkin AJ, Hemani G, Suderman M, Gaunt TR, Lyttleton O, McArdle WL, 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;25:191–201.
35. Zheng Y, Joyce BT, Colicino E, Liu L, Zhang W, Dai Q, et al. Blood epigenetic age may predict cancer incidence and mortality. *EBioMedicine*. 2016;5:68–73.
36. Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. *Int J Epidemiol*. 2015;44:1388–96.
37. Vetter VM, Kalies CH, Sommerer Y, Spira D, Drewelies J, Regitz-Zagrosek V, et al. Relationship between 5 epigenetic clocks, telomere length, and functional capacity assessed in older adults: cross-sectional and longitudinal analyses. *J Gerontol Ser A*. 2022;2022:1–10.
38. Carter A, Bares C, Lin L, Glover B, Bowden M, Zucker RA, et al. Sex-specific and generational effects of alcohol and tobacco use on epigenetic age acceleration in the Michigan longitudinal study. *Drug Alcohol Depend Rep*. 2022;4:100077.
39. Copeland WE, Shanahan L, McGinnis EW, Aberg KA, Oord EJCG. Early adversities accelerate epigenetic aging into adulthood: a 10-year, within-subject analysis. *J Child Psychol Psychiatry*. 2022;63:1308–15.
40. Esteban-Cantos A, Montejano R, Rodríguez-Centeno J, Saiz-Medrano G, De Miguel R, Barruz P, et al. The journal of infectious diseases longitudinal changes in epigenetic age acceleration in aviremic human immunodeficiency virus-infected recipients of long-term antiretroviral treatment. *J Infect Dis*. 2022;225:287–94.
41. Fraszczyk E, Thio CHL, Wackers P, Martijn, Dollé ET, Bloks VW, et al. DNA methylation trajectories and accelerated epigenetic aging in incident type 2 diabetes. *GeroScience*. 2022;44:2671–84.
42. Iftimovici A, Kebir O, Jiao C, He Q, Krebs M-O, Chaumette B. Dysmaturational longitudinal epigenetic aging during transition to psychosis. *Schizophr Bull Open*. 2022. <https://doi.org/10.1093/schizbullopen/sgac030>.
43. McGill MG, Pokhvisneva I, Clappison AS, McEwen LM, Beijers R, Tollenaar MS, et al. Maternal prenatal anxiety and the fetal origins of epigenetic aging. *Biol Psychiatry*. 2022;91:303–12.
44. Pinho GM, Martin JGA, Farrell C, Haghani A, Zoller JA, Zhang J, et al. Hibernation slows epigenetic ageing in yellow-bellied marmots. *Nat Ecol Evol*. 2022;6:418–26.
45. Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. Cohort profile: the 'children of the 90s'-the index offspring of the Avon longitudinal study of parents and children. *Int J Epidemiol*. 2013;42:111–27.
46. Fraser A, Macdonald-wallis C, Tilling K, Boyd A, Golding J, Davey Smith G, et al. Cohort profile: the Avon longitudinal study of parents and children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42:97–110.
47. Breen EC, Sehl ME, Shih R, Langfelder P, Wang R, Horvath S, et al. Accelerated aging with HIV begins at the time of initial HIV infection. *Science*. 2022;25:104488.
48. Sehl ME, Carroll JE, Horvath S, Bower JE. The acute effects of adjuvant radiation and chemotherapy on peripheral blood epigenetic age in early stage breast cancer patients. *npj Breast Cancer*. 2020;6:1–5.
49. Min JL, Hemani G, Smith GD, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics*. 2018;34:3983–9.
50. Higgins-Chen AT, Thrush KL, Wang Y, Minter CJ, Kuo PL, Wang M, et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nat Aging*. 2022;2:644.
51. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinform*. 2012;13:1–16.
52. Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén SE, Greco D, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE*. 2012;7:e41361.
53. Hägg S, Jylhävä J. Sex differences in biological aging with a focus on human studies. *Elife*. 2021;10:1–27.
54. Nakamura E, Miyao K. Sex differences in human biological aging. *J Gerontol Ser A Biol Sci Med Sci*. 2008;63:936–44.
55. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol*. 2016;17:1–23.
56. Engelbrecht HR, Merrill SM, Gladish N, MacIsaac JL, Lin DTS, Ecker S, et al. Sex differences in epigenetic age in Mediterranean high longevity regions. *Front Aging*. 2022;3:1–16.
57. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, et al. DNA methylation aging clocks: challenges and recommendations. *Genome Biol BioMed Central*. 2019. <https://doi.org/10.1186/s13059-019-1824-y>.
58. Quinn EB, Hsiao CJ, Maisha FM, Mulligan CJ. Low birthweight is associated with epigenetic age acceleration in the first 3 years of life. *Evol Med Public Heal*. 2023;11:251–61.
59. Kuzawa CW, Ryan CP, Adair LS, Lee NR, Carba DB, MacIsaac JL, et al. Birth weight and maternal energy status during pregnancy as predictors of epigenetic age acceleration in young adults from metropolitan Cebu, Philippines. *Epigenetics*. 2022;17:1535.
60. van Lieshout RJ, McGowan PO, de Vega WC, Savoy CD, Morrison KM, Saigal S, et al. Extremely low birth weight and accelerated biological aging. *Pediatrics*. 2021. <https://doi.org/10.1542/peds.2020-001230>.
61. Mathewson KJ, McGowan PO, de Vega WC, Morrison KM, Saigal S, Van Lieshout RJ, et al. Cumulative risks predict epigenetic age in adult survivors of extremely low birth weight. *Dev Psychobiol*. 2021;63:e22222.
62. Min JL, Hemani G, Hannon E, Dekkers KF, Castillo-Fernandez J, Luijk R, et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat Genet*. 2021;53:1311–21.
63. Zhu Y, Simpkin AJ, Suderman MJ, Lussier AA, Walton E, Dunn EC, et al. A structured approach to evaluating life-course hypotheses: moving beyond analyses of exposed versus unexposed in the -omics context. *Am J Epidemiol*. 2021;190:1101–12.
64. Smith BJ, Smith ADAC, Dunn EC. Statistical modeling of sensitive period effects using the structured life course modeling approach (SLCMA). *Curr Top Behav Neurosci*. 2022;53:215–34.

Publisher's Note

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