American Journal of Epidemiology

Vol. 192, No. 5
https://doi.org/10.1093/aje/kwad025
Advance Access publication:
January 30, 2023

Practice of Epidemiology

Use of Correct and Incorrect Methods of Accounting for Age in Studies of Epigenetic Accelerated Aging: Implications and Recommendations for Best Practices

Nancy Krieger*, Jarvis T. Chen, Christian Testa, Ana Diez Roux, Kate Tilling, Sarah Watkins, Andrew J. Simpkin, Matthew Suderman, George Davey Smith, Immaculata De Vivo, Pamela D. Waterman, and Caroline Relton

* Correspondence to Dr. Nancy Krieger, Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, 677 Huntington Avenue, Kresge 717, Boston, MA 02115 (e-mail: nkrieger@hsph.harvard.edu).

Initially submitted March 31, 2022; accepted for publication January 27, 2023.

Motivated by our conduct of a literature review on social exposures and accelerated aging as measured by a growing number of epigenetic "clocks" (which estimate age via DNA methylation (DNAm) patterns), we report on 3 different approaches in the epidemiologic literature—1 incorrect and 2 correct—on the treatment of age in these and other studies using other common exposures (i.e., body mass index and alcohol consumption). Among the 50 empirical articles reviewed, the majority ($n = 29$; 58%) used the incorrect method of analyzing accelerated aging detrended for age as the outcome and did *not* control for age as a covariate. By contrast, only 42% used correct methods, which are either to analyze accelerated aging detrended for age as the outcome *and* control for age as a covariate ($n = 16$; 32%) or to analyze raw DNAm age as the outcome *and* control for age as a covariate ($n = 5$; 10%). In accord with prior demonstrations of bias introduced by use of the incorrect approach, we provide simulation analyses and additional empirical analyses to illustrate how the incorrect method can lead to bias towards the null, and we discuss implications for extant research and recommendations for best practices.

accelerated aging; air pollution; Avon Longitudinal Study of Parents and Children, DNA methylation; epigenetic age; epigenetic clocks; racism; socioeconomic position

Abbreviations: ALSPAC, Avon Longitudinal Study of Parents and Children; BMI, body mass index; DNAm, DNA methylation; SEP, socioeconomic position.

Accounting for age in epidemiologic studies is fundamental to the field (1–5), including but not limited to research focused on measures of biological aging. In the past decade, technological advances have enabled development of epigenetic clocks, which have become a widely used tool in aging research (6–8) (see Web Table 1, available at https://doi.org/10.1093/aje/kwad025). An epigenetic clock uses DNA methylation (DNAm) patterns to estimate the biological age of an individual or biological specimen, and these estimates are termed "DNAm age" (also referred to as "epigenetic age") (6–8). Above-average DNAm age relative to chronological age constitutes "accelerated aging," which in turn is associated, as hypothesized, with increased risk of numerous adverse health outcomes, including cardiometabolic disease, cancer, and younger age at death (6–8).

In this paper, we report on problems we encountered, which to our knowledge have not previously been documented, when reviewing recent epidemiologic research focused on such epigenetic clocks as the outcome (Web Tables 2–4). At issue is use of 3 methods to account for chronological age—one of which, though widely used, is incorrect, and 2 of which are correct (Table 1). These methods are to: 1) incorrectly analyze accelerated aging detrended for age (typically measured as the residual of regressing DNAm age on chronological age) as the outcome and *not* control for age as a covariate; 2) correctly analyze accelerated aging detrended for age as the outcome *and also* control for age as a covariate; and 3) correctly analyze raw DNAm age as the outcome *and also* control for age as a covariate (Table 1). We demonstrate that use of the incorrect

Table 1. Different Approaches to the Treatment of Age in Epidemiologic Analyses of Epigenetic Accelerated Aging as an Outcome, 2014–2022[a]

Article Type and Approach to Accounting for Age in the Analyses	No. of Analyses[b]				No. of Articles[b]					
	Social Exposure		Health Characteristic		Social Exposure			Health Characteristic		
	No.	%	No.	%	Racism	Economic Adversity	Air Pollution	BMI[c] and Alcohol Consumption	BMI Only	Alcohol Consumption Only
Review articles	(*n* = 7 Analyses)		(*n* = 2 Analyses)							
No discussion regarding heterogeneity of approaches to controlling or note for age as a covariate if the outcome employed is age-detrended epigenetic accelerated age	7	100	2	100	N/A	N/A	N/A	N/A	N/A	N/A
Empirical articles	(*n* = 33 Analyses)		(*n* = 17 Analyses)							
Method 1: Use detrended age acceleration and *do not* adjust for age as a covariate.	18	55	11	65	3	13	2	1	5	5
Method 2: Use detrended age acceleration and *do* adjust for age as a covariate.	10	30	6	35	1	8	1	2	2	2
Method 3: Use raw DNAm age and *do* adjust for age as a covariate.	5	15	0	0	1	1	3	0	0	0
All empirical articles	33	100	17	100	5	22	6	3	7	7

Abbreviations: BMI, body mass index; DNAm, DNA methylation; N/A, not applicable.

[a] The articles reviewed comprised review articles and empirical analyses focused on social exposures (racism, economic adversity, air pollution) and health characteristic (BMI and alcohol consumption), as identified by literature searches in PubMed (National Library of Medicine, Bethesda, Maryland) and Web of Science (Clarivate Analytics PLC, Philadelphia, Pennsylvania). The literature search for the social epidemiologic exposures included articles indexed through November 16, 2021; the search for BMI and alcohol exposures included articles indexed through January 6, 2022.

[b] See Web Table 3 for a list of social exposure articles included in each table cell, and see Web Table 4 for the corresponding list of articles on BMI and alcohol exposure. The analyses of 3 articles were separately counted twice, because 1 article included separate analyses independently using racism and economic adversity as exposures (26) and 2 articles included separate analyses independently using social exposures and health characteristics (27, 28).

[c] Weight (kg)/height(m)^2.

approach can bias findings towards the null and also renders it difficult to interpret study findings, given the possibilities of residual confounding by age, let alone compare results across studies. Such problems in the incorrect treatment of age and other potential confounders have long been recognized (9–13), including most recently in genetic epidemiology research (14, 15).

Motivating our analysis was a comprehensive literature review update we conducted in November 2021 (Web Table 2) regarding epidemiologic investigations of epigenetic accelerated aging (as an outcome) in relation to exposures involving racial discrimination, economic adversity, and air pollution, relevant to analyses we have in progress (16). We extended the review to include 2 other common exposures assessed in the epigenetic literature: body mass index (BMI; weight (kg)/height $(m)^2$) and alcohol consumption (Web Table 2). Informing our concerns, associations between age and these social exposures and socially patterned health characteristics and practices can potentially be shaped by not only chronological age but also birth cohort and period effects, hence secular trends, reflecting the larger societal context (e.g., legal age for enrollment in school; legal age to purchase alcohol; being born before or after the imposition or abolition of "Jim Crow" laws legalizing racial discrimination) (4, 5). Depending on the causal process at issue, age can thus potentially confound and/or modify the specified exposure-outcome associations (17, 18). To illustrate how, in relation to epigenetic aging, the incorrect method can bias results towards the null, we provide both simulation analyses and new empirical analyses, and we discuss implications for the extant literature and recommendations for best practices.

METHODS

Literature review

The initial objective of our literature review was to identify review articles and empirical studies, published through November 2022, that explicitly focused on relationships between DNAm age acceleration and exposure to 3 types of adversity: racism, social class injustice, and higher levels of air pollution (see Web Table 2 for search strategy and search terms). For each article, we assessed how the investigators accounted for age 1) in their conceptual discussion of the literature and 2) for the empirical studies, in their statistical analyses (Web Table 3), noting that some of the researchers also carried out analyses in which DNAm age acceleration was used as a predictor of health outcomes. We included review articles because of their importance in critically evaluating extant literature for the weaknesses and strengths of both evidence and methods used to obtain this evidence (see Web Table 2). Additionally, as a check on whether the findings for the reviewed social epidemiologic articles were specific to this field, in January 2022 we conducted an analogous search of articles focused on DNAm age acceleration in relation to the exposures of BMI and alcohol consumption (Web Table 4).

Statistical methods: Frisch-Waugh-Lovell theorem and the "partialing out" interpretation in multiple regression

A fixture of many a basic statistics course on ordinary least squares regression, especially in econometrics, is a discussion of the "partialing out" interpretation in multiple regression (12, 13). Consider, for example, a model with $k = 2$ independent variables,

$$\mathbb{E}(Y_i) = \beta_0 + \beta_1 X_i + \beta_2 Z_i,$$

for $i = 1, \ldots n$ individuals. One way to estimate β_1 is

$$\beta_1 = \frac{\sum_i \tilde{X}_i y_i}{\sum_i \tilde{X}_i^2},$$

where the $\tilde{X}_i$ are the ordinary least squares residuals from a simple regression of X on Z. As Woolridge notes (12, p. 69), the residuals $\tilde{X}_i$ have a zero sample average, and so $\hat{\beta}_1$ is the usual slope estimate from simple regression. We can think of the residuals $\tilde{X}_i$ as the part of X_i that is uncorrelated with Z_i. That is, $\tilde{X}_i$ is what remains of X_i after the effect of Z_i has been *partialed out* or *netted out*. More generally, as Angrist and Pischke explain in *Mostly Harmless Econometrics* (13), the usual ordinary least squares estimator of $\boldsymbol{\beta}$,

$$\boldsymbol{\beta} = \mathbb{E}\left[X_i^T X\right]^{-1} \mathbb{E}\left[X_i^T Y_i\right],$$

is a $k \times 1$ vector with kth element

$$\beta_k = \frac{\mathrm{Cov}\left(Y_i, \tilde{X}_{ki}\right)}{\mathrm{Var}\left(\tilde{X}_{ki}\right)},$$

where $\tilde{X}_{ki}$ is the residual from a regression of X_{ki} on all of the other covariates. The formula "shows us that each coefficient in a multivariate regression is the bivariate slope coefficient for the corresponding regressor after partialing out all the other covariates" (13, p. 35). This result is known as the Frisch-Waugh-Lovell theorem (9, 10), which recognizes, respectively, the contributions of Frisch and Waugh in identifying this problem in 1933 and the extension of this work by Lovell in 1963. Angrist and Pischke note that this "regression anatomy formula," as they term it (13, p. 36), can also be written

$$\beta_k = \frac{\mathrm{Cov}\left(\tilde{Y}_i, \tilde{X}_{ki}\right)}{\mathrm{Var}\left(\tilde{X}_{ki}\right)},$$

where $\tilde{Y}_i$ is the residual from a regression of Y_i on every covariate except X_{ki}. They state, "This works because the fitted values removed from $\tilde{Y}_{ki}$ are uncorrelated with $\tilde{X}_{ki}$" (13, p. 36).

To our knowledge, no epigenetic studies have explicitly addressed the implications of the Frisch-Waugh-Lovell theorem for analyzing DNAm age acceleration. However, a 2011 article by Demissie and Cupples (14), focused on

genetic association studies, does derive the bias induced by failing to "partial out" analyses correctly and considers the implications for analogous kinds of investigations. We refer readers to this paper for the technical derivation of this bias; see also Web Appendix 1.

Here, we flag that the "partialing out" result does *not* apply to the original motivating example of age-detrended epigenetic clocks analyzed in relation to covariates where no additional modeling of age is included in the regression model. In this setting, one would have the age-detrended epigenetic clock as $\tilde{Y}_i$ but would be regressing it on non–age-detrended X variables. As Angrist and Pischke (13) noted,

$$\frac{\mathrm{Cov}(\tilde{Y}_i, X_{ki})}{\mathrm{Var}\,(X_{ki})} = \left[\frac{\mathrm{Cov}(\tilde{Y}_i, \tilde{X}_{ki})}{\mathrm{Var}(\tilde{X}_{ki})}\right]\left[\frac{\mathrm{Var}(\tilde{X}_{ki})}{\mathrm{Var}\,(X_{ki})}\right] \neq \beta_k,$$

unless X_{ki} is uncorrelated with the other covariates.

Thus, when regressing age-detrended epigenetic clocks on other exposures and covariates that have not been similarly age-detrended, we expect β estimates to be biased unless these variables are uncorrelated with age (14).

We additionally underscore that in the omitted-variable setting (in which there is no adjustment for age or no detrending of DNAm for chronological age; see Web Figures 1 and 2), the bias depends on 1) the relationship between age and the outcome, 2) the correlation between age and the exposure, and 3) the relative variation in age and the exposure, whereas in the detrended setting, the magnitude of the bias depends on the magnitude of β_1 and the correlation between age and the exposure. Also, while omitted-variable bias can bias results towards or away from the null, in the detrended setting, the direction of the bias is always towards the null. Intuitively, failing to adjust for age at all is a bigger problem than age-detrending only the outcome, since the resulting bias can go in any direction (Web Figures 1 and 2). Moreover, in applied settings, the extent of the bias when using detrended outcome data may be small relative to the uncertainty in the estimate of $\hat{\beta}_1$. Also important to consider is the functional form of the relationship between chronological age and DNAm age (i.e., linear vs. nonlinear); while most of the studies that included age as a covariate in Web Tables 3 and 4 used only age, 2 studies included terms for both age and age^2.

Simulations

We employed simulations to illustrate the implications of using the 3 analytical approaches we identified in the literature:

- *Method 1*: Detrended DNAm age is modeled as the outcome with no additional adjustment for chronological age.
- *Method 2*: Detrended DNAm age is modeled as the outcome with chronological age included in the model.
- *Method 3*: Raw DNAm is modeled as the outcome with chronological age included in the model.

To do so, we developed 5 simulation scenarios as summarized in Table 2.

- In simulation 1, DNAm age (M) is an outcome that depends on an exposure (X) and chronological age (Z), and the true relationship between Z and M is linear.
- In simulation 2, DNAm age (M) is an outcome that depends on an exposure (X) and chronological age (Z), and the true relationship between Z and M includes a squared term for chronological age.

Additionally, although the focus of our paper is on studies where DNAm age is the outcome, we also wanted to illustrate the consequences of detrending DNAm age and choices around the inclusion of chronological age in models when DNAm is a predictor of an outcome and included on the right-hand side of the model or when it is analyzed as a mediator (also approaches employed in a handful of the papers listed in Web Tables 3 and 4). Thus:

- In simulation 3, DNAm age is an exposure analyzed in relation to an outcome Y that depends only on DNAm and chronological age.
- In simulation 4, the outcome depends on DNAm age, chronological age, and an additional exposure X whose relationship to the outcome is also confounded by chronological age, and the simulation considers the case where all relationships are linear.
- In simulation 5, the outcome depends on DNAm age, chronological age, and an additional exposure X whose relationship to the outcome is also confounded by chronological age, and the model introduces a squared term for chronological age.

We considered simulation scenarios 2 and 5 to explore the potential for bias when the age-detrending model for obtaining detrended DNAm age is misspecified.

To tie the simulations to previously published data, we chose parameters for the data-generating models from relationships between the epigenetic clock GrimAge, chronological age, and a socioeconomic index in the Multi-Ethnic Study of Atherosclerosis, as reported by Schmitz et al. (19). Because there were virtually no articles in our literature review that presented data sufficient to inform simulation of outcomes, we simulated data for a hypothetical outcome using the values of parameters for the data-generating processes summarized in Table 2. We also conducted new empirical analyses, showing that results obtained with each method applied to identical data obtained from the Avon Longitudinal Study of Parents and Children (ALSPAC) (20).

For each simulation (except simulation 3) we varied the strength of the correlation ρ_{xz} between the exposure (X) and chronological age (Z) from 0.1 to 0.5. Since under simulation 3 we do not expect bias when using detrended DNAm age, we simulated under a fixed correlation of 0.2. We conducted 5,000 iterations for each simulation, and we summarized the mean value and 95% quantiles of the sampling distribution of the estimators of the target estimand for each scenario using the 3 analytical methods.

Table 2. Simulation Scenarios, Data-Generating Models, Target Estimands, and Parameters Used in a Study of Approaches to the Treatment of Age in Epidemiologic Analyses of Epigenetic Accelerated Aging as an Outcome, 2014–2022[a]

Simulation	Description	Parameters
1	DNAm age (M) as an outcome that depends on an exposure (X) and chronological age (Z) $\begin{Bmatrix} X_i \\ Z_i \end{Bmatrix} \sim MVN \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix}$ $M_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_M$ Target estimand: β_1	$\mu_x = 2.82$ $\sigma_x = 0.91$ $\mu_z = 70$ $\sigma_z = 11$ $\rho_{xz} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ $\beta_0 = 79.4$ $\beta_1 = 0.5$ $\beta_2 = 0.744$ $\epsilon_M = 6$
2	DNAm age (M) as an outcome that depends on an exposure (X) and chronological age (Z), where the true relationship between epigenetic age and chronological age is nonlinear $\begin{Bmatrix} X_i \\ Z_i \end{Bmatrix} \sim MVN \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix}$ $M_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 Z_i^2 + \epsilon_M$ Target estimand: β_1	$\mu_x = 2.82$ $\sigma_x = 0.91$ $\mu_z = 70$ $\sigma_z = 11$ $\rho_{xz} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ $\beta_0 = 79.4$ $\beta_1 = 0.5$ $\beta_2 = 0.744$ $\beta_3 = -0.004$ $\epsilon_M = 6$
3	DNAm age (M) as an exposure analyzed in relation to an outcome (Y) that depends only on M and chronological age (Z) $\begin{Bmatrix} X_i \\ Z_i \end{Bmatrix} \sim MVN \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma & z\sigma_z^2 \end{bmatrix}$ $M_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_M$ $Y_i = \theta_0 + \theta_1 M_i + \theta_2 Z_i + \epsilon_Y$ Target estimand: θ_1	$\mu_x = 2.82$ $\sigma_x = 0.91$ $\mu_z = 70$ $\sigma_z = 11$ $\rho_{xz} = 0.2$ $\beta_0 = 79.4$ $\beta_1 = 0.5$ $\beta_2 = 0.744$ $\epsilon_M = 6$ $\theta_0 = 20$ $\theta_1 = 2.0$ $\theta_2 = 0.5$ $\epsilon_Y = 5$
4	DNAm age (M) as an exposure analyzed in relation to an outcome (Y) that depends on M, chronological age (Z), and an additional covariate (X) that is also associated with chronological age $\begin{Bmatrix} X_i \\ Z_i \end{Bmatrix} \sim MVN \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix}$ $M_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_M$ $Y_i = \theta_0 + \theta_1 M_i + \theta_2 Z_i + \theta_3 X_i + \epsilon_Y$ Target estimand: θ_1	$\mu_x = 2.82$ $\sigma_x = 0.91$ $\mu_z = 70$ $\sigma_z = 11$ $\rho_{xz} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ $\beta_0 = 79.4$ $\beta_1 = 0.5$ $\beta_2 = 0.744$ $\epsilon_M = 6$ $\theta_0 = 20$ $\theta_1 = 2.0$ $\theta_2 = 0.5$ $\theta_3 = 1.0$ $\epsilon_Y = 5$

Table continues

RESULTS

Literature review

Common to all the literature we reviewed, the standard approach to estimating raw DNAm age was to derive it from each epigenetic clock's algorithm, which converts the DNA methylation data (typically obtained from blood samples) into estimates of DNAm age (Web Tables 1, 3, and 4). Following standard practice (6–8), the most commonly employed measure of accelerated epigenetic aging used in the empirical studies was the residual obtained from regressing the observed raw DNAm age on chronological age; a handful used DNAm age minus chronological age (Web Tables 3 and 4).

Table 2. Continued

Simulation	Description	Parameters
5	DNAm age (M) as an exposure analyzed in relation to an outcome (Y) that depends on M, chronological age (Z), and an additional covariate (X), where the true relationship between epigenetic age and chronological age is nonlinear and the true relationship between age and the outcome is nonlinear $\begin{Bmatrix} X_i \\ Z_i \end{Bmatrix} \sim MVN \begin{bmatrix} \sigma_x^2 & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \sigma_z^2 \end{bmatrix}$ $M_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 Z_i^2 + \epsilon_M$ $Y_i = \theta_0 + \theta_1 M_i + \theta_2 Z_i + \theta_3 Z_i^2 + \theta_4 X_i + \epsilon_Y$ Target estimand: θ_1	$\mu_x = 2.82$ $\sigma_x = 0.91$ $\mu_z = 70$ $\sigma_z = 11$ $\rho_{xz} = \{0.1, 0.2, 0.3, 0.4, 0.5\}$ $\beta_0 = 79.4$ $\beta_1 = 0.5$ $\beta_2 = 0.744$ $\beta_3 = -0.004$ $\epsilon_M = 6$ $\theta_0 = 20$ $\theta_1 = 2.0$ $\theta_2 = 0.5$ $\theta_3 = -0.002$ $\theta_4 = 1.0$ $\epsilon_Y = 5$

Abbreviation: DNAm, DNA methylation.

[a] For each simulation scenario, we analyzed the data using the following 3 methods: 1) DNAm age is detrended for chronological age by regressing epigenetic age on chronological age and saving the residuals. Detrended DNAm age is included in the model and chronological age is omitted. 2) DNAm age is detrended for chronological age by regressing DNAm age on chronological age and saving the residuals. Detrended DNAm age is included in the model along with chronological age. 3) Raw DNAm age is included in the model along with chronological age.

Among the 40 relevant articles we identified (7 review, 33 empirical) that focused on epigenetic accelerated aging as an outcome in relation to exposures involving racial discrimination, economic adversity, and air pollution, none of the review articles explicitly discussed the different approaches taken in the empirical literature regarding how age was treated in their statistical models (Table 1; Web Table 3). Among the 33 identified empirical analyses, the most frequent approach (used in 18 analyses; 55%) was to 1) incorrectly use the age-detrended DNAm age as the outcome and 2) *not* control for age in models examining the associations of this outcome with the specified exposures (Table 1; Web Table 3). However, 10 of the analyses (30%) that used this same outcome *did* correctly control for age, and 5 analyses (15%) used raw DNAm age as the outcome and correctly controlled for age as a covariate (Table 1; Web Table 3).

A similar mix of incorrect and correct approaches was evident in the 2 review articles and 17 empirical studies for which BMI, alcohol, or both served as the exposure measure(s) (Table 1; Web Table 4). For example, among the empirical analyses, investigators in 11 studies (65%) incorrectly reported on the associations of an age-detrended DNAm age with these exposures unadjusted for age, whereas 6 (35%) used this same outcome but correctly adjusted for age in their statistical analyses; none analyzed raw DNAm age as the outcome. Thus, among the total of 50 empirical analyses reviewed that focused on social exposures or on BMI and alcohol consumption as exposures, investigators in 29 studies (58%) incorrectly analyzed age-detrended DNAm age as the outcome and did *not* control for age as a covariate; 16 (32%) used this outcome and *did* correctly control for age; and 5 (10%) analyzed raw DNAm age as the outcome *and* correctly controlled for age as a covariate.

Derivation of bias

As described above, we derived the bias under method 1 in the setting of simulation scenario 1 and show that it is a function of the true relationship between the exposure X and DNAm age M and the correlation ρ_{xz} between the exposure and chronological age (Z):

$$\text{Bias}(\tilde{\beta}_1) = \mathbb{E}(\tilde{\beta}_1) - \beta_1 = \beta_1(1 - \rho_{xz}^2 - 1) = \beta_1\rho_{xz}^2,$$

where $\tilde{\beta}_1$ is the estimated regression coefficient for the exposure X when detrended DNAm age is regressed on the exposure and chronological age Z is not included in the model, β_1 is the true relationship between the exposure and DNAm age, and ρ_{xz} is the correlation between the exposure X and chronological age Z. Conversely, we do not expect bias when raw DNAm age is the outcome and age is included as a covariate in the model (method 3) or when detrended DNAm age is the outcome and age is included as a covariate in the model (method 2).

Simulations

Figure 1A shows the distributions of estimates for analytical methods 1–3 under simulation scenario 1. Method 3 (raw DNAm age is modeled as the outcome with chronological age included in the model) yields unbiased estimates of the

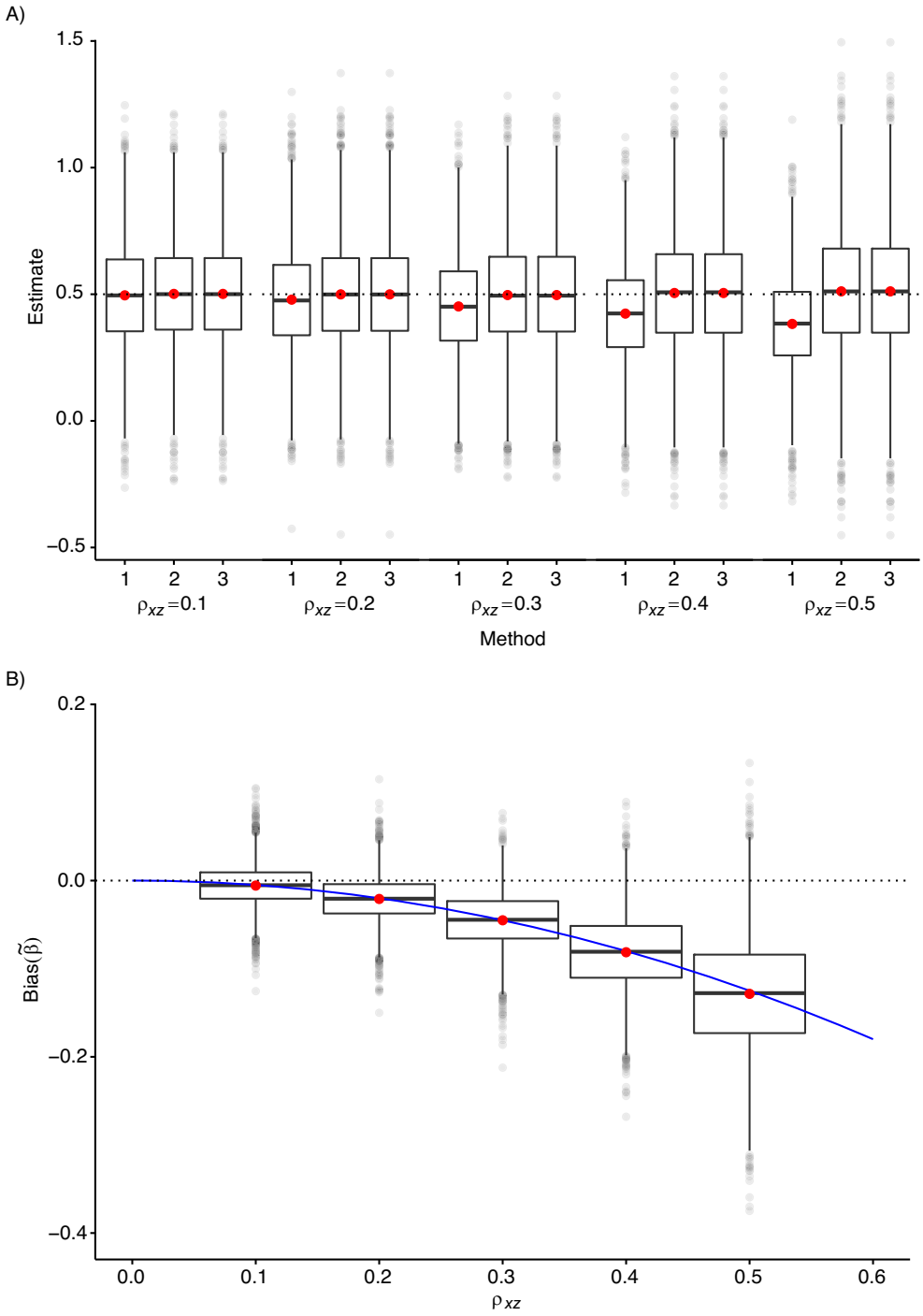

Figure 1. Simulation results for bias in estimates of the association between epigenetic accelerated aging and a hypothetical socioeconomic position metric. A) Distribution of estimates using analytical methods 1–3 for simulation scenario 1 for varying values of ρ_{xz}; B) distribution of bias over 5,000 simulations under analytical method 1 for simulation scenario 1 for varying values of ρ_{xz}. The solid blue line in panel B shows the predicted bias as derived in the Methods section. Per convention for box-and-whisker plots, the top and bottom of the box represent the 75th and 25th percentiles, the horizontal line inside the box represents the mean value, the whiskers are the 10th and 90th percentiles, and circles represent outliers.

association between a hypothetical index for socioeconomic position (SEP) and accelerated aging (since the analytical model corresponds directly to the data-generating model). Method 2 (detrended DNAm age is modeled as the outcome with chronological age included in the model) yields estimates identical to those of method 1. However, method 1

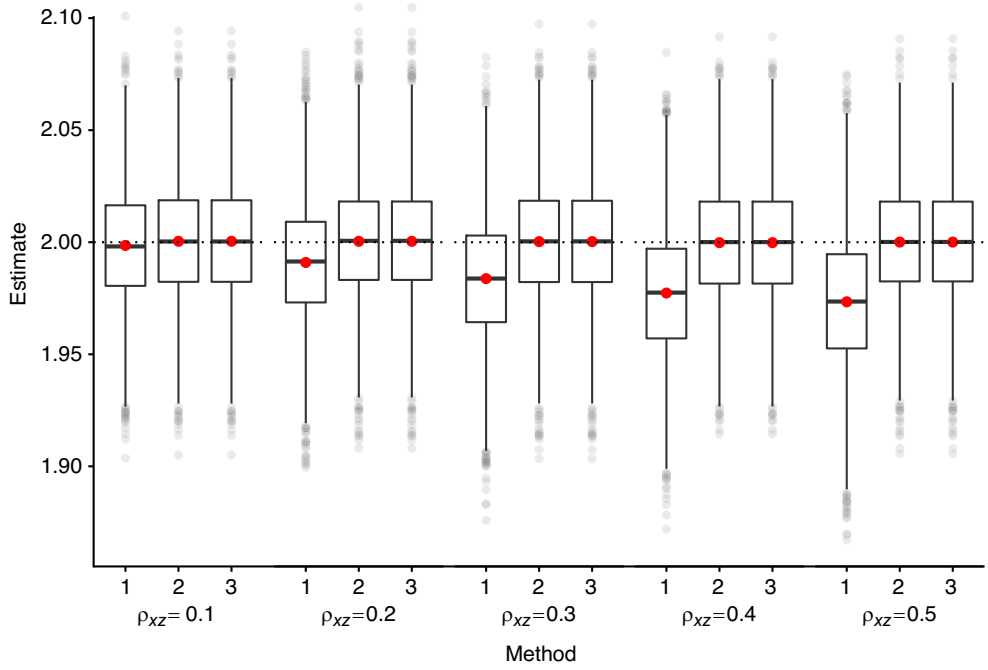

Figure 2. Simulation results for bias when the true relationship between chronological age and DNA methylation age is quadratic and a linear specification for chronological age is assumed. The graph shows the distribution of estimates over 5,000 simulations under analytical methods 1–3 for simulation scenario 2 for varying values of ρ_{xz}. Per convention for box-and-whisker plots, the top and bottom of the box represent the 75th and 25th percentiles, the horizontal line inside the box represents the mean value, the whiskers are the 10th and 90th percentiles, and circles represent outliers.

(detrended DNAm age is modeled as the outcome with no additional adjustment for chronological age) yields biased estimates of the association with SEP index. As illustrated in Figure 1B, the bias incurred by method 1 increases as the correlation between SEP index and chronological age increases, and it matches the bias predicted by the derivation above.

In Figure 2 we consider simulation scenario 2, where the true relationship between chronological age and DNAm age is quadratic but where, under analytical method 1, a linear specification for chronological age is assumed when regressing raw DNAm age on chronological age and saving the residuals. As the correlation between SEP index and chronological age increases, the extent of bias increases.

Figure 3 illustrates the results obtained under the restrictive setting of simulation scenario 3, where DNAm age and chronological age are the only predictors of a continuous outcome *Y*. Under this scenario, all 3 analytical methods yield identical unbiased estimates of the association between DNAm age and the outcome. We note that under method 1, detrended DNAm age meets the conditions under the Frisch-Waugh-Lovell theorem (9, 10) for the estimated regression coefficient to be unbiased, since there are no other covariates in the model.

In contrast, Figure 4 illustrates the increasing bias incurred by analytical method 1 under simulation scenario 4, where DNAm age and SEP index are both independent predictors of the outcome *Y* and both covariates are associated with chronological age. Because analytical method 1 fails to control for chronological age, which is a confounder of the SEP index association with the outcome, the parameter estimate for the SEP index is misestimated, and thus in turn the estimate for the effect of DNAm age is confounded, even though under analytical method 1 DNAm age has been detrended for chronological age. In Figure 4, however, we see that analytical methods 2 and 3 continue to yield unbiased estimates of the association between DNAm age and the outcome.

Under simulation scenario 5, when the true relationship between chronological age and the outcome is quadratic (Figure 5), analytical method 1 yields substantially biased estimates of the association between DNAm age and the outcome. Notably, though the detrending model assumed (incorrectly) a linear relationship between chronological age and DNAm age, analytical methods 2 and 3 still yield unbiased estimates of the association between DNAm age and the outcome when the correct functional relationship with chronological age is included in the model.

Empirical demonstration of bias

In Web Figure 3, we depict the observed and expected estimates of association for 4 epigenetic clocks (Horvath, Hannum, GrimAge, and PhenoAge) for 14 outcomes with demonstrated associations with epigenetic age, using data

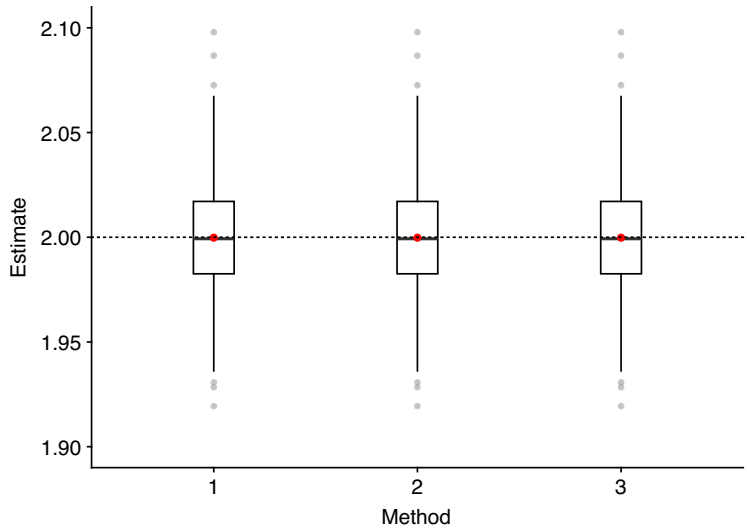

Figure 3. Simulation results for bias when DNA methylation age and chronological age are the only predictors of a continuous outcome *Y*. The graph shows the distribution of estimates over 5,000 simulations under analytical methods 1–3 for simulation scenario 3 for a fixed value of ρ_{xz} = 0.2. As anticipated, there is no bias when using detrended DNAm age (analytical method 1) when there are no other covariates. Per convention for box-and-whisker plots, the top and bottom of the box represent the 75th and 25th percentiles, the horizontal line inside the box represents the mean value, the whiskers are the 10th and 90th percentiles, and circles represent outliers.

from 4,355 individuals aged 17–24 years in ALSPAC (20, 21); see Web Appendix 2 for detailed descriptions of the ALSPAC population and model results (Web Tables 5 and 6).

As expected, use of the incorrect approach (i.e., analytical method 1) resulted, for most pairs, in attenuation of associations as compared with those obtained using analytical

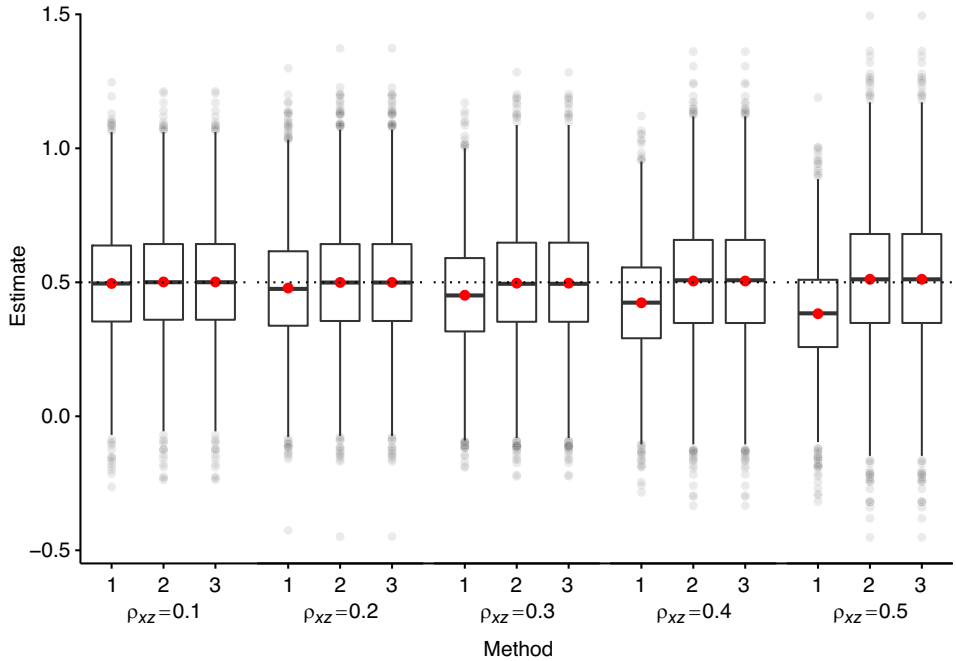

Figure 4. Simulation results for bias when DNA methylation age and socioeconomic position index are both independent predictors of the outcome *Y* and both covariates are associated with chronological age. The graph shows the distribution of estimates over 5,000 simulations under analytical methods 1–3 for simulation scenario 4 for varying values of ρ_{xz}. Per convention for box-and-whisker plots, the top and bottom of the box represent the 75th and 25th percentiles, the horizontal line inside the box represents the mean value, the whiskers are the 10th and 90th percentiles, and circles represent outliers.

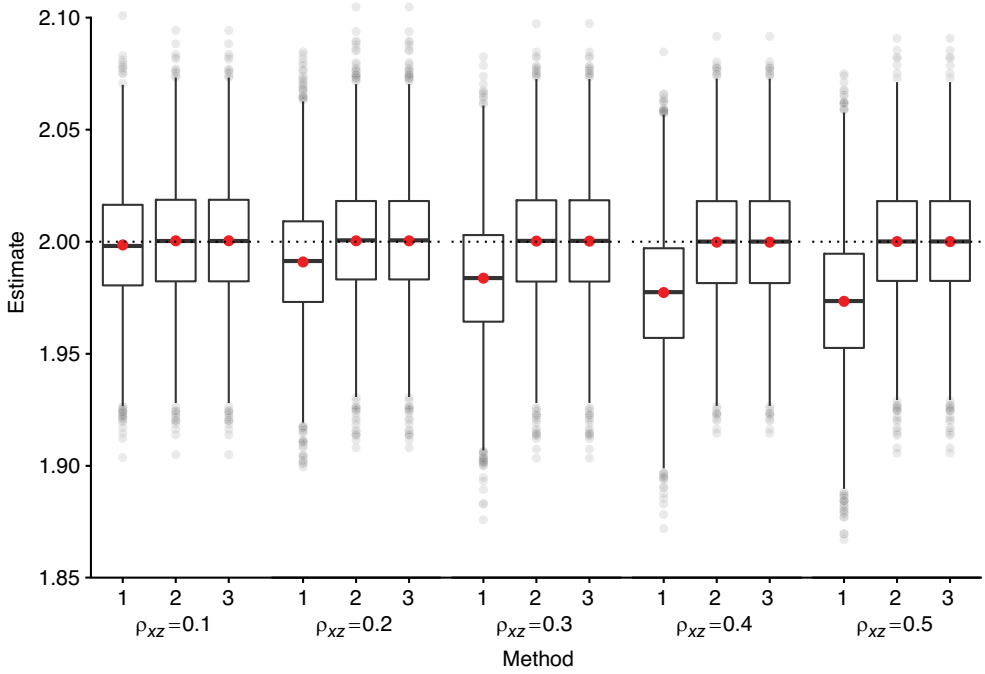

Figure 5. Simulation results for bias when the true relationship between chronological age and the continuous outcome *Y* is quadratic. The graph shows the distribution of estimates over 5,000 simulations under analytical methods 1–3 for simulation scenario 5 for varying values of ρ_{xz}. Per convention for box-and-whisker plots, the top and bottom of the box represent the 75th and 25th percentiles, the horizontal line inside the box represents the mean value, the whiskers are the 10th and 90th percentiles, and circles represent outliers.

methods 2 and 3; and these attenuations were greater for variables with higher correlations with age. In those few pairs whose associations were larger using analytical method 1, this was probably due to sampling variability and was also observed in our simulations.

DISCUSSION

The widespread use of both correct and incorrect approaches for accounting for age in epidemiologic studies that investigate DNAm age acceleration is concerning. Considered together, among the 50 empirical analyses we reviewed in relation to diverse social exposures involving racial discrimination, economic adversity, and air pollution, as well as BMI and alcohol consumption, fully 58% of the 50 empirical analyses we reviewed used incorrect methods (i.e., did not control for age in models whose outcome was accelerated aging detrended for age), and only 21 (42%) of the analyses employed the appropriate approach of controlling for age as a covariate when their outcome comprised either age-detrended DNAm age acceleration or raw DNAm age. Yet, as nearly a century of research has shown (9–15), if an outcome is detrended for age, so too must be the exposure(s) and the covariate(s).

Three implications of our findings stand out. First, results of extant epigenetic clock studies using the incorrect method are presumably biased towards the null. Second, the results of these studies cannot be directly compared with those from studies that used the correct methods. Third, given publication bias against null findings (22), it is plausible that there may be underreporting of meaningful significant associations between accelerated epigenetic aging and social or other exposures that were biased towards the null by use of incorrect methods.

For best practices, we accordingly recommend use of one of the 2 correct methods:

- Method 2, which uses the age-detrended measures of DNAm age in analyses that also control for age (in relation to the exposure and covariates).
- Method 3, which employs raw DNAm age as the outcome and controls for age as a covariate, and which likely is preferable, due to ease of modeling and greater data transparency.

As Wang et al. (23) note, when raw DNAm age is modeled as the outcome and chronological age is included in the model as a covariate, effect estimates for covariates can still be interpreted as effects on accelerated aging. Additionally, we urgently recommend that explicit adjustment for chronological age be employed in reanalyses of existing studies that have not accounted for age confounding in relation to exposures and covariates, so as to 1) generate more accurate estimates of association and 2) enable more rigorous and robust comparisons of results across studies.

Moreover, as simulation scenarios 2 and 5 showed, care should be taken to explore the correct functional form of the relationship between chronological age and DNAm age in the data. It is our opinion that method 3 makes this

easier to do, since the raw DNAm age and chronological age variables are available to the data analyst and can thus be examined in a careful model-building process. However, as long as due diligence is applied in modeling chronological age even when using detrended DNAm age (even when a linear relationship has been assumed in the detrending process), method 2 will also yield valid results.

One concern for comparison across different studies is technical variation in the measurement of DNAm both within and between studies, including the use of different arrays (or other methods for quantifying DNAm), limiting the ability to compare findings between studies (24, 25). In particular, not all DNAm sites are represented on each chip. As such, some sites in the epigenetic clock algorithm may be missing, especially from the Infinium MethylationEPIC BeadChip (Illumina, Inc., San Diego, California), as most epigenetic clocks were developed using the Illumina HumanMethylation450 BeadChip (450K). These missing sites are likely to mean that raw DNAm age summary measures will not be directly comparable across studies if they do not use the same chip. While this can be a concern when reporting overall mean values for DNAm age across different studies, we anticipate that this will not be problematic when reporting effect estimates from models, since control for chronological age in the model effectively does the same thing as detrending for age using the residual method. When different chips are used in the same study, chip effects should clearly be included in the model as well. Further work should be done to consider the implications for analyses of DNAm age in relation to dichotomous, count, or time-to-event outcomes.

In conclusion, the widespread use of incorrect treatment of age in epidemiologic studies using an epigenetic clock as an outcome, despite well-known methodological critiques of consequent biases (9–15), warrants serious reflection. In the case of age, every individual by definition not only has a specific chronological age tied to time period at birth but is also characterized by being a member of a birth cohort—and both period and birth cohort are highly likely to shape the kinds of exposures which an individual encounters over the life course (1–5). Thus, age, by virtue of simultaneously being both a social and a biological variable, necessarily has a high likelihood of being a relevant exposure, confounder, and potentially modifier of any exposure-outcome association, rendering it crucial to employ correct methods to account for age in epidemiologic investigations.

ACKNOWLEDGMENTS

Author affiliations: Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States (Nancy Krieger, Jarvis T. Chen, Christian Testa, Pamela D. Waterman); Dean's Office and Department of Epidemiology and Biostatistics, Dornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania, United States (Ana Diez Roux); MRC Integrative Epidemiology Unit, Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom (Kate Tilling, Sarah Watkins, Matthew Suderman, George Davey Smith, Caroline Relton); School of Mathematical and Statistical Sciences, National University of Ireland, Galway, Galway, Ireland (Andrew J. Simpkin); and Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States (Immaculata De Vivo).

This work was funded by the National Institute of Minority Health and Health Disparities, US National Institutes of Health (grant R01MD014304 to N.K. and C.R., as multiple Principal Investigators), with the portions involving ASLPAC being funded by the Medical Research Council and the Wellcome Trust (grant 217065/Z/19/Z) and the University of Bristol (Bristol, United Kingdom) and the portions pertaining to ARIES (https://clinicaltrials.gov/ct2/show/NCT02581891) being funded by the Biotechnology and Biological Sciences Research Council (grants BBI025751/1 and BB/I025263/1). Supplementary funding for generation of the DNA methylation data included in ARIES was obtained from the Medical Research Council, the Economic and Social Research Council, the National Institutes of Health, and other sources. ARIES is maintained under the auspices of the MRC Integrative Epidemiology Unit at the University of Bristol (grants MC_UU_12013/2, MC_UU_12013/8, and MRC_UU_12013/9).

The ALSPAC website contains details on all of the data that are available through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data/). The question of data availability is not applicable for the simulation analyses.

For the ALSPAC component of our paper, we are extremely grateful to all of the families who took part in the 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.

The views expressed in this article are those of the authors and do not reflect the views of any of the funders.

Conflict of interest: none declared.

REFERENCES

1. Morris JN. *Uses of Epidemiology*. Edinburgh, United Kingdom: E & S Livingston; 1957.
2. Kuh D, Ben-Sholomo Y. *A Life Course Approach to Chronic Disease Epidemiology*. 2nd ed. New York, NY: Oxford University Press; 2004.
3. Yashi A, Jazwinski S, Fulop T. *Aging and Health: A Systems Biology Perspective*. (Interdisciplinary Topics in Gerontology, vol. 40). Basel, Switzerland: S. Karger AG; 2014.
4. Estes C, DiCarlo NB. *Aging A–Z: Concepts Toward Emancipatory Gerontology*. New Brunswick, NJ: Routledge; 2019.

5. Krieger N. *Ecosocial Theory, Embodied Truths, and the People's Health*. New York, NY: Oxford University Press; 2021.
6. Maroni RE, Suderman M, Chen BH, et al. Tracking the epigenetic clock across the human life course: a meta-analysis of longitudinal cohort data. *J Gerontol A Biol Sci Med Sci.* 2019;74(1):57–61.
7. Ryan CP. "Epigenetic clocks": theory and applications in human biology. *Am J Hum Biol.* 2021;33(3):e23488.
8. Noroozi R, Ghafouri-Fard S, Pisarek A, et al. DNA methylation-based age clocks: from age prediction to age reversion. *Ageing Res Rev.* 2021;68:101314.
9. Frisch R, Waugh FV. Partial time regressions as compared with individual trends. *Econometrica.* 1933;1(4):387–401.
10. Lovell M. Seasonal adjustment of economic time series and multiple regression analysis. *J Am Stat Assoc.* 1963;58(304): 993–1010.
11. Rosenbaum PR, Rubin DB. Difficulties with regression analyses of age-adjusted rates. *Biometrics.* 1984;40(2): 437–443.
12. Wooldridge JM. *Introductory Econometrics: A Modern Approach*. 6th ed. Mason, OH: South-Western Cengage Learning; 2016.
13. Angrist JD, Pischke J-S. *Mostly Harmless Econometrics: An Empiricist's Companion*. 1st ed. Princeton, NJ: Princeton University Press; 2008.
14. Demissie S, Cupples LA. Bias due to two-stage residual-outcome regression analysis in genetic association studies. *Genet Epidemiol.* 2011;35(7):592–596.
15. Che R, Motsinger-Reif AA, Brown CC. Loss of power in two-stage residual-outcome regression analysis in genetic association studies. *Genet Epidemiol.* 2012;36(8):890–894.
16. Krieger N, Relton CL. DNA methylation & adversity: pathways from exposures to health inequities. (NIH project no. 1R01MD014304-01). https://reporter.nih.gov/search/T1qnoiJCukWRK0Mcvk-CJA/project-details/9811618. Published 2019. Accessed March 28, 2022.
17. VanderWeele TJ. *Explanation in Causal Inference: Methods for Mediation and Interaction*. 1st ed. New York, NY: Oxford University Press; 2015.
18. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol.* 2019;34(3):211–219.
19. Schmitz LL, Zhao W, Ratliff SM, et al. The socioeconomic gradient in epigenetic ageing clocks: evidence from the Multi-Ethnic Study of Atherosclerosis and the Health and Retirement Study. *Epigenetics.* 2022;17(6): 589–611.
20. Boyd A, Golding J, Macleod 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(1):111–127.
21. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol.* 2013;42(1): 97–110.
22. Bespalov A, Steckler T, Skolnick P. Be positive about negatives—recommendations for the publication of negative (or null) results. *Eur Neuropsychopharmacol.* 2019;29(12): 1312–1320.
23. Wang C, Koutrakis P, Gao X, et al. Associations of annual ambient $PM_{2.5}$ components with DNAm PhenoAge acceleration in elderly men: the Normative Aging Study. *Environ Pollut.* 2020;258:113690.
24. Zindler T, Frieling H, Neyazi A, et al. Simulating ComBat: how batch correction can lead to the systematic introduction of false positive results in DNA methylation microarray studies. *BMC Bioinformatics.* 2020;21(1): 271.
25. Sugden K, Hannon EJ, Arseneault L, et al. Patterns of reliability: assessing the reproducibility and integrity of DNA methylation measurement. *Patterns (N Y).* 2020;1(2): 100014.
26. Simons RL, Lei MK, Klopack E, et al. The effects of social adversity, discrimination, and health risk behaviors on the accelerated aging of African Americans: further support for the weathering hypothesis. *Soc Sci Med.* 2021;282: 113169.
27. Quach A, Levine ME, Tanaka T, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging (Albany NY).* 2017;9(2):419–446.
28. Zhao W, Ammous F, Ratliff S, et al. Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. *Int J Environ Res Public Health.* 2019; 16(17):3141.