

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

### **Title:** Associations of Socioeconomic Disparities With Buccal DNA-Methylation Measures Of Biological Aging

**Authors:** Raffington<sup>1</sup>, L., Schwaba<sup>2</sup>, T., Aikins<sup>1</sup>, M., Richter<sup>3,4</sup>, D., Wagner<sup>5,6,7</sup>, G.G., Harden<sup>2</sup>, K.P., Belsky<sup>8,9</sup>, D.W., & Tucker-Drob<sup>2</sup>, E.M.

#### **Affiliations:**

<sup>1</sup> Max Planck Research Group Biosocial – Biology, Social Disparities, and Development; Max Planck Institute for Human Development

<sup>2</sup> Population Research Center, The University of Texas at Austin, Austin, TX, USA

<sup>3</sup> SHARE Berlin Institute, Berlin, Germany

<sup>4</sup> Free University Berlin, Educational Science and Psychology, Berlin, Germany

<sup>5</sup> Max-Planck Institute for Human Development, Berlin, Germany

<sup>6</sup> Federal Institute for Population Research, Wiesbaden and Berlin, Germany

<sup>7</sup> German Socio-Economic Panel Study (SOEP), Berlin, Germany

<sup>8</sup> Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA

<sup>9</sup> Robert N Butler Columbia Aging Center, Columbia University Mailman School of Public Health, New York, NY, USA

**Corresponding Author:** Laurel Raffington, Max Planck Institute for Human Development; Max Planck Research Group Biosocial – Biology, Social Disparities, and Development; Lentzeallee 94, 14195 Berlin, Germany. E-mail: [raffington@mpib-berlin.mpg.de](mailto:raffington@mpib-berlin.mpg.de)

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

### Abstract

**Background:** Individuals who are socioeconomically disadvantaged are at increased risk for aging-related diseases and perform less well on tests of cognitive function. The Weathering Hypothesis proposes that these disparities in physical and cognitive health arise from an acceleration of biological processes of aging. Theories of how life adversity is biologically embedded identify epigenetic alterations, including DNA methylation (DNAm), as a mechanistic interface between the environment and health. Consistent with the Weathering hypothesis and theories of biological embedding, recently developed DNAm algorithms have revealed profiles reflective of more advanced aging and lower cognitive function among socioeconomically-at-risk groups. These DNAm algorithms were developed using blood-DNA, but social and behavioral science research commonly collect saliva or cheek-swab DNA. This discrepancy is a potential barrier to research to elucidate mechanisms through which socioeconomic disadvantage affects aging and cognition. We therefore tested if social gradients observed in blood-DNAm measures could be reproduced using buccal-cell DNA obtained from cheek swabs.

**Results:** We analyzed three DNAm measures of biological aging and one DNAm measure of cognitive performance, all of which showed socioeconomic gradients in previous studies: the PhenoAge and GrimAge DNAm clocks, DunedinPACE, and Epigenetic-*g*. We first computed blood-buccal cross-tissue correlations in  $n=21$  adults (GEO111165). Cross-tissue correlations were low-to-moderate across ( $r=.25$  to  $r=.48$ ). We next conducted analyses of socioeconomic gradients using buccal DNAm data from SOEP-G ( $n=1128$ , 57% female; age mean=42 yrs, SD=21.56, range 0-72). Associations of socioeconomic status with DNAm measures of aging were in the expected direction, but were smaller as compared to reports from blood DNAm datasets ( $r=-.08$  to  $r=-.13$ ).

**Conclusions:** Our findings are consistent with the hypothesis that socioeconomic disadvantage is associated with DNAm indicators of worse physical and cognitive health. However, relatively low cross-tissue correlations and attenuated effect-sizes for socioeconomic gradients in buccal DNAm compared with reports from analysis of blood DNAm suggest that, in order to take full advantage of buccal-DNA samples, DNAm algorithms customized to buccal DNAm are needed.

**Keywords:** aging; DNA methylation; biological aging; pace of aging; cognition; biomarker; lifespan; social determinants of health

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

### Background

Individuals who are socioeconomically disadvantaged are at increased risk for aging-related diseases and exhibit lower average levels of cognitive function across the life course, (Gkiouleka et al., 2018; Lövdén et al., 2020, p. 202; Tucker-Drob, 2019). Studies of humans and other animals identify several biological pathways through which social factors drive disease, including dysregulation of immune and metabolic systems in response to chronic stress (Snyder-Mackler et al., 2020). These pathways overlap substantially with the biology that mediates aging-related health declines (López-Otín et al., 2013). This overlap is consistent with the Weathering Hypothesis, which proposes that social adversity accelerates biological processes of aging (Geronimus et al., 2006).

Biological aging can be conceptualized as the progressive loss of system integrity that occurs with advancing chronological age (Kirkwood, 2005). The current state-of-the-art for quantification of biological aging in epidemiological studies of humans is a family of DNA methylation (DNAm) measurements. Epigenetic changes, including DNAm, are among the hallmarks of aging and are theorized to be key transducers of biological embedding of social adversity (Hertzman & Boyce, 2010; López-Otín et al., 2013). DNAm measures of biological aging that are most strongly predictive of disease, disability, and mortality are also consistently associated with social determinants of health (Oblak et al., 2021; Raffington & Belsky, 2022). In addition, there is evidence for social patterning of a DNAm measurement quantifying cognitive performance (McCartney et al., 2022), which parallels well-documented socioeconomic disparities in cognitive function across the life course (Lövdén et al., 2020). These DNAm measures open opportunities to study mechanisms of social disparities in physical and cognitive health and to guide the development and evaluation of interventions to address them.

A barrier to achieving this potential is that DNAm is specific to types of tissues and cells; it is a critical mechanism of cellular differentiation and determinant of cellular phenotype (Bakulski et al., 2016). Most DNAm algorithms used to study social gradients in health were developed from analysis of DNA derived from blood samples. Therefore, the ideal setting for their application is blood-derived DNA methylation. However, collection of blood samples is not feasible in some studies. For these studies, alternative sources of DNA, such as saliva and buccal tissue (*i.e.*, inner cheek) may be easier to obtain. The extent to which algorithms developed from blood-derived DNA can provide reliable and valid measurements in alternative tissues remains uncertain.

In two prior projects, we followed up algorithms developed to measure biological aging and cognitive functioning from blood DNAm in saliva samples collected from a pediatric cohort (Raffington, Belsky, et al., 2021; Raffington, Tanksley, et al., in press). In those studies, we were able to replicate several observations made from blood samples. First, the DNAm measure of the pace of biological aging (*i.e.*, a previous iteration of DunedinPACE) exhibited a parallel socioeconomic gradient in the pediatric saliva samples as had been observed previously in blood DNAm datasets from adults. Second, the DNAm measure of cognitive functioning

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

Epigenetic-*g* exhibited parallel association with children's performance on cognitive tests as had been observed previously in a blood DNAm dataset from adults. In contrast, the PhenoAge and GrimAge DNAm measures of biological age showed no social gradient in the pediatric saliva samples, in contrast to results from studies of blood samples (Schmitz et al., 2021).

Saliva is composed of a mix of leukocytes (which are also the source of blood-derived DNA samples) and epithelial cells. Buccal sample-derived DNA comes predominantly from epithelial cells. It is unclear whether DNAm measures computed in buccal DNAm will show similar evidence of trans-tissue validation. Here, we examined whether the same socioeconomic gradients in biological aging and DNAm-predicted cognitive performance apparent in blood DNAm analyses could be reproduced in analysis of buccal DNAm. The analysis we report is based on a pre-registration plan filed with OSF (<https://osf.io/msjgc>). Where our work has developed beyond this original pre-registration, we note it in the text. We first tested cross-tissue correlations of DNAm measures of biological aging (*i.e.*, PhenoAge Accel., GrimAge Accel., DunedinPACE) and DNAm-predicted cognitive performance (*i.e.*, Epigenetic-*g*) in buccal and blood DNAm datasets generated from the same individuals using the public dataset GEO111165 (n=21). Next, we examined association of chronological age with buccal DNAm measures in n=1128 participants from SOEP-G (57% female; age mean=42 yrs, SD=21.56, range 0-72). Finally, we tested associations of socioeconomic status with DNAm algorithms computed from buccal-cell DNAm in the same SOEP-G sample.

## Results

(1) Cross-tissue correlations between blood and buccal samples were low-to-moderate

We evaluated the correspondence between buccal and blood DNAm measures in an auxiliary dataset that collected both buccal and blood samples from the same n=21 people (Braun et al., 2019); Illumina EPIC array dataset in Gene Expression Omnibus accession GSE11116, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE111165>).

Cross-tissue correlations between blood and buccal samples of the DNAm measures were low-to-moderate across measures ( $r=0.25$  to  $r=0.48$ ). Means of DNAm measures were higher in buccal compared to blood samples, with the exception of Epigenetic-*g*, for which mean comparisons are not possible because beta-methylation values are standardized prior to computation (see **Table 1**).

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Table 1.** *Blood-buccal cross-tissue correlations of blood-based DNA-methylation measures (n=21).*

	Mean differences				Cross-tissue Correlation	
	Mean	SD	95% CI	p	r	
PhenoAge Accel.						0.25
	<i>blood</i>	7.25	8.35	45.03, 56.23	<0.001	
	<i>buccal</i>	57.88	9.56			
GrimAge Accel.						0.48
	<i>blood</i>	19.04	4.93	12.02, 18.08	<0.001	
	<i>buccal</i>	34.09	4.78			
DunedinPACE			0.39, 0.50	<0.001		0.31
	<i>blood</i>	1.07	0.11			
	<i>buccal</i>	1.52	0.06			
Epigenetic-g				–	–	0.46
	<i>blood</i>	0	0.33			
	<i>buccal</i>	0	0.23			

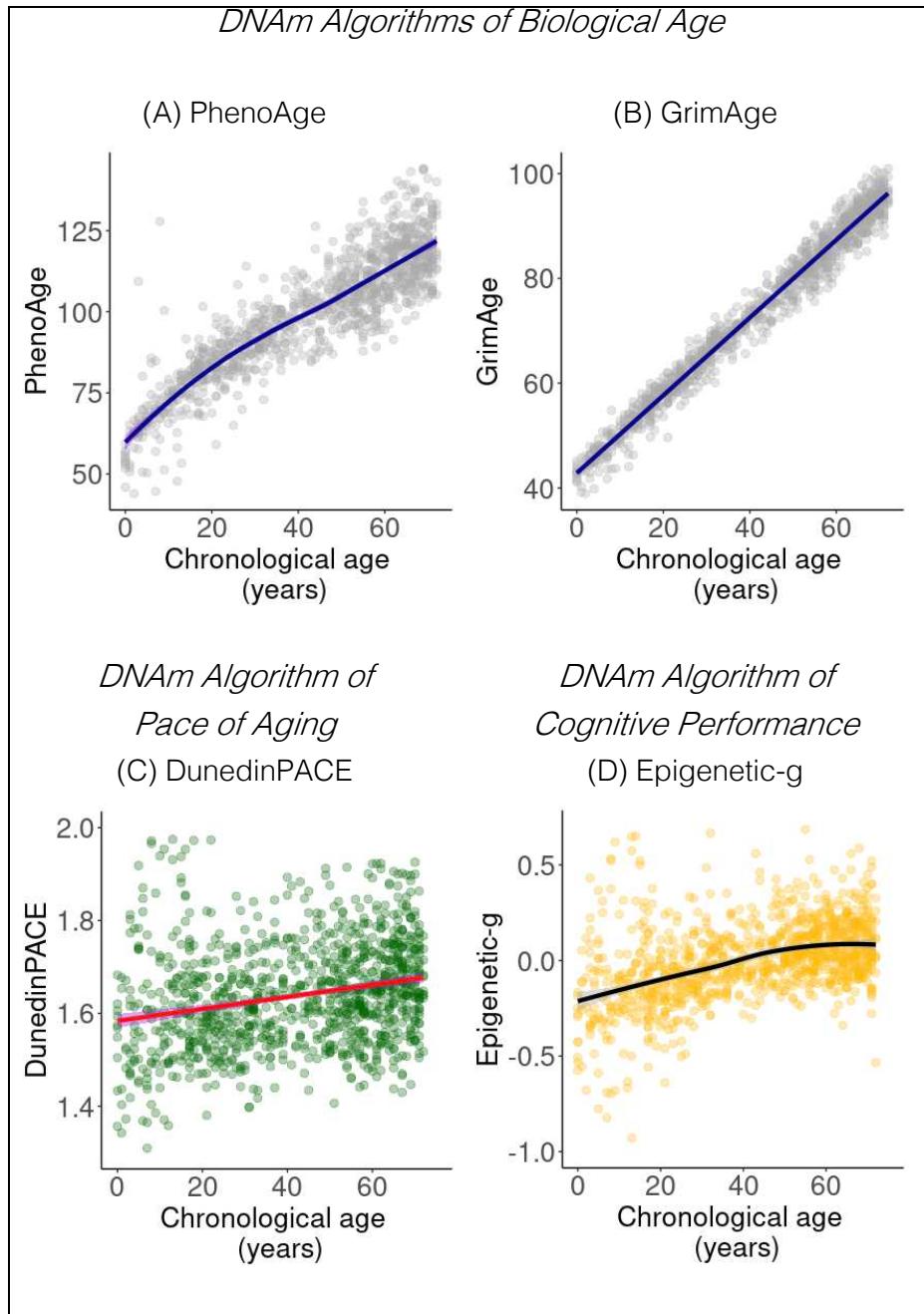
Means, standard deviations (SD), and blood-buccal cross-tissue correlations of DNA-methylation measures of accelerated biological aging (*i.e.*, PhenoAge Acceleration, GrimAge Acceleration), pace of aging (*i.e.*, DunedinPACE), and DNAm-predicted cognitive performance (*i.e.*, Epigenetic-*g*). Mean comparisons for Epigenetic-*g* are not possible because beta-methylation values are standardized prior to computation. Based on n=21 people from Gene Expression Omnibus accession GSE11116 (chronological age mean = 32.24, SD = 16.05).

## (2) Chronological age gradients in biological aging are reproduced in buccal DNAm

We examined associations of chronological age with buccal DNAm algorithms. For PhenoAge, strong association with chronological age is expected. In SOEP-G, participants' buccal DNAm PhenoAge values were highly correlated with their chronological ages (PhenoAge  $r=0.89$ , 95% CI = 0.88, 0.90,  $p<0.001$ ). GrimAge calculations include information about participant chronological age and, as a result, show very strong correlations ( $r=0.99$ , 95% CI = 0.99, 0.99,  $p<0.001$ ). In contrast to PhenoAge and GrimAge, which estimate biological age values, DunedinPACE estimates the pace of aging. Consistent with prior reports from blood DNAm datasets and with biodemography theory, which proposes that the pace of aging accelerates as we grow older (Belsky et al., 2022; Finch & Crimmins, 2016), participants' DunedinPACE values were moderately correlated with their chronological ages ( $r=0.24$ , 95% CI = 0.18, 0.29,  $p<0.001$ ). We also observed positive age trends for Epigenetic-*g*, mirroring known

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

patterns of cognitive development; values increased across the first half of the lifespan and then stabilized in late middle age ( $r=0.45$ , 95% CI = 0.40, 0.49,  $p<0.001$ ; age in years unstandardized  $b = 0.008$ , 95% CI = 0.006 – 0.011,  $p<0.001$ ; age squared unstandardized  $b = -0.001$ , 95% CI = -0.001- 0.000,  $p=0.001$ ). Age patterning of DNAm measures is shown in in **Figure 1**.



## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Figure 1. Chronological age and buccal DNAm algorithms.** Panel (A-B) plot associations of chronological age with buccal DNAm algorithms of biological aging, for which strong associations are expected: (A) PhenoAge and (B) GrimAge. Panel (C) plots association of chronological age with the pace of aging, DunedinPACE. Panel (D) plots association of chronological age with a DNAm algorithm of cognitive performance, Epigenetic-*g*.

### (3) Socioeconomic disadvantage is associated with accelerated biological aging in Germany

We tested associations of socioeconomic status (SES) with DNAm measures of biological aging computed from buccal-cell DNAm in SOEP-G. SES was measured as a composite of household income and educational levels (highest in household). Consistent with reports from blood DNAm datasets, participants with higher SES had younger biological ages and slower pace of aging ( $r's = -0.08$  to  $-0.13$ ,  $p's < 0.011$ , **Table 2**).

Next, according to our pre-registered analysis plan, we tested whether the association of SES with DNAm measures of aging differed by chronological age. This interaction was statistically significant for PhenoAge and GrimAge Acceleration (SES by continuous age interaction on PhenoAge std  $b = -0.11$ , 95% CI =  $-0.17$ ,  $-0.05$ ,  $p < 0.001$ ; Grimage std  $b = -0.07$ , 95% CI =  $-0.13$ ,  $-0.02$ ,  $p = 0.011$ ). There were no age differences in the SES association with DunedinPACE ( $p$ -value for continuous age interaction = 0.916). To further illustrate the interaction, we stratified the sample into older and younger participants. Among the older participants (aged > 42 years,  $n = 576$ ), the SES association with PhenoAge Acceleration was  $r = -0.14$ , 95% CI =  $-0.22$ ,  $-0.06$ ,  $p < 0.001$  and with GrimAge Acceleration was  $r = -0.18$ , 95% CI =  $-0.26$ ,  $-0.10$ ,  $p < 0.001$ . In contrast, among younger participants (aged < 42 years,  $n = 482$ ), the SES association with PhenoAge Acceleration was  $r = 0.03$ , 95% CI =  $-0.06$ ,  $0.12$ ,  $p = 0.494$  and with GrimAge Acceleration was  $r = -0.04$ , 95% CI =  $-0.13$ ,  $0.05$ ,  $p = 0.352$ . In sum, SES was associated with PhenoAge and GrimAge Acceleration only for older participants, whereas low SES was associated with DunedinPACE across age groups. **Figure 2** shows the association of socioeconomic status with DNAm by age.

Association of socioeconomic status with Epigenetic-*g* was in the expected direction, but was small and not statistically different from zero at the alpha=0.05 level (see **Table 4** and **Figure 2D**). Excluding smokers and accounting for body mass index did not substantially affect associations with SES (see **Figure S1**).

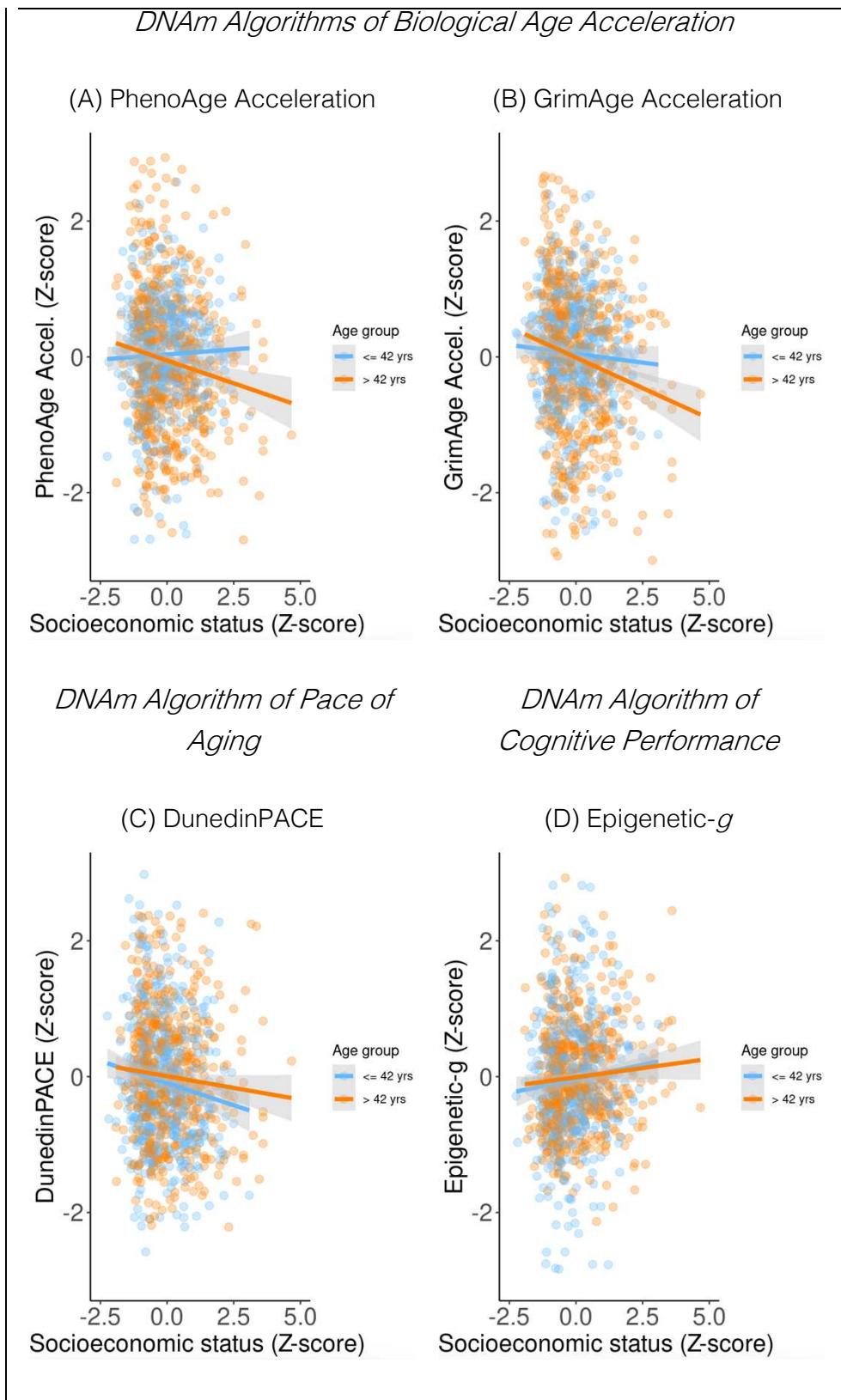
## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Table 2.** *Associations of socioeconomic status with buccal DNA-methylation measures.*

	<i>r</i>	95% CI	<i>P</i>
PhenoAge Accel.	-0.08	-0.139, -0.018	0.011
GrimAge Accel	-0.13	-0.190, -0.071	<0.001
DunedinPACE	-0.10	-0.154, -0.034	0.002
Epigenetic-g	0.06	-0.003, 0.117	0.064

Associations of socioeconomic status (average z-scored household income and education) with buccal DNA-methylation measures of accelerated biological aging (*i.e.*, PhenoAge Acceleration, GrimAge Acceleration), pace of biological aging (*i.e.*, DunedinPACE) and DNAm-predicted cognitive performance (*i.e.*, Epigenetic-*g*).

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION



## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Figure 2. Socioeconomic status and buccal DNAm algorithms.** Panel (A-B) plot associations of socioeconomic status with buccal DNAm algorithms of accelerated biological aging: (A) PhenoAge Acceleration and (B) GrimAge Acceleration. Panel (C) plots association of socioeconomic status with the pace of aging, DunedinPACE. Panel (D) plots association of socioeconomic status with a DNAm algorithm of cognitive performance, Epigenetic-*g*.

## Discussion

We tested if socioeconomic gradients in DNAm measurements of biological aging and cognitive performance, which are apparent in blood DNAm analyses, could be reproduced in analysis of buccal DNAm. Our findings are consistent with the Weathering Hypothesis that socioeconomic disadvantage is associated with accelerated biological aging. However, effect-sizes were approximately 50% lower than those reported in previously published analyses of blood DNAm datasets. Such studies have reported associations of magnitude of approximately  $r=.20$ , ranging from  $r=.10$  to  $r=.37$  (Raffington & Belsky, 2022), whereas here we report associations of magnitude of approximately  $r=.10$ , ranging from  $r=.079$  to  $r=.13$ . Similarly, associations of socioeconomic status with buccal DNAm-predicted cognitive performance were attenuated by approximately 50% and not statistically different from zero, in contrast to studies of blood and saliva DNAm datasets, which have reported associations with socioeconomic measures of magnitude  $r=.11$  and  $r=.14$  (McCartney et al., 2022; Raffington, Tanksley, et al., in press, note larger effect sizes for neighborhood-level socioeconomic contexts). Moreover, cross-tissue correspondence of DNAm indices was low-to-moderate. Collectively, these findings suggest that in order to take full advantage of buccal DNA samples, it will be important to develop DNAm indices that are customized to buccal DNAm.

One observation from our buccal DNAm data is that SES was associated with more PhenoAge Acceleration and GrimAge Acceleration only for older participants, whereas in the case of DunedinPACE the socioeconomic gradient was evident for both young and old participants. This pattern of results is consistent with findings from saliva DNAm in children and adolescents, which showed no association of PhenoAge and GrimAge with household SES, but did identify an association with DunedinPACE (Raffington, Belsky, et al., 2021). One possible explanation for this result is that measures of biological age, such as PhenoAge and GrimAge, which were designed to quantify differences in mortality risk among midlife and older adults, may be less sensitive to early stages in the biological embedding of social disadvantage. Replication of this result in other datasets and across tissues is needed.

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

### Conclusion

Our findings are consistent with the hypothesis that socioeconomic disadvantage is associated with accelerated biological aging in Germany. However, cross-tissue correspondence of DNAm indices was low-to-moderate and effect-sizes for SES associations estimated from buccal DNAm were attenuated by roughly 50% compared with reports from blood DNAm datasets. Development of DNAm measures of biological aging and cognitive performance that are customized to buccal DNAm should be a research priority.

### Methods

#### 1. Participants

SOEP-G participants were from the SOEP-IS cohort, which is based on a random sample of German households and contains a rich array of information on socioeconomic context, household dynamics, personality, and health (Koellinger et al., 2021). 6,576 people were originally invited to participate in the 2019 wave of the SOEP-IS with the aim to collect saliva for genotyping, 2598 of whom provided a valid genetic sample. ~98% of the genotyped SOEP-IS sample is of high genetic similarity to European reference groups. See Koellinger et al. (2021) for more information on the genotyped SOEP-IS cohort called SOEP-G.

Residual frozen DNA samples from n=1128 individuals from the n=2598 genotyped SOEP-G cohort were selected for DNAm extraction based on the availability of funds (see **Table 3** for descriptive statistics). Exclusion and inclusion criteria were: (1) exclusion of 5 samples due to sex mismatch between self-reported and genetic sex, (2) inclusion of all samples from children and adolescents (*i.e.*, under or equal to 18 yrs) whose residual DNA samples contained at least 50ng of DNA, (3) inclusion of adults that had (a) at least 250ng of DNA left, (b) had a DNA call rate of at least 0.975, (c) were not parents of selected children and adolescents so that the maximum number of different households were included, and (d) extended the age distribution continuously past 18 years so that all younger adults were included. The ID list was randomized so that plate effects were not confounded with chronological age. In addition, 24 samples were randomly selected as technical duplicates. The final sample of n=1128 unrelated participants (490 male, 638 female) consisted of 872 adults and 256 children and adolescents (age *mean*=41.88 yrs, *SD*=21.56, *range* 0-72, see supplementary Figure S1 for density plot of age distribution). 95% of participants were born in Germany.

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Table 3.** Descriptive statistics of the analytic sample after DNA-methylation based exclusions (N= 1058).

Sample	N	M	SD
Age (years)	1058	42.42	21.17
Sex, female	610	58%	–
DunedinPACE <sup>a</sup>	1058	1.64	0.11
PhenoAge <sup>b</sup>	1058	99.15	18.81
GrimAge <sup>c</sup>	1058	74.3	15.9
Epigenetic-g	1058	0	0.21
Household income (Euro)	1044	3318.07	1859.59
Household income / persons in household (Euro)	1044	1497.82	827.05
Maximum household education (years)	1042	13.34	2.76
Age- and sex-normed body mass index	876	22.55	5.8
PedBE	1058	30.21	10.86
Self-reported smoking, yes	87		

<sup>a</sup> A value of 1 reflects the average Pace of Aging in the Dunedin Study birth cohort over the age 26-45 follow-up period. A value of 1.01 therefore reflects a pace of aging 1% faster than the Dunedin Study norm.

<sup>b</sup> PhenoAge represents the age in years at which average mortality risk in NHANES III matches the mortality risk predicted by the PhenoAge algorithm.

<sup>c</sup> GrimAge represents the age in years at which average mortality risk in the Framingham Heart Study Offspring cohort matches predicted mortality risk.

## Measures

### *DNA-methylation preprocessing and exclusions.*

DNA was extracted from buccal swabs collected using Isohelix IS SK-1S Dri-Capsules (Koellinger et al 2021). DNA extraction and methylation profiling was conducted by the Human Genomics Facility (HuGe-F) at the Erasmus Medical Center in Rotterdam, Netherlands. The Infinium MethylationEPIC v1 manifest B5 kit (Illumina, Inc., San Diego, CA) was used to assess methylation levels at 865,918 CpG sites.

DNA methylation preprocessing was primarily conducted with Illumina's GenomeStudio software and open-source *R*(version 4.2.0) packages 'minfi' (Aryee et al., 2014) and 'ewastools' (Heiss & Just, 2018). We generated 20 control metrics in GenomeStudio as described in the BeadArray

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

Controls Reporter Software Guide from Illumina (note similar parameters can be computed using the `ewastools 'control_metrics()'` function). Samples falling below the Illumina-recommended cut-offs were flagged and further investigated. Flagged samples were classified as failed if 1. all types of poor bisulfite conversion and all types of poor bisulfite conversion background; 2. all types of bisulfite conversion background falling below 0.5; 3. all types of poor hybridization; 4. all types of poor specificity (excluded n=42).

As a second step, we identified unreliable data points resulting from low fluorescence intensities by filtering using detection p-values, calculated from comparing fluorescence intensities to a noise distribution. We removed probes with only background signal in a high proportion of samples (proportion of samples with detection p-value > 0.01 is > 0.1). We also removed probes for which a high proportion of samples had low bead numbers (proportion of samples with bead number < 3 is > 0.1). Further, we removed probes with SNPs at the CG or single base extension position as well as cross-reactive probes for EPIC arrays (McCartney et al., 2016; Pidsley et al., 2016).

We used minfi's 'preprocessNoob' (Triche et al., 2013) to correct for background noise and color dye bias and 'BMIQ' to account for probe-type differences (Teschendorff & Widschwendter, 2012).

Cell composition was estimated using HEpiDISH, which is an iterative hierarchical version of the EpiDISH *R* package using robust partial correlations (<https://github.com/sjczheng/EpiDISH>). Because epithelial cell types are the dominant cell type in buccal samples, we applied a threshold of 0.5 for epithelial cell proportions to reliably call a "buccal sample" and excluded samples that failed this metric (n=28). All samples were from the same batch. Final analytic sample size after DNAm exclusions was N=1058.

In GSE111165 blood samples, DNAm algorithms were residualized for reference-free cell composition and plate (Houseman et al., 2016).

### *DNA-methylation algorithms*

Our pre-registered analysis focused on two DNAm measures developed from blood DNAm data and which we had previously followed-up in saliva DNAm data (*i.e.*, DunedinPACE and Epigenetic-*g*) as well as a buccal-based algorithm of chronological age to be used as a data quality control measure (*i.e.*, PedBE). For comparative purposes, we report additional results for two further DNAm measures developed from blood DNAm, the PhenoAge and GrimAge clocks (Levine et al., 2018; Lu et al., 2019). We include these measures, which are among the best-evidenced DNAm biomarkers of aging, to help contextualize findings for DunedinPACE and Epigenetic-*g*. See **Table 4** for description of DNA-methylation algorithm computations.

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Table 4.** *Description of DNA-methylation algorithm computations.*

DNAm algorithm	Description
PhenoAge	<p>PhenoAge was first modeled from physiological markers and chronological age (Levine et al., 2018). This first-stage algorithm was then applied to a new sample in which it was modeled from DNA methylation to derive the final DNA methylation clock. PhenoAge represents the age in years at which average mortality risk in NHANES III matches the mortality risk predicted by the PhenoAge algorithm.</p> <p>PhenoAge was computed using DNAm principal components, which have been found to increase reliability (Higgins-Chen et al., 2022), using code available at <a href="https://github.com/MorganLevineLab/PC-Clocks">https://github.com/MorganLevineLab/PC-Clocks</a>. Using 24 technical replicates of samples in SOEP, we estimated the intraclass correlation coefficient (ICC). PhenoAge showed excellent reliability (ICC= 0.982). PhenoAge Acceleration was computed by residualizing PhenoAge for chronological age.</p>
GrimAge	<p>GrimAge was developed with a set of physiological indicators modeled from DNAm using machine learning analysis and then these DNA methylation algorithms along with age, sex, and a DNAm algorithm of smoking history were applied to model mortality (Lu et al., 2019). GrimAge represents the age in years at which average mortality risk in the Framingham Heart Study Offspring cohort matches predicted mortality risk.</p> <p>GrimAge was computed using DNAm principal components, which have been found to increase reliability (Higgins-Chen et al., 2022), using code available at <a href="https://github.com/MorganLevineLab/PC-Clocks">https://github.com/MorganLevineLab/PC-Clocks</a>. GrimAge showed excellent reliability (ICC= 0.999). GrimAge Acceleration was computed by residualizing GrimAge for chronological age.</p>
DunedinPACE	<p>DunedinPACE was developed as a DNA methylation measure of the pace of aging in the Dunedin Study birth cohort (Belsky et al., 2022). The Dunedin Study Pace of Aging is a composite phenotype derived from analysis of longitudinal change in biomarkers of organ-system integrity. Initially developed from analysis of three waves of biomarker data accumulated over a 12-year period (Belsky et al., 2015). Pace of Aging has recently been extended to a fourth measurement occasion spanning 20 years of follow-up (Elliott et al., 2021). DunedinPACE was developed from this second iteration of the Pace of Aging.</p> <p>Briefly, DNAm algorithm development was conducted using a subset of EPIC</p>

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

array probes that were also included on Illumina's earlier 450k array and that were identified as having relatively higher test-retest reliability (Sugden et al., 2020). Elastic-net regression machine learning analysis was used to fit Pace of Aging to DNAm data generated from blood samples collected when participants were aged 45 years. The elastic net regression produced a 173-CpG algorithm. Increments of DunedinPACE correspond to "years" of physiological change occurring per 12-months of chronological time. A value of 1 reflects the average Pace of Aging in the Dunedin Study birth cohort over the age 26-45 follow-up period. A value of 1.01 therefore reflects a pace of aging 1% faster than the Dunedin Study norm. DunedinPACE was calculated based on the published algorithm using code available at <https://github.com/danbelsky/DunedinPACE/>. 14 of the 173 CpG probes that are part of DunedinPACE were not present in our dataset. Buccal DunedinPACE showed good reliability (ICC= 0.74).

**Epigenetic-*g*** Epigenetic-*g* was computed using a blood-based algorithm from an epigenome-wide association study (EWAS) in BayesR+ of general cognitive function (*g*) in 9162 adults (59% females; mean age 49.8 years, SD 13.6, range 18–93) in the Generation Scotland Study (McCartney et al., 2022). Briefly, general cognitive function was derived from the first unrotated principal component of logical memory, verbal fluency and digit symbol tests, and vocabulary. Cognitive phenotypes were corrected for age, sex, BMI and an epigenetic smoking score. Epigenetic-*g* includes all CpG sites in the EWAS. The weights for each CpG are the mean posterior effect sizes from the EWAS model of *g*. Prior to computation of Epigenetic-*g* in the present study, methylation values were scaled within each CpG site (*mean* = 0, SD = 1) and calculated based on the published algorithm using code available at [https://gitlab.com/danielmccartney/ewas\\_of\\_cognitive\\_funct](https://gitlab.com/danielmccartney/ewas_of_cognitive_funct). Epigenetic-*g* showed good reliability (ICC= 0.84).

**PedBE** As a data quality control, we examined associations of chronological age with the Pediatric-Buccal-Epigenetic (PedBE) clock, which was computed using a buccal-based algorithm of chronological age in individuals aged 0 to 20 years old (McEwen et al., 2019). Elastic net penalized regression was used to select 94 CpGs from a training dataset of 1032 subjects. PedBE was calculated based on the published algorithm using code available at <https://github.com/kobor-lab/Public-Scripts/blob/master/PedBE.Md>. All 94 CpG probes were present in our dataset. PedBE showed excellent reliability (ICC= 0.967). PedBE was strongly associated with chronological age, indicating good data quality ( $r = 0.91$ , 95% CI = 0.90, 0.92,  $p < 0.001$ ).

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Socioeconomic status.** We deviated from our pre-registered analysis plan by testing associations with socioeconomic status (average z-scored household income and education) rather than examining income and education separately, to reduce the number of statistical comparisons. Monthly household net income in Euros from all sources (e.g. employment, pensions, unemployment benefits, maternity benefits, higher education grants, military or civil service pay, compulsory child support, etc.) was reported by the self-defined head of household. In the 2% of cases with missing income values, information about determinants of household income and past data were used to impute estimated values (for more information see page 27 [https://www.diw.de/documents/publikationen/73/diw\\_01.c.787445.de/diw\\_ssp0844.pdf](https://www.diw.de/documents/publikationen/73/diw_01.c.787445.de/diw_ssp0844.pdf)).

Household income was divided by the number of persons in the household and sqrt transformed to correct for skew (this deviated from our preregistration plan; sqrt-transformation improved normality of distribution more than log-transformation in shapiro wilks test).

Given the wide age range of participants, we indexed educational attainment as the highest degree obtained by any individual in the household. Educational attainment was converted to number of educational years (no degree = 7 years, lower school degree = 9 years, intermediary school = 10 years, degree for, a professional coll. = 12 years, high school degree = 13 years, other = 10 years) with additional occupational training added (apprenticeship = +1.5 years, technical schools (including health) = +2 years, civil servants apprenticeship = +1.5 years, higher technical college = +3 years, university degree = +5 years).

### *Covariates.*

**Body mass index** (BMI). Height (in cm) and weight (in kg) were measured via self-report and transformed to sex- and age-normed BMI z-scores

**Smoking.** Participant self-reported current or past smoking across multiple waves. Across questions and waves, if a participant ever responded that they smoked currently or in the past, they were identified as a smoker. If a participant ever responded that they never smoked and never responded that they did smoke, they were identified as a never-smoker.

## Declarations

**Ethics approval and consent to participate.** Participants or the parents of minor participants consented to the archiving, extraction, and analysis of DNA-based measures. Ethical approval for DNA-based research was received by the Research Ethics Review Board of Vrije Universiteit Amsterdam, School of Business and Economics (application number 20181018.1.pkr730) and the Ethics Council of the Max Planck Society (application number 2019\_16).

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

**Consent for publication.** Not applicable.

**Availability of data and materials.** The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests.** The authors declare that they have no competing interests.

**Funding.** This research was supported by National Institutes of Health (NIH) grants R01HD083613 and R01HD092548. KPH and EMTD are Faculty Research Associates of the Population Research Center at the University of Texas at Austin, which is supported by a NIH grant P2CHD042849. EMTD is a member of the Center on Aging and Population Sciences (CAPS) at The University of Texas at Austin, which is supported by NIH grant P30AG066614. KPH and EMTD were also supported by Jacobs Foundation Research Fellowships. DWB received supported from R01AG073402. The funding bodies had no role in the design, collection, analysis or interpretation of the study.

**Authors' contributions.** Laurel Raffington developed the study concept and design, performed and supervised data analysis, and drafted the manuscript. Ted Schwaba performed data analysis and provided critical revisions. Muna Aikins provided critical revisions. David Richter and Gert G. Wagner developed the study concept and provided critical revisions. K. Paige Harden, Daniel W. Belsky, and Elliot M. Tucker-Drob developed the study concept and design, supervised data analysis, drafted the manuscript, and provided critical revisions. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

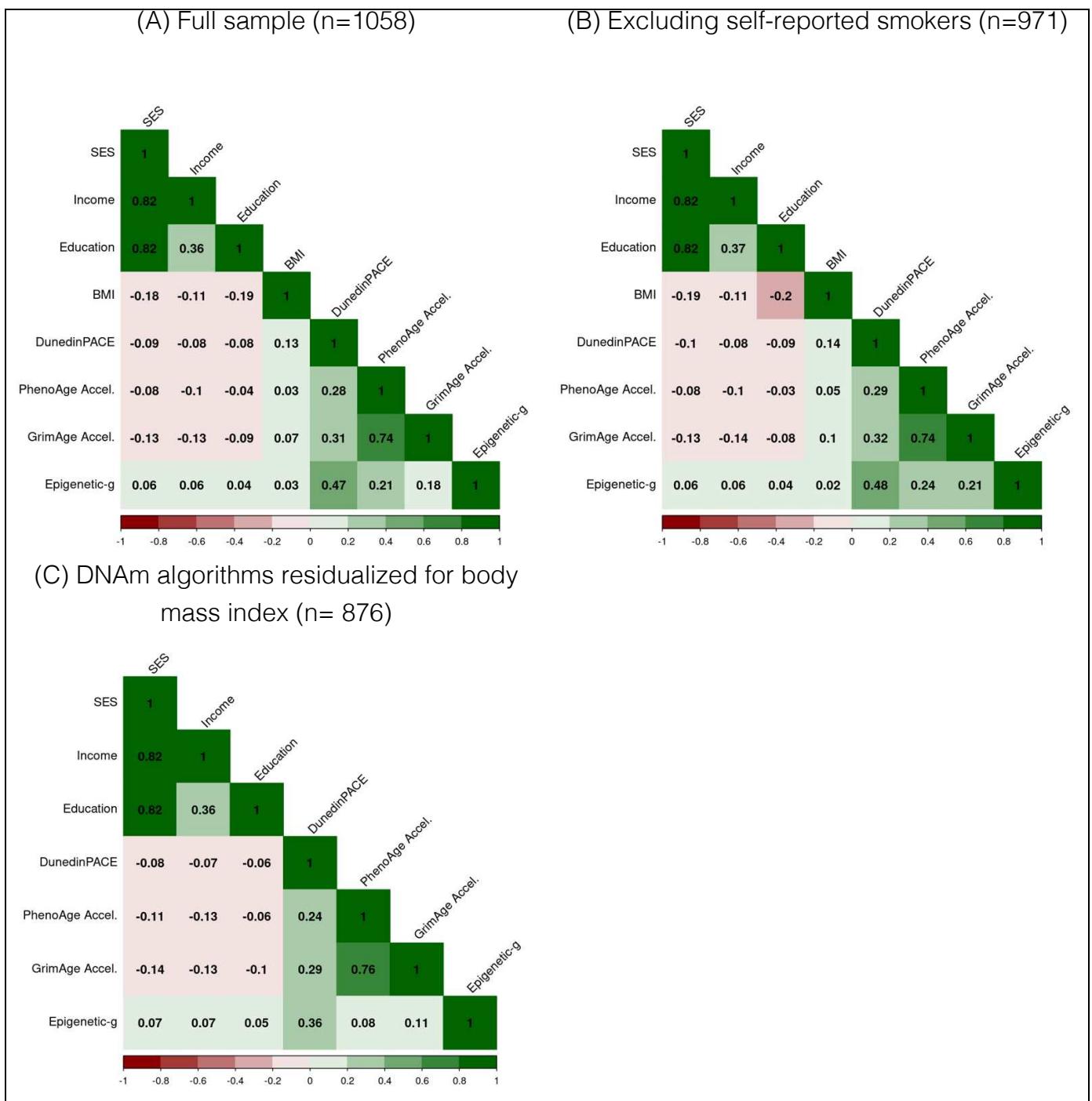
**Acknowledgements.** We thank the participants of SOEP-G.

## List of abbreviations

- DNAm – DNA methylation
- SES – Socioeconomic status

## Supplemental Results

## Supplemental Figures



**Supplementary Figure S1.** Correlation matrix of socioeconomic variables with buccal DNA-methylation (DNAm) algorithms in (A) full sample, (B) excluding self-reported smokers, and (C) residualizing DNAm algorithms for body mass index (BMI). In all plots DNAm measures were residualized for chronological age. BMI was age- and sex-normed.

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

### References

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, 30(10), 1363–1369. <https://doi.org/10.1093/bioinformatics/btu049>

Bakulski, K. M., Halladay, A., Hu, V. W., Mill, J., & Fallin, M. D. (2016). Epigenetic Research in Neuropsychiatric Disorders: The “Tissue Issue.” *Current Behavioral Neuroscience Reports*, 3(3), 264–274. <https://doi.org/10.1007/s40473-016-0083-4>

Belsky, D. W., Caspi, A., Corcoran, D. L., Sugden, K., Poulton, R., Arseneault, L., Baccarelli, A., Chamarti, K., Gao, X., Hannon, E., Harrington, H. L., Houts, R., Kothari, M., Kwon, D., Mill, J., Schwartz, J., Vokonas, P., Wang, C., Williams, B. S., & Moffitt, T. E. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife*, 11, e73420. <https://doi.org/10.7554/eLife.73420>

Belsky, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., Harrington, H., Israel, S., Levine, M. E., Schaefer, J. D., Sugden, K., Williams, B., Yashin, A. I., Poulton, R., & Moffitt, T. E. (2015). Quantification of biological aging in young adults. *Proceedings of the National Academy of Sciences*, 112(30), E4104–E4110. <https://doi.org/10.1073/pnas.1506264112>

Braun, P. R., Han, S., Hing, B., Nagahama, Y., Gaul, L. N., Heinzman, J. T., Grossbach, A. J., Close, L., Dlouhy, B. J., Howard, M. A., Kawasaki, H., Potash, J. B., & Shinozaki, G. (2019). Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Translational Psychiatry*, 9(1), 47. <https://doi.org/10.1038/s41398-019-0376-y>

Elliott, M. L., Caspi, A., Houts, R. M., Ambler, A., Broadbent, J. M., Hancox, R. J., Harrington, H., Hogan, S., Keenan, R., Knodt, A., Leung, J. H., Melzer, T. R., Purdy, S. C., Ramrakha, S., Richmond-Rakerd, L. S., Righarts, A., Sugden, K., Thomson, W. M., Thorne, P. R., ... Moffitt, T. E. (2021). Disparities in the pace of biological aging among midlife adults of the same chronological age have implications for future frailty risk and policy. *Nature Aging*, 1(3), 295–308. <https://doi.org/10.1038/s43587-021-00044-4>

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

Finch, C. E., & Crimmins, E. M. (2016). Constant molecular aging rates vs. The exponential acceleration of mortality. *Proceedings of the National Academy of Sciences*, 113(5), 1121–1123. <https://doi.org/10.1073/pnas.1524017113>

Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). “Weathering” and Age Patterns of Allostatic Load Scores Among Blacks and Whites in the United States. *American Journal of Public Health*, 96(5), 8.

Gkiouleka, A., Huijts, T., Beckfield, J., & Bambra, C. (2018). Understanding the micro and macro politics of health: Inequalities, intersectionality & institutions - A research agenda. *Social Science & Medicine*, 200, 92–98.  
<https://doi.org/10.1016/j.socscimed.2018.01.025>

Heiss, J. A., & Just, A. C. (2018). Identifying mislabeled and contaminated DNA methylation microarray data: An extended quality control toolset with examples from GEO. *Clinical Epigenetics*, 10(1), 73. <https://doi.org/10.1186/s13148-018-0504-1>

Hertzman, C., & Boyce, T. (2010). How Experience Gets Under the Skin to Create Gradients in Developmental Health. *Annual Review of Public Health*, 31(1), 329–347.  
<https://doi.org/10.1146/annurev.publhealth.012809.103538>

Higgins-Chen, A. T., Thrush, K. L., Wang, Y., Minteer, C. J., Kuo, P.-L., Wang, M., Niimi, P., Sturm, G., Lin, J., Moore, A. Z., Bandinelli, S., Vinkers, C. H., Vermetten, E., Rutten, B. P. F., Geuze, E., Okhuijsen-Pfeifer, C., van der Horst, M. Z., Schreiter, S., Gutwinski, S., ... Levine, M. E. (2022). A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. *Nature Aging*, 2(7), 644–661. <https://doi.org/10.1038/s43587-022-00248-2>

Houseman, E. A., Kile, M. L., Christiani, D. C., Ince, T. A., Kelsey, K. T., & Marsit, C. J. (2016). Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. *BMC Bioinformatics*, 17(1), 259. <https://doi.org/10.1186/s12859-016-1140-4>

Kirkwood, T. B. L. (2005). Understanding the Odd Science of Aging. *Cell*, 120(4), 437–447.  
<https://doi.org/10.1016/j.cell.2005.01.027>

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

Koellinger, P. D., Okbay, A., Kweon, H., Schweinert, A., Karlsson Linnér, R., Goebel, J., Richter, D., Reiber, L., Zweck, B. M., Belsky, D. W., Biroli, P., Mata, R., Tucker-Drob, E. M., Harden, K. P., Wagner, G., & Hertwig, R. (2021). *Cohort Profile: Genetic data in the German Socio-Economic Panel Innovation Sample (Gene-SOEP)* [Preprint]. Genetics. <https://doi.org/10.1101/2021.11.06.467573>

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, 10(4), 573–591. <https://doi.org/10.18632/aging.101414>

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., & Kroemer, G. (2013). The Hallmarks of Aging. *Cell*, 153(6), 1194–1217. <https://doi.org/10.1016/j.cell.2013.05.039>

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and Cognitive Functioning Across the Life Span. *Psychological Science in the Public Interest*, 21(1), 6–41. <https://doi.org/10.1177/1529100620920576>

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, 11(2), 303–327. <https://doi.org/10.18632/aging.101684>

McCartney, D. L., Hillary, R. F., Conole, E. L. S., Banos, D. T., Gadd, D. A., Walker, R. M., Nangle, C., Flaig, R., Campbell, A., Murray, A. D., Maniega, S. M., Valdés-Hernández, M. del C., Harris, M. A., Bastin, M. E., Wardlaw, J. M., Harris, S. E., Porteous, D. J., Tucker-Drob, E. M., McIntosh, A. M., ... Marioni, R. E. (2022). Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biology*, 23(1), 26. <https://doi.org/10.1186/s13059-021-02596-5>

McCartney, D. L., Walker, R. M., Morris, S. W., McIntosh, A. M., Porteous, D. J., & Evans, K. L. (2016). Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. *Genomics Data*, 9, 22–24. <https://doi.org/10.1016/j.gdata.2016.05.012>

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

McEwen, L. M., O'Donnell, K. J., McGill, M. G., Edgar, R. D., Jones, M. J., MacIsaac, J. L., Lin, D. T. S., Ramadori, K., Morin, A., Gladish, N., Garg, E., Unternaehrer, E., Pokhvisneva, I., Karnani, N., Kee, M. Z. L., Klengel, T., Adler, N. E., Barr, R. G., Letourneau, N., ... Kobor, M. S. (2019). The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proceedings of the National Academy of Sciences*, 201820843. <https://doi.org/10.1073/pnas.1820843116>

Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E., & Boks, M. P. (2021). A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Research Reviews*, 69, 101348. <https://doi.org/10.1016/j.arr.2021.101348>

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., Van Djik, S., Muhlhausler, B., Stirzaker, C., & Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, 17(1), 208. <https://doi.org/10.1186/s13059-016-1066-1>

Raffington, L., & Belsky, D. W. (2022). Integrating DNA Methylation Measures of Biological Aging into Social Determinants of Health Research. *Current Environmental Health Reports*, 15.

Raffington, L., Belsky, D. W., Kothari, M., Malanchini, M., Tucker-Drob, E. M., & Harden, K. P. (2021). Socioeconomic Disadvantage and the Pace of Biological Aging in Children. *Pediatrics*, e2020024406. <https://doi.org/10.1542/peds.2020-024406>

Raffington, L., Tanksley, P. T., Sabhlok, A., Vinnik, L., Mallard, T., King, L. S., Goosby, B., Harden, K. P., & Tucker-Drob, E. M. (2021). *Socially stratified epigenetic profiles are associated with cognitive functioning in children and adolescents* [Preprint]. Genetics. <https://doi.org/10.1101/2021.08.19.456979>

Schmitz, L. L., Zhao, W., Ratliff, S. M., Goodwin, J., Miao, J., Lu, Q., Guo, X., Taylor, K. D., Ding, J., Liu, Y., Levine, M., & Smith, J. A. (2021). *The Socioeconomic Gradient in Epigenetic Aging Clocks: Evidence from the Multi-Ethnic Study of Atherosclerosis and the Health and Retirement Study* [Preprint]. Epidemiology. <https://doi.org/10.1101/2021.03.01.21252660>

## SOCIOECONOMIC DISPARITIES AND BUCCAL DNA-METHYLATION

Snyder-Mackler, N., Burger, J. R., Gaydosh, L., Belsky, D. W., Noppert, G. A., Campos, F. A., Bartolomucci, A., Yang, Y. C., Aiello, A. E., O'Rand, A., Harris, K. M., Shively, C. A., Alberts, S. C., & Tung, J. (2020). Social determinants of health and survival in humans and other animals. *Science*, 368(6493). <https://doi.org/10.1126/science.aax9553>

Sugden, K., Hannon, E. J., Arseneault, L., Belsky, D. W., Corcoran, D. L., Fisher, H. L., Houts, R. M., Kandaswamy, R., Moffitt, T. E., Poulton, R., Prinz, J. A., Rasmussen, L. J. H., Williams, B. S., Wong, C. C. Y., Mill, J., & Caspi, A. (2020). Patterns of Reliability: Assessing the Reproducibility and Integrity of DNA Methylation Measurement. *Patterns*, 100014. <https://doi.org/10.1016/j.patter.2020.100014>

Teschendorff, A. E., & Widschwendter, M. (2012). Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. *Bioinformatics*, 28(11), 1487–1494.  
<https://doi.org/10.1093/bioinformatics/bts170>

Triche, T. J., Weisenberger, D. J., Van Den Berg, D., Laird, P. W., & Siegmund, K. D. (2013). Low-level processing of Illumina Infinium DNA Methylation BeadArrays. *Nucleic Acids Research*, 41(7), e90–e90. <https://doi.org/10.1093/nar/gkt090>

Tucker-Drob, E. M. (2019). Cognitive Aging and Dementia: A Life-Span Perspective. *Annual Review of Developmental Psychology*, 1(1), 177–196. <https://doi.org/10.1146/annurev-devpsych-121318-085204>