

RESEARCH

Open Access



Associations of socioeconomic disparities with buccal DNA-methylation measures of biological aging

L. Raffington^{1*}, T. Schwaba², M. Aikins¹, D. Richter^{3,4}, G. G. Wagner^{5,6,7}, K. P. Harden², D. W. Belsky^{8,9} and E. M. Tucker-Drob²

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 ($r = .25$ 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 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

*Correspondence:

L. Raffington

raffington@mpib-berlin.mpg.de

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

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 [1–3]. 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 [4]. These pathways overlap substantially with the biology that mediates aging-related health declines [5]. This overlap is consistent with the weathering hypothesis, which proposes that social adversity accelerates biological processes of aging [6].

Biological aging can be conceptualized as the progressive loss of system integrity that occurs with advancing chronological age [7]. 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 [5, 8]. Consistent with the hypothesis that DNAm measures of biological aging capture both of these epigenetic mechanisms, the DNAm measures of biological aging that show the most consistent and strongest prediction of aging-related disease, disability, and mortality also show consistent associations with social determinants of health (i.e., the GrimAge, PhenoAge, and DunedinPACE epigenetic clocks; [9, 10]). The current study, therefore, focuses on these DNAm measures. Consistent with the hypothesis that DNAm measures of biological aging capture both of these epigenetic mechanisms, the DNAm measures of biological aging that show the most consistent and strongest prediction of aging-related disease, disability, and mortality also show consistent associations with social determinants of health (i.e., the GrimAge, PhenoAge, and DunedinPACE epigenetic clocks) [9, 10]. In addition, there is evidence for social patterning of a DNAm measurement quantifying cognitive performance (i.e., Epigenetic-g) [11], which parallels well-documented socioeconomic disparities in cognitive function across the life course [2]. 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 [12]. 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 [13, 14]. 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 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 [15].

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

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 [16], 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).

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 [17, 18], participants’ DunedinPACE values were modestly 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 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, $p<0.001$; and age squared unstandardized $b=-0.001$, 95% CI=-0.001–0.000, $p=0.001$). Age patterning of DNAm measures is shown in Fig. 1.

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

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

	Mean differences				Cross-tissue correlation r
	Mean	SD	95% CI	p	
PhenoAge Accel			45.03, 56.23	< 0.001	0.25
Blood	7.25	8.35			
Buccal	57.88	9.56			
GrimAge Accel			12.02, 18.08	< 0.001	0.48
Blood	19.04	4.93			
Buccal	34.09	4.78			
DunedinPACE			0.39, 0.50	< 0.001	0.31
Blood	1.07	0.11			
Buccal	1.52	0.06			
Epigenetic- g			–	–	0.46
Blood	0	0.33			
Buccal	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)

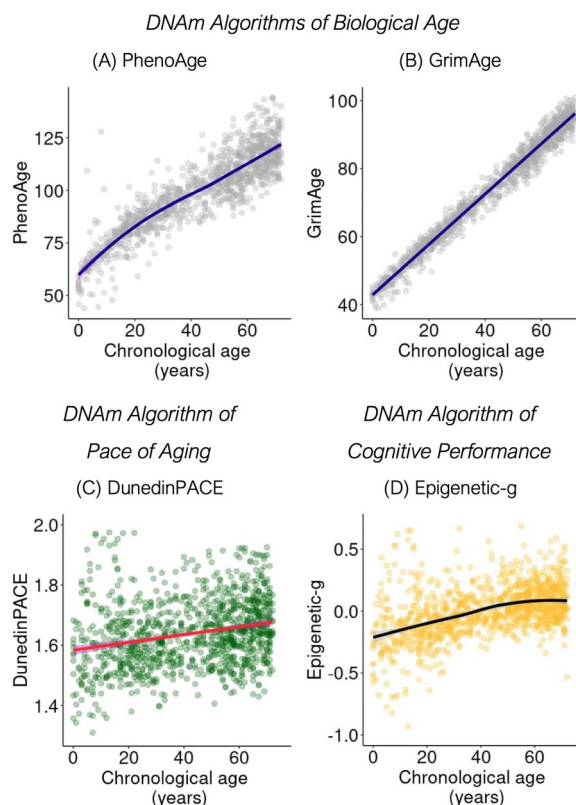


Fig. 1 Chronological age and buccal DNAm algorithms. Panel **A–B** plot associations of chronological age with buccal DNAm algorithms of biological aging, for which strong associations are expected: **A** PhenoAge and **B** GrimAge. Panel **C** plots association of chronological age with the pace of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

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)

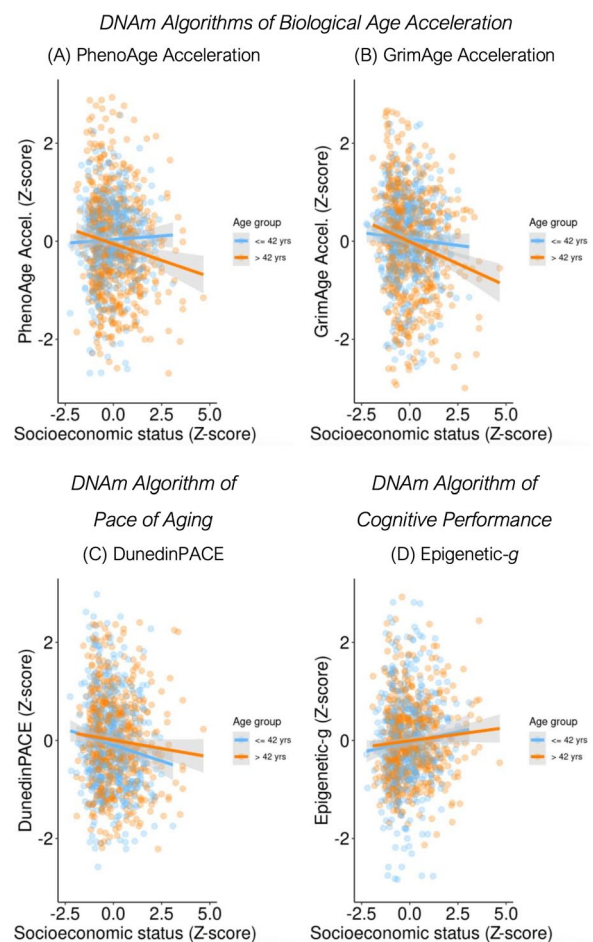
from blood DNAm datasets, participants with higher SES had younger biological ages and slower pace of aging (r 's = − 0.08 to − 0.13, p 's < 0.011, Table 2).

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 (mean split). 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 α = 0.05 level (see Table S4 and Fig. 2D). Excluding smokers and accounting for body mass index did not substantially affect associations with SES (see Additional file 1: Figure S1).

Discussion

We tested if socioeconomic gradients in DNAm measurements of biological aging and cognitive performance, which are apparent in blood DNAm analyses, could be

**Fig. 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

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 [10], 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 $r=.14$ [11, 19] (note larger effect sizes for neighborhood-level socioeconomic contexts in 19). 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 [13]. 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.

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

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 [20]. Six thousand five hundred seventy-six 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. [20] 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 five 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 50 ng of DNA, (3) inclusion of adults that had (a) at least 250 ng 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 Additional file 1: Figure S1 for density plot of age distribution). 95% of participants were born in Germany.

Measures

DNA-methylation preprocessing and exclusions

DNA was extracted from buccal swabs collected using Isohelix IS SK-1S Dri-Capsules [20]. 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)

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 ^a	1058	1.64	0.11
PhenoAge ^b	1058	99.15	18.81
GrimAge ^c	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		

^a 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

^b PhenoAge represents the age in years at which average mortality risk in NHANES III matches the mortality risk predicted by the PhenoAge algorithm

^c GrimAge represents the age in years at which average mortality risk in the Framingham Heart Study Offspring cohort matches predicted mortality risk

was used to assess methylation levels at 865,918 CpG sites.

DNAm preprocessing was primarily conducted with Illumina's GenomeStudio software and open-source *R* (version 4.2.0) packages 'minfi' [21] and 'ewastools' [22]. We generated 20 control metrics in GenomeStudio as described in the BeadArray 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; and 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 [23, 24].

We used minfi's 'preprocessNoob' [25] to correct for background noise and color dye bias and 'BMIQ' to account for probe-type differences [26].

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 [27].

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 [28, 29]. 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 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). Household income was divided by the number of persons in the household and sqrt transformed to correct for skew (this deviated from our pre-registration plan; sqrt-transformation improved normality of distribution more than log-transformation in Shapiro–Wilk 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.

Table 4 Description of DNA-methylation algorithm computations

DNAm algorithm	Description
PhenoAge	PhenoAge was first modeled from physiological markers and chronological age [28]. 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. PhenoAge was computed using DNAm principal components, which have been found to increase reliability [30], using code available at https://github.com/MorganLevineLab/PC-Clocks . 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.
GrimAge	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 [29]. GrimAge represents the age in years at which average mortality risk in the Framingham Heart Study Offspring cohort matches predicted mortality risk. GrimAge was computed using DNAm principal components, which have been found to increase reliability [30], using code available at https://github.com/MorganLevineLab/PC-Clocks . GrimAge showed excellent reliability (ICC = 0.999). GrimAge Acceleration was computed by residualizing GrimAge for chronological age.
DunedinPACE	DunedinPACE was developed as a DNA-methylation measure of the pace of aging in the Dunedin Study birth cohort [17]. 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 [31]. Pace of Aging has recently been extended to a fourth measurement occasion spanning 20 years of follow-up [32]. DunedinPACE was developed from this second iteration of the Pace of Aging. Briefly, DNAm algorithm development was conducted using a subset of EPIC array probes that were also included on Illumina's earlier 450 k array and that were identified as having relatively higher test–retest reliability [33]. 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 in 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/danbe/lsky/DunedinPACE/ . Fourteen 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 (<i>g</i>) in 9162 adults (59% females; mean age 49.8 years, SD 13.6, range 18–93) in the Generation Scotland Study [11]. 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 <i>g</i> . 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 . 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 developed to predict chronological age in individuals aged < 20 years from buccal-cell DNAm, i.e., the same tissue type examined here [34]. While the pediatric sample used to develop the PedBE clock is considerably younger than our lifespan sample, it is one of the few aging-related DNAm indicators developed using buccal cells. Correspondence between PedBE and chronological age in our sample increases confidence in the quality of our buccal-cell DNAm data. 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, $p < 0.001$).

Abbreviations

DNAm DNA methylation
SES Socioeconomic status

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-023-01489-7>.

Additional file 1: Fig. S1. Correlation matrix of socioeconomic variables with buccal DNA-methylation algorithms in full sample, excluding self-reported smokers, and residualizing DNAm algorithms for body mass index. In all plots, DNAm measures were residualized for chronological age. BMI was age- and sex-normed.

Acknowledgements

We thank the participants of SOEP-G.

Author contributions

LR developed the study concept and design, performed and supervised data analysis, and drafted the manuscript. TS performed data analysis and provided critical revisions. MA provided critical revisions. DR and GGW developed the study concept and provided critical revisions. KPH, DWB, and EMTD 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. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. 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 support from R01AG073402. KPH received support from R01HD092548. EMTD received support from RF1AG073593. The funding bodies had no role in the design, collection, analysis or interpretation of the study.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

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).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Max Planck Research Group Biosocial – Biology, Social Disparities, and Development, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. ²Department of Psychology, The University of Texas at Austin, Austin, TX, USA. ³SHARE Berlin Institute, Berlin, Germany. ⁴Educational Science and Psychology, Free University Berlin, Berlin, Germany. ⁵Max Planck Institute for Human Development, Berlin, Germany. ⁶Federal Institute for Population Research, Wiesbaden, Berlin, Germany. ⁷German Socio-Economic Panel Study (SOEP), Berlin, Germany. ⁸Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA. ⁹Robert N Butler Columbia Aging Center, Columbia University Mailman School of Public Health, New York, NY, USA.

Received: 7 December 2022 Accepted: 20 April 2023

Published online: 28 April 2023

References

- Gkiouleka A, Huijts T, Beckfield J, Bamba C. Understanding the micro and macro politics of health: inequalities, intersectionality & institutions—a research agenda. *Soc Sci Med*. 2018;200:92–8. <https://doi.org/10.1016/j.socscimed.2018.01.025>.
- Lövdén M, Fratiglioni L, Glymour MM, Lindenberg U, Tucker-Drob EM. Education and cognitive functioning across the life span. *Psychol Sci Public Interest*. 2020;21(1):6–41. <https://doi.org/10.1177/1529100620920576>.
- Tucker-Drob EM. Cognitive aging and dementia: a life-span perspective. *Annu Rev Dev Psychol*. 2019;1(1):177–96. <https://doi.org/10.1146/annurev-devpsych-121318-085204>.
- Snyder-Mackler N, Burger JR, Gaydos L, Belsky DW, Noppert GA, Campos FA, Bartolomucci A, Yang YC, Aiello AE, O’Rand A, Harris KM, Shively CA, Alberts SC, Tung J. Social determinants of health and survival in humans and other animals. *Science*. 2020. <https://doi.org/10.1126/science.aax9553>.
- López-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153(6):1194–217. <https://doi.org/10.1016/j.cell.2013.05.039>.
- Geronimus AT, Hicken M, Keene D, Bound J. “Weathering” and age patterns of allostatic load scores among blacks and whites in the United States. *Am J Public Health*. 2006;96(5):8.
- Kirkwood TBL. Understanding the odd science of aging. *Cell*. 2005;120(4):437–47. <https://doi.org/10.1016/j.cell.2005.01.027>.
- Hertzman C, Boyce T. How experience gets under the skin to create gradients in developmental health. *Annu Rev Public Health*. 2010;31(1):329–47. <https://doi.org/10.1146/annurev.publhealth.012809.103538>.
- Oblak L, van der Zaag J, Higgins-Chen AT, Levine ME, Boks MP. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Age Res Rev*. 2021;69:101348. <https://doi.org/10.1016/j.arr.2021.101348>.
- Raffington L, Belsky DW. Integrating DNA methylation measures of biological aging into social determinants of health research. *Curr Environ Health Rep*. 2022;15:196.
- McCartney DL, Hillary RF, Conole ELS, Banos DT, Gadd DA, Walker RM, Nangle C, Flaig R, Campbell A, Murray AD, Maniega SM, Valdés-Hernández MC, Harris MA, Bastin ME, Wardlaw JM, Harris SE, Porteous DJ, Tucker-Drob EM, McIntosh AM, Marioni RE. Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biol*. 2022;23(1):26. <https://doi.org/10.1186/s13059-021-02596-5>.
- Bakulski KM, Halladay A, Hu VW, Mill J, Fallin MD. Epigenetic research in neuropsychiatric disorders: the “tissue issue.” *Curr Behav Neurosci Rep*. 2016;3(3):264–74. <https://doi.org/10.1007/s40473-016-0083-4>.
- Raffington L, Belsky DW, Kothari M, Malanchini M, Tucker-Drob EM, Harden KP. Socioeconomic disadvantage and the pace of biological aging in children. *Pediatrics*. 2021. <https://doi.org/10.1542/peds.2020-024406>.
- Raffington L, Tanksley PT, Sabhlok A, Vinnik L, Mallard T, King LS, Goosby B, Harden KP, Tucker-Drob EM. Socially stratified epigenetic profiles are associated with cognitive functioning in children and adolescents. *Psychol Sci*. 2023;16:170.
- Schmitz LL, Zhao W, Ratliff SM, Goodwin J, Miao J, Lu Q, Guo X, Taylor KD, Ding J, Liu Y, Levine M, Smith JA. The socioeconomic gradient in epigenetic aging clocks: evidence from the multi-ethnic study of atherosclerosis and the health and retirement study [Preprint]. *Epidemiology*. 2021. <https://doi.org/10.1101/2021.03.01.21252660>.
- Braun PR, Han S, Hing B, Nagahama Y, Gaul LN, Heinzman JT, Grossbach AJ, Close L, Dlouhy BJ, Howard MA, Kawasaki H, Potash JB, Shinzaki G. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. *Transl Psychiatry*. 2019;9(1):47. <https://doi.org/10.1038/s41398-019-0376-y>.
- Belsky DW, Caspi A, Corcoran DL, Sugden K, Poulton R, Arseneault L, Baccarelli A, Chamarti K, Gao X, Hannon E, Harrington HL, Houts R, Kothari M, Kwon D, Mill J, Schwartz J, Vokonas P, Wang C, Williams BS, Moffitt TE. DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife*. 2022;11:e73420. <https://doi.org/10.7554/eLife.73420>.
- Finch CE, Crimmins EM. Constant molecular aging rates vs. The exponential acceleration of mortality. *Proc Natl Acad Sci*. 2016;113(5):1121–3. <https://doi.org/10.1073/pnas.1524017113>.
- Raffington L, Tanksley PT, Sabhlok A, Vinnik L, Mallard T, King LS, Goosby B, Harden KP, Tucker-Drob EM. Socially stratified epigenetic profiles are associated with cognitive functioning in children and adolescents [Preprint]. *Genetics*. 2021. <https://doi.org/10.1101/2021.08.19.456979>.
- Koellinger PD, Okbay A, Kweon H, Schweinert A, Karlsson Linnér R, Goebel J, Richter D, Reiber L, Zweck BM, Belsky DW, Biroli P, Mata R, Tucker-Drob EM, Harden KP, Wagner G, Hertzog R. Cohort profile: genetic data in the German Socio-Economic Panel Innovation Sample (Gene-SOEP) [Preprint]. *Genetics*. 2021. <https://doi.org/10.1101/2021.11.06.467573>.
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363–9. <https://doi.org/10.1093/bioinformatics/btu049>.
- Heiss JA, Just AC. Identifying mislabeled and contaminated DNA methylation microarray data: an extended quality control toolset with examples from GEO. *Clin Epigenetics*. 2018;10(1):73. <https://doi.org/10.1186/s13148-018-0504-1>.
- McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites

- on the illumina inffinium MethylationEPIC BeadChip. *Genomics Data*. 2016;9:22–4. <https://doi.org/10.1016/j.gdata.2016.05.012>.
24. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, Van Dijk S, Muhlhäusler B, Stirzaker C, Clark SJ. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol*. 2016;17(1):208. <https://doi.org/10.1186/s13059-016-1066-1>.
 25. Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of illumina inffinium DNA methylation BeadArrays. *Nucleic Acids Res*. 2013;41(7):e90. <https://doi.org/10.1093/nar/gkt090>.
 26. Teschendorff AE, Widschwendter M. Differential variability improves the identification of cancer risk markers in DNA methylation studies profiling precursor cancer lesions. *Bioinformatics*. 2012;28(11):1487–94. <https://doi.org/10.1093/bioinformatics/bts170>.
 27. Houseman EA, Kile ML, Christiani DC, Ince TA, Kelsey KT, Marsit CJ. Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. *BMC Bioinf*. 2016;17(1):259. <https://doi.org/10.1186/s12859-016-1140-4>.
 28. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, Aviv A, Lohman K, Liu Y, Ferrucci L, Horvath S. An epigenetic biomarker of aging for lifespan and healthspan. *Aging*. 2018;10(4):573–91. <https://doi.org/10.18632/aging.101414>.
 29. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Assimes TL, Ferrucci L, Horvath S. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*. 2019;11(2):303–27. <https://doi.org/10.18632/aging.101684>.
 30. Higgins-Chen AT, Thrush KL, Wang Y, Minter CJ, Kuo P-L, Wang M, Niimi P, Sturm G, Lin J, Moore AZ, Bandinelli S, Vinkers CH, Vermetten E, Rutten BPF, Geuze E, Okhuijsen-Pfeifer C, van der Horst MZ, Schreiter S, Gutwinski S, Levine ME. A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. *Nature Aging*. 2022;2(7):644–61. <https://doi.org/10.1038/s43587-022-00248-2>.
 31. Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, Harrington H, Israel S, Levine ME, Schaefer JD, Sugden K, Williams B, Yashin AI, Poulton R, Moffitt TE. Quantification of biological aging in young adults. *Proc Natl Acad Sci*. 2015;112(30):E4104–10. <https://doi.org/10.1073/pnas.1506264112>.
 32. Elliott ML, Caspi A, Houts RM, Ambler A, Broadbent JM, Hancox RJ, Harrington H, Hogan S, Keenan R, Knodt A, Leung JH, Melzer TR, Purdy SC, Ramrakha S, Richmond-Rakerd LS, Righarts A, Sugden K, Thomson WM, Thorne PR, Moffitt TE. 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*. 2021;1(3):295–308. <https://doi.org/10.1038/s43587-021-00044-4>.
 33. Sugden K, Hannon EJ, Arseneault L, Belsky DW, Corcoran DL, Fisher HL, Houts RM, Kandaswamy R, Moffitt TE, Poulton R, Prinz JA, Rasmussen LJH, Williams BS, Wong CCY, Mill J, Caspi A. Patterns of reliability: assessing the reproducibility and integrity of DNA methylation measurement. *Patterns*. 2020. <https://doi.org/10.1016/j.patter.2020.100014>.
 34. McEwen LM, O'Donnell KJ, McGill MG, Edgar RD, Jones MJ, MacIsaac JL, Lin DTS, Ramadori K, Morin A, Gladish N, Garg E, Unternaehrer E, Pokhvisneva I, Karnani N, Kee MZL, Klengel T, Adler NE, Barr RG, Letourneau N, Kobor MS. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells. *Proc Natl Acad Sci*. 2019. <https://doi.org/10.1073/pnas.1820843116>.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

