Comparison of DNA methylation clocks in black South African men

Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Epigenomics

Comparison of DNA methylation clocks in black South African men

Héléne T Cronjé*[,1,2], Cornelie Nienaber-Rousseau[2,3], Josine L Min[4,5], Fiona R Green[6], Hannah R Elliott[4,5] & Marlien Pieters[2,3]

[1]Department of Public Health, Section of Epidemiology, University of Copenhagen, Copenhagen, 2200, Denmark
[2]Centre of Excellence for Nutrition, North-West University, Potchefstroom, 2520, South Africa
[3]Medical Research Council Unit for Hypertension & Cardiovascular Disease, North-West University, Potchefstroom, 2520, South Africa
[4]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 6BN, UK
[5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, BS8 6BN, UK
[6]School of Biosciences & Medicine, Faculty of Health & Medical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK
*Author for correspondence: toinet.cronje@sund.ku.dk

Aims: DNA methylation clocks are widely used to estimate biological age, although limited data are available on non-European ethnicities. This manuscript characterizes the behavior of five DNA methylation clocks in 120 older black South African men. **Methods:** The age estimation accuracy of the Horvath, Hannum and skin and blood clocks and the relative age-related mortality risk and predicted time to death portrayed by the PhenoAge and GrimAge biomarkers are investigated, respectively. **Results:** The results confirm the tendency of DNA methylation clocks to underestimate the biological age of older individuals. GrimAge more accurately characterizes biological decline in this African cohort compared with PhenoAge owing to the unique inclusion of smoking-related damage in the GrimAge estimate. **Conclusions:** Each clock provides a different fraction of information regarding the aging body. It is essential to continue studying under-represented population groups to ensure methylation-derived indicators are robust and useful in all populations.

First draft submitted: 24 August 2020; Accepted for publication: 29 January 2021; Published online: 8 March 2021

Keywords: biological age • GrimAge • PhenoAge • phenotypic age • smoking

As global life expectancy continues to increase, the chronic disease burden expands and the need for a better understanding of how to promote healthy aging is emphasized [1]. Chronological age is an integral component of frailty, noncommunicable disease risk and mortality [2]. Although easily accessible and standardized, chronological age as a biomarker is limited by its inability to portray changes in biological functionality accurately over the life span, especially in later life [3]. A group of peers, for example, may be the same chronological age, yet exhibit a spectrum of age-related deterioration [3]. For this reason, extensive efforts have been made to develop markers that are able to reflect biological aging better than years since birth do [3,4]. Ultimately, the availability of such markers could allow for improved targeted intervention through the identification of high-risk, functionally declining individuals before clinical symptoms appear [4].

DNA methylation (DNAm) refers to attachment of a methyl group to a DNA base. DNAm changes accumulate with age [5,6] and are thought to mediate the effects of environmental risk factors on disease [7]. DNAm levels at specific CpG sites (referred to as clock CpGs) can also be used to predict chronological age [4,8,9]. These predictors are termed 'epigenetic clocks' and quantify (in years) a biological age estimate [4]. Residuals from the regression of epigenetic age on chronological age are defined as DNAm age acceleration (AA, DNAmAA). Positive DNAmAA (biological age is projected to be older than chronological age) has been associated with diabetes [10], cancer [11], cardiovascular disease [12] and all-cause mortality [11,13], although causality still has to be established [4,14].

To date, multiple DNAmAge clocks have been developed, with some variation in composition and outcome (Table 1). The Horvath [8], Hannum [15] and skin and blood (SB) [16] clocks are widely used because of their ability to robustly predict either the chronological age of unknown donors or biological age discrepancies in various tissues

Future Medicine

Table 1. Descriptive characteristics of widely used epigenetic clocks.

Name	Estimates	DNAm data description†	Clock computation	Ref.
Horvath	DNAmAge reflects CA in the majority of healthy tissues Where DNAmAge and CA differ, DNAmAA is calculated as a measure of faster/slower relative biological aging	27/450K probe analysis of 51 tissues and cell types	ENR of CA on DNAm returned 353 CA-associated CpGs	[8]
Hannum		450K probe analysis of whole blood	ENR of CA on DNAm returned 71 CA-associated CpGs	[15]
Skin and blood		450K/EPIC probe analysis of peripheral tissues used for ex vivo studies	ENR of CA on DNAm returned 391 CA-associated CpGs	[16]
PhenoAge	Biological age based on physiological dysregulation that relates to mortality risk	450K/EPIC probe analysis of whole blood	ENR of PA on DNAm returned 513 PA-associated CpGs	[17]
GrimAge	Biological age calculated as an estimation of time to death	450K/EPIC probe analysis of whole blood	ENR of seven protein markers and smoking pack-years on DNAm delivered eight DNAm surrogates that amounted to a collective biomarker of 1030 CpGs	[18]

PA = weighted average of albumin, alkaline phosphatase, creatinine, C-reactive protein, serum glucose, mean cell volume, lymphocyte percentage, red cell distribution width, white blood cell count and chronological age.
†Naming of two platforms denotes the use of only overlapping probes; 27/450K and 450K/EPIC refer to various Illumina Infinium MethylationEPIC BeadChip kits.
CA: Chronological age; DNAm: DNA methylation; DNAmAge: DNA methylation age; DNAmAA: DNA methylation age acceleration; ENR: Elastic net regression; PA: Phenotypic age.

in a single individual. Because these age predictor clocks were developed with chronological age as the sole outcome of interest, they often fail to capture the inter-individual methylation differences that discern biological decline above that of advancing age itself [3,4]. The more recent development of mortality predictor clocks PhenoAge [17] and GrimAge [18] addresses this limitation. Instead of relying solely on chronological age, these models incorporate a composite outcome of aging-related clinical measurements that differentiate between healthy and unhealthy aging (Table 1). The mortality predictor clocks are particularly important in the context of health research because they have been developed using longitudinal data and well-defined mortality outcomes and are, therefore, able to generate useful biological age biomarkers. This gives cohorts without such data, such as the Prospective Urban and Rural Epidemiology (PURE) study in the North West province of South Africa (PURE-SA-NW), an opportunity to use methylation-derived biomarkers supported by causal models to indicate the health risks of study populations.

Although these clocks are widely used and studied, a number of limitations in their use have been identified. First, a systematic underestimation of Horvath and Hannum DNAmAge in older adults has been reported [19], necessitating further investigation into the use of these clocks in older populations. Second, although ethnic differences in the behavior of epigenetic clocks have been reported [17,20,21], most of the current literature represents data obtained from individuals of European ancestry only, with very limited information available on other ethnic groups. In addition to genetic confounding, environmental differences between populations also influence the physiological features these models are built on, particularly when lifestyle factors are incorporated, such as in the case of the GrimAge clock, where the damaging effects of smoking are captured. For this reason, replication in various cohorts should continually be pursued and clocks refined when previously unmeasured confounding becomes known.

This manuscript therefore compares the behavior of five epigenetic clocks [8] in a group of black South African men between the ages of 45 and 88. The age estimation accuracy of the Horvath, Hannum and SB clocks (and the potential issue of underestimation in older adults) and the relative age-related mortality risk and predicted time to death portrayed by the PhenoAge and GrimAge estimators are investigated, respectively, in a non-European population with a particularly high prevalence of smoking. Because many datasets of existing cohorts possess DNAm data, mostly generated using the Illumina platforms, understanding the multitude of uses of such DNAm data is essential. For this reason, analyses are restricted to the most frequently used DNAm clocks in the epigenetic literature [14,18,20,22,23] developed using the Illumina platform and available via the online methylation age calculator (https://dnamage.genetics.ucla.edu/new) [8,16].

Methods

Study population

The international PURE study comprises cohorts from 27 countries tracking participants over a period of 20 years. This article reports cross-sectional data from the 2015 arm of the PURE-SA-NW study. A subset of ostensibly healthy black South African men (n = 120) with available peripheral blood samples were randomly selected for this study, provided they tested negative for HIV at the time of data collection. Eligibility was restricted to reduce

Table 2. Descriptive characteristics of age, DNAmAge and DNAmAA in the study sample.

	Mean ± SD	Correlation with chronological age	
		r	p-value
Chronological age (years)	63 ± 10		
Age predictor clocks (years)			
HorvathAge	59 ± 8	0.58	2.3E–12
HannumAge	47 ± 8	0.64	2.2E–15
SBAge	54 ± 8	0.70	7.4E–19
IEAA	0 ± 6.4		
EEAA	0 ± 7.6		
SBAA	0 ± 5.5		
Mortality predictor clocks (years)			
PhenoAge	47 ± 9	0.51	2.7E–09
GrimAge	64 ± 9	0.80	1.9E–28
PhenoAA	0 ± 7.5		
GrimAA	0 ± 5.3		

DNAmAge: DNA methylation age; DNAmAA: DNA methylation age acceleration; EEAA: Extrinsic epigenetic age acceleration; IEAA: Intrinsic epigenetic age acceleration; SBAge: Skin and blood age; SBAA: Skin and blood age acceleration; SD: Standard deviation.

confounding by sex and CD4 T-cell count. Additional information on the international cohort [24] and this substudy [25] has been published.

Data collection

Participants reported their current smoking and alcohol use via interview. If applicable, they reported the frequency and quantity of intake, age at the start of use and previous attempts at abstinence. The Cobas Integra 400 (Roche Diagnostics, IN, USA) was used to quantify serum albumin, creatinine and high-sensitivity CRP. Peripheral blood collected in fluoride tubes was used to measure glucose concentrations using the same device. Serum alkaline phosphatase was quantified using the Cobas Integra 400 plus (Roche Diagnostics, Basel, Switzerland).

Peripheral blood samples were used for DNA extraction. Genome-wide methylation data were generated by Bristol Bioresource Laboratories (Bristol Medical School, University of Bristol, Bristol, UK) using the standard protocol of the Infinium MethylationEPIC BeadChip platform (Illumina, CA, USA). Quality control, sample filtering and functional normalization were done using the meffil [26] package in R 3.4.3 [27]. A detailed description of the DNA extraction, quality control, methylation quantification and data processing protocols has been published [25]. Lymphocyte proportions were methylation-derived [25,28].

Cell counts & DNAmAge

The IDOL-optimized L-DMR library for whole blood samples [28] was used to estimate the distribution of B, CD4+ T, CD8+ T, neutrophil, monocyte and natural killer cells. For the estimation of DNAmAge and DNAmAA, a widely used online calculator (https://dnamage.genetics.ucla.edu/new) was used [8]. The following age and AA estimates were used as provided in the calculator output: DNAmAge, DNAmAgeHannum, extrinsic epigenetic AA (EEAA), DNAmAgeSkinBloodClock, DNAmPhenoAge, DNAmGrimAge, BloodskinAA, PhenoAA and GrimAA. Intrinsic epigenetic age acceleration (IEAA) was calculated using the residuals from a linear model, where DNAmAge was used as the outcome and chronological age, the calculator's plasmaBlast, CD8pCD28nCD45Ran and CD8.naive and IDOL-estimated CD4+ T cells, natural killer cells, monocytes and neutrophils as predictors. As opposed to the EEAA estimate, which includes leukocyte cell distribution in its algorithm, IEAA represents a cell count-adjusted AA estimate. It should be noted that the Illumina Infinium MethylationEPIC array used to quantify methylation data for the current study excludes 19 of the 353 Horvath and six of the 71 Hannum clock CpGs [8,15]. The absence of these CpGs has previously been reported not to compromise the accuracy of these clocks [29]. All further cell count adjustments were performed using cell estimates from the IDOL package.

Statistical analysis

Data normality was evaluated using the Shapiro–Wilk test. The mean and standard deviation of chronological age, DNAmAge and DNAmAA estimates in this study population are reported alongside Pearson's correlation

Table 3. Comparison of seven clinical components of phenotypic age between current and never smokers and each component's association with PhenoAge and chronological age.

Clinical components of phenotypic age	Geometric mean ± SD			PhenoAge correlation		PhenoAA correlation	
	Never smoker (n = 56)	Current smoker (n = 61)	*t*-test p-value	*r*	p-value	*r*	p-value
Age, years	63 ± 10	63 ± 10	0.78	0.51	2.7E–09*	0.00	0.96
Albumin, g/l	42.1 ± 1.19	44.2 ± 1.17	0.10	-0.15	0.10	-0.11	0.23
ALP, U/l	83.9 ± 1.45	85.1 ± 1.39	0.82	0.04	0.63	0.04	0.69
Creatinine, μmol/l	0.08 ± 0.03	0.07 ± 0.02	0.04*	0.10	0.30	-0.06	0.54
CRP, mg/dl	3.62 ± 4.23	3.00 ± 3.37	0.45	0.39	1.1E–05*	0.36*	9.0E–05*
Glucose, mmol/l	5.21 ± 1.20	4.99 ± 1.16	0.19	-0.03	0.72	-0.07	0.45
Lymphocyte, %	37.2 ± 9.30	36.6 ± 10.5	0.75	-0.31	4.8E–04*	-0.33*	2.9E–04*

All components, apart from age and lymphocyte %, were log-transformed. Lymphocyte estimates are methylation-derived. Phenotypic age components for which no data are available are mean cell volume, white blood cell count and red cell distribution width.
*$p < 0.05$.
ALP: Alkaline phosphatase; SD: Standard deviation.

Table 4. Adjusted group means of aging-related phenotypes for current versus never smokers.

Outcome	Model	Smoking status Group mean ± SE		
		Never (n = 56)	Current (n = 61)	p-value
GrimAge	1	62.2 ± 0.68	66.6 ± 0.65	2.0E–05
	2	62.0 ± 0.79	66.7 ± 0.76	3.5E–04
GrimAA	1	-2.27 ± 0.68	2.19 ± 0.65	2.0E–05
	2	-2.4 ± 0.79	2.31 ± 0.76	3.5E–04
DNAmPackY	1	19.9 ± 1.6	35.1 ± 1.6	4.2E–09
	2	19.5 ± 1.9	35.5 ± 1.8	6.5E–07
DNAmLeptin	1	6448 ± 541	8514 ± 520	1.1E–02
	2	6450 ± 634	8512 ± 604	4.5E–02
DNAmPAI-1	1	16657 ± 503	18757 ± 483	5.5E–03
	2	16956 ± 586	18479 ± 559	1.1E–01

Variables prefaced by DNAm are the methylation-derived surrogates of the following components used for the GrimAge estimate: model 1 = outcome ~ smoking status + chronological age + body mass index + education + white blood cell count; model 2 = outcome ~ smoking status + chronological age + body mass index + education + white blood cell count + alcohol consumption status.
DNAmPackY: DNA methylation pack-years; GrimAA: Grim age acceleration; SE: Standard error.

coefficients between chronological age and each of the DNAmAge estimates. Figures 1 and 2 were compiled with the ggplot2 and BlandAltmanLeh packages. In Table 3, chronological age, lymphocyte fraction and logarithmically transformed akaline phosphatase, serum albumin, creatinine, high-sensitivity CRP and glucose concentrations are correlated with PhenoAge and PhenoAA. The means of these clinical components of phenotypic age are also compared between current and never smokers using *t*-tests.

Type III analysis of variance models in the car package was applied to linear regression objects to quantify the differences in outcome means between current smokers and never smokers (Table 4). Adjusted group means and standard errors were extracted using the effects package. Two models were run for each outcome: a model adjusting for chronological age, BMI, education and white blood cell (WBC) count and a model adjusting for alcohol consumption with the former variables. Covariates were chosen based on the current literature to ease comparability [30]. R version 3.5.0 was used for all analyses [27].

Results

This study evaluated the behavior of five epigenetic clocks in 120 black men who participated in the PURE-SA-NW study. Chronological ages ranged from 45–88 years. Approximately half of the 120 participants were current smokers and alcohol consumers (61 and 58, respectively), and 48 participants (40% of the cohort) were using both substances at the time of data collection. The majority of the study population were of normal weight, with only

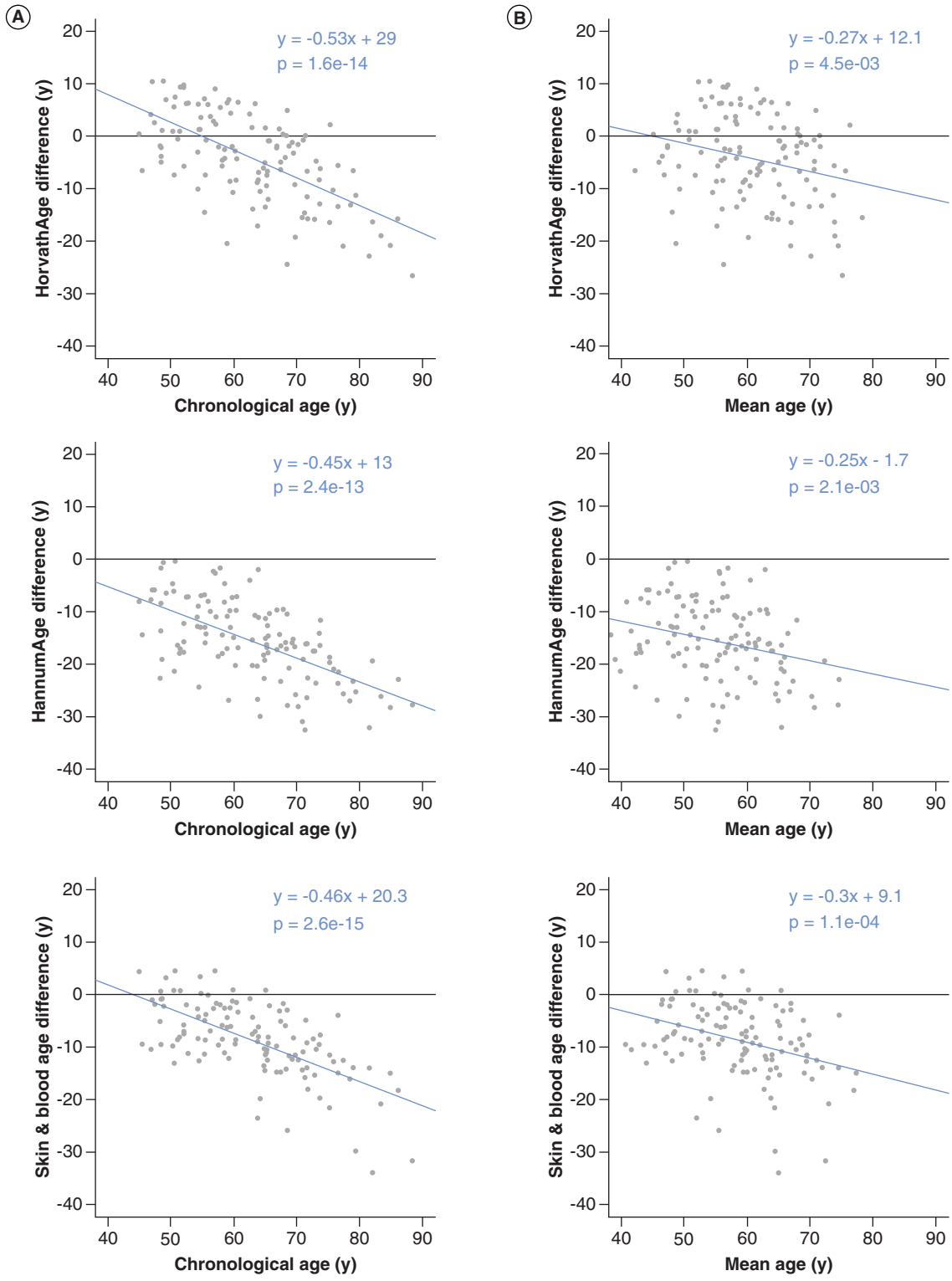

Figure 1. Scatter plots illustrating the relative difference in biological versus chronological age by three age predictor clock DNAmAge estimates. (A) Scatter plots depicting the relationship between chronological age and age difference when subtracting chronological age from DNAmAge estimates. **(B)** Bland–Altman plots depicting the mean of each DNAmAge and chronological age against the difference between each DNAmAge and chronological age. The line of best fit from a linear regression model is formulated as $y = mx + c$ in the top right corner and depicted in blue on the plot. The p-value represents the statistical significance of the linear regression model.
DNAmAge: DNA methylation age.

10.2217/epi-2020-0333

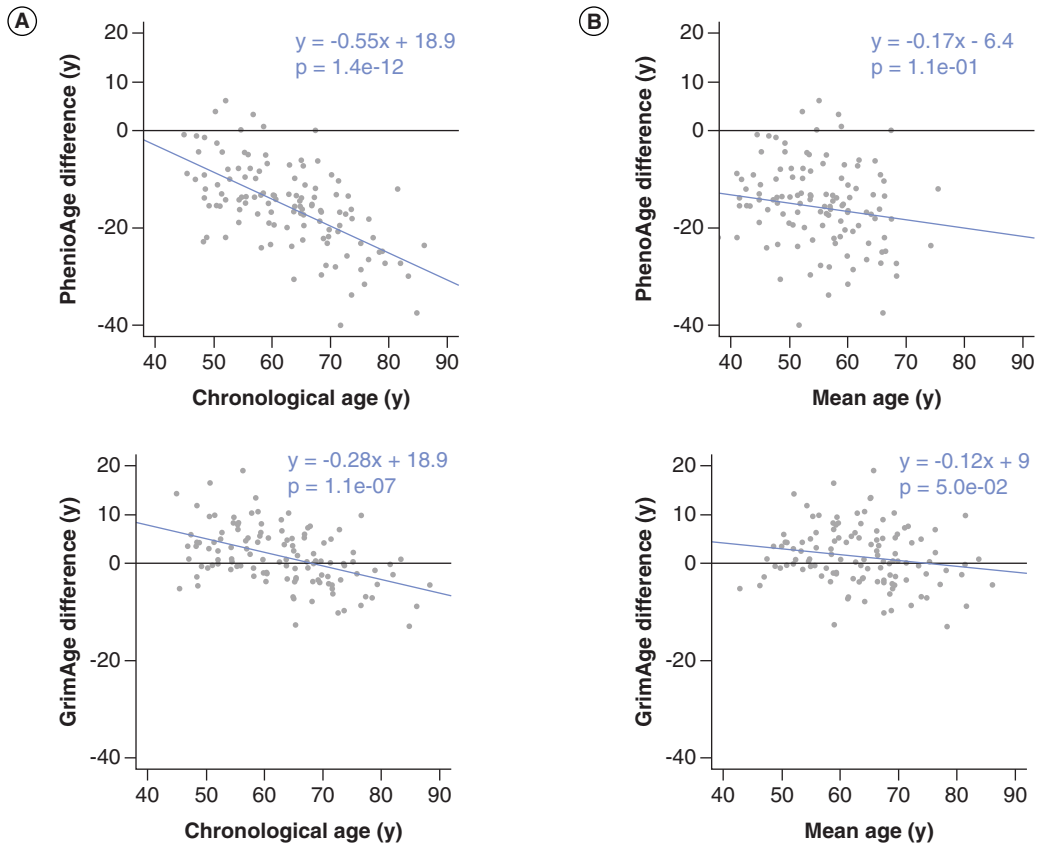

Figure 2. Scatter plots illustrating the relative difference in biological versus chronological age by two mortality predictor clock DNAmAge estimates. (A) Scatter plots depicting the relationship between chronological age and age difference when subtracting chronological age from DNAmAge estimates. **(B)** Bland–Altman plots depicting the mean of each DNAmAge and chronological age against the difference between each DNAmAge and chronological age. The line of best fit from a linear regression model is formulated as y = mx + c in the top right corner and depicted in blue on the plot. The p-value represents the statistical significance of the linear regression model.
DNAmAge: DNA methylation age.

21% classified as overweight (BMI ≥25–29.9 kg/m^2) and 8% as obese (BMI ≥30 kg/m^2). Fifty-five percent of the study population had undergone 1–7 years of schooling, whereas 23% had undergone 8–12 years. The remainder of the population had received no formal education.

Comparison of biological age & AA estimates with chronological age

Table 2 reports the mean and standard deviation of three age predictor and two mortality predictor DNAmAge and DNAmAA estimates. The association of each DNAmAge estimate with chronological age is also reported. Figures 1 and 2 depict scatter plots comparing the age estimations by each of the age predictor and mortality predictor clocks. Chronological age is plotted against the difference between each clock's estimated and chronological age, and Bland–Altman plots depict the relationship between the mean of each estimated and chronological age against the difference between each estimated and chronological age. The age and mortality predictor clocks are discussed separately in the following sections because of the integral differences in their aims and outcomes.

Age predictor clocks

SBAge presented the highest correlation with chronological age, although HorvathAge gave the closest estimate of mean age (Table 2). HorvathAge correlated with HannumAge and SBAge, with a correlation coefficient of 0.62 (p = 7.7E–14) and 0.64 (p = 2.1E–5), respectively. HannumAge and SBAge associated with each other more strongly (r = 0.71; p = 6.2E–20). In terms of AA measures, Horvath's IEAA correlated with Hannum's EEAA, with

a correlation coefficient of 0.31 (p = 6.0E–4), whereas equally strong correlations of 0.41 were observed between SBAA and the IEAA (p = 3.8E–06) and EEAA (p = 2.8E–06) estimates, respectively.

All three clocks tended to underestimate chronological age in a systematic fashion, with this deviation increasing with age (Figure 1). The trend of underestimation was stronger for HorvathAge than HannumAge and SBAge, although HannumAge was consistently lower than the other two. The Horvath clock was most accurate at a chronological age of 55 years (the age at which the regression line intercepted the y-axis) (Figure 1). The Hannum and SB clocks seemed to be most accurate for individuals younger than those represented in the PURE-SA-NW cohort (at chronological ages of 29 and 44 years, respectively). Additional adjustment for WBC count variation did not significantly alter the degree of underestimation.

Mortality predictor clocks

A 17 year difference was observed between the biological age estimation of the PhenoAge and GrimAge clocks (47 vs 64 years) (Table 2). Biological age and AA estimates from the PhenoAge and GrimAge clocks correlated with a coefficient of only 0.51 (p = 4.2E–09) and 0.21 (p = 0.02), respectively. Similar to the Hannum and SB clock estimates, the PhenoAge clock likely performs best at ages outside the reported cohort's age range, as shown in the simulation on PURE-SA-NW data where the y-axis intercept is 34.3 years (Figure 2). For the GrimAge clock, however, optimal estimation is at age 67.4 years, which is the highest optimal age of all tested clocks and is also closest to the mean age of the PURE-SA-NW cohort. Additional adjustment for WBC count variation did not significantly alter the degree of underestimation.

Role of smoking in the mortality predictor clocks

Because more than half of the PURE-SA-NW study population were current smokers, it was hypothesized that the difference in the contribution of smoking to the PhenoAge and GrimAge estimates might explain the large discrepancy between the performances of these two clocks. PhenoAge of 47.4 ± 9 and 46.8 ± 9 years, respectively, was observed when never and current smokers were compared (p = 0.71). No differences between PhenoAA of smokers (-0.35 ± 7.96) and nonsmokers (0.49 ± 7.23) were observed (p = 0.55) either. By contrast, GrimAge of nonsmokers was on average 5 years younger than that of smokers (61.7 ± 9 vs 66.7 ± 9; p = 0.003), whereas a 4.5 year difference was observed in AA comparisons (-2.39 ± 4.54 vs 2.20 ± 5.12; p = 1.2E–06).

To explore this further, the association of smoking, PhenoAge and PhenoAA with seven of the ten clinical components (protein markers) of phenotypic age that were available for the PURE-SA-NW study cohort are provided (Table 3). The association of smoking with each of the eight methylation surrogate markers that comprise the GrimAge estimate are also reported (Table 4).

Apart from albumin, none of the available clinical components of phenotypic age differed significantly between smokers and nonsmokers. PhenoAge correlated with chronological age, CRP and lymphocyte percentage, whereas PhenoAA correlated with only the latter two (Table 3).

Apart from the DNAm pack-years component, the DNAmPAI-1 and DNAmLeptin also differed between current and never smokers (Table 4). Because alcohol consumption often coincides with smoking habits, a sensitivity analysis was performed, additionally adjusting for alcohol use (model 2). This did not significantly alter the associations observed with smoking status, apart from an attenuating effect on DNAmPAI-1.

As a second sensitivity test, potential bias in the self-reported smoking status data was investigated by analyzing the association between the continuous methylation percentage of cg05575921 and each of the PhenoAge and GrimAge constituents. Previous literature has suggested that cg05575921 is a less biased estimate of smoking behavior than self-report [31]. Again, only GrimAge and GrimAA, not PhenoAge or PhenoAA, associated with cg05575921. Similar to self-reported smoking comparisons, the only PhenoAge component associated with cg05575921 was creatinine. For the GrimAge components, associations with the DNAm pack-years, DNAmPAI-1 and DNAmB2M were observed. Table 4 includes only results at $p < 0.05$. For the full table, see Supplementary Table 1.

Discussion

This study compared five epigenetic clocks previously proven to be highly effective in reaching their respective aims of chronological age prediction (age predictor clocks) and prediction of mortality-related functional decline (mortality predictor clocks). Previous findings of discrepancies between the predicted biological ages generated by the different clocks were confirmed as well as the tendency of all tested clocks to underestimate the biological age

of older individuals. GrimAge provided the most accurate reflection of the biological aging of a study sample of continental African men with a high prevalence of smoking and alcohol use.

Comparison of biological age & AA estimates with chronological age

Age predictor clocks

The average chronological ages predicted by the three age predictor clocks spanned 12 years (47–59 years) and were all younger than the average calendar chronological ages reported by the PURE-SA-NW participants (63 ± 10 years). The ages predicted by these clocks are known to differ within the same dataset, in part because of the limited overlap in represented CpGs and the likelihood that each clock (and the clock CpGs it contains) captures varying degrees of cell count, environmental and ethnic influences or confounding [4,23,32]. The SB clock shares 45 loci with the Hannum clock and 60 loci with the Horvath clock [16], whereas the latter two share only five loci [8,15]. This larger overlap and the similarity of the target tissue on which these clocks were trained (whole blood) are reflected in the stronger correlation between the PURE-SA-NW estimated Hannum and SB age estimates than their respective correlations with HorvathAge.

Correlations between the ages estimated by the age predictor clocks and chronological ages in the PURE-SA-NW cohort were comparatively weaker than those previously reported. In the SBAge validation analysis conducted by Horvath *et al.*, chronological and predicted age correlated at a coefficient of 0.96 for HorvathAge, 0.97 for HannumAge and 0.98 for SBAge [16]. Weaker correlations observed in the current data compared with Horvath's validation data are likely due to the wider range of ages represented in the validation than the current study (19–82 compared with 45–88 years).

In agreement with SBAge having the strongest correlation with chronological age, SBAA also displayed the least amount of AA variance, followed by the IEAA and EEAA measures. These differences in AA variation probably occurred because of the varying role of WBC count within the different DNAmAge algorithms from which each AA was derived. Based on the fact that WBC composition changes with age, EEAA incorporates WBC changes in a weighted manner by aggregating the HannumAge estimate with plasmablasts, naive cytotoxic T cells and exhausted cytotoxic T cells [15,33]. This can, however, introduce inter-individual variation because WBC composition is influenced by factors apart from aging itself, such as sex, medication use, disease and ethnicity and genetic factors [33–37]. For IEAA, however, inter-individual variation is reduced by adjusting for cell counts. The adjustment is specifically made to optimize the performance of the multitissue predictor (HorvathAge) when it is applied to blood samples [8,33]. SBAA consequently outperforms both of these estimates, as it was initially developed for blood samples and, therefore, needs no additional WBC count adjustments, nor does it introduce potential additional variance by incorporating differential WBC composition [16].

The SB clock was developed to improve the age estimation accuracy of the Horvath clock (by limiting training data to a smaller number of cell types) and the Hannum clock (by using larger training datasets and increasing represented loci). In PURE-SA-NW, the superiority of the SB clock was confirmed by its higher correlation with chronological age than the other clocks and its improvement in AA variation in samples derived from whole blood.

Mortality predictor clocks

The PhenoAge clock estimated the PURE-SA-NW cohort at a much lower age-related mortality risk than suggested by both chronological age and time to death estimate reflected by GrimAge. Although the GrimAge and PhenoAge estimates cannot be directly compared in terms of age estimation and were not designed to reflect chronological age, the ideal would be that these clocks reveal a comparable and complementary estimation of biological decline.

The role of chronological age in the development of the PhenoAge and GrimAge clocks is critical in untangling their behavior in this investigation. For the PhenoAge algorithm, chronological age was incorporated as one of ten clinical markers associated with the risk of mortality, which were aggregated to represent phenotypic age. Chronological age is, therefore, used in a manner similar to that seen with the age predictor clocks, the difference being its aggregation with other clinical markers, rather than being the single outcome measure.

The GrimAge clock, however, incorporated chronological age as an adjustment variable in the CpG selection models. Age was selected together with sex, pack-years and seven protein markers to best predict time to death. A penalized regression, adjusted for sex and chronological age, was first applied to select CpGs associated with each of the GrimAge proteins and pack-years. After all the selected CpGs had been combined to form the GrimAge marker, linear transformation based on forcing the GrimAge mean and variance to match chronological age was applied.

Based on these protocols, a weaker correlation between chronological age and PhenoAge than HorvathAge, HannumAge and SBAge was expected because of the diluted contribution of chronological age in the prediction models. However, even compared with external data, the PhenoAge clock performed particularly poorly in the PURE-SA-NW cohort [17,32]. For example, the correlations between PhenoAge and chronological age reported for four independent validation cohorts, investigated by Levine *et al.* in the PhenoAge development article, were much stronger than what was observed for the PURE-SA-NW participants (r = 0.51 in PURE-SA-NW vs 0.66, 0.69, 0.78 and 0.89 in the respective validation cohorts) [17]. For GrimAge, however, the relationship observed in the current study with chronological age (r = 0.81) was strikingly similar to that reported for the same validation cohorts mentioned earlier (reported here in the same order: r = 0.79, 0.80, 0.82 and 0.89) [18]. The reduced accuracy of the PhenoAge compared with GrimAge chronological age correlation in the PURE-SA-NW participants compared with other cohorts potentially reflects environmental or ethnic confounding in the PURE-SA-NW group that is not captured by the design of the PhenoAge clock but is captured by the GrimAge clock. In agreement with this hypothesis of ethnic variability, Levine *et al.* reported notable differences in PhenoAge between Hispanic, non-Hispanic black and non-Hispanic white samples (p = 5.1E–05) [17]. When investigating the standard deviations of PhenoAA compared with GrimAA, Lu *et al.* found an overall larger variance for PhenoAA as well as larger interethnic differences, with the largest variation reported for African American cohorts, followed by white and Hispanic groups [18]. The observation in the current study of larger variation in PhenoAA than GrimAA is in agreement with that of both Lu *et al.* and a more recent comparison by Zhao *et al.* in an independent African American study cohort [32]. Zhao *et al.* reported PhenoAge estimates that were on average 13 years younger than chronological age (44 vs 57 years) compared with the much more congruent behavior observed in the GrimAge estimate (54 years), which is analogous to the current investigation's findings. Because of the limited literature, it remains difficult to confirm whether the comparatively weaker performance of the PhenoAge clock is the result of population-specific ethnic or environmental confounding or a general limitation of the PhenoAge clock.

Population-specific confounding could result from either different associations between the clinical markers used in the design of the clock and mortality risk in the study population investigated or differing CpG phenotypic age associations. Although the role of ethnic confounding could not be investigated in the PURE-SA-NW cohort, the premise of genetic confounding related to PhenoAge was recently strengthened by evidence revealing that at least 50 of the clock CpGs comprising the PhenoAge algorithm capture ancestry-specific data [22]. Environmental confounding, however, was one hypothesis that could be tested in the PURE-SA-NW study. It is possible that the PhenoAge clock may not have accurately captured the age acceleration associated with smoking behavior in the PURE-SA-NW cohort, resulting in severe biological age underestimation. Although the PhenoAge algorithm was not developed to encapsulate smoking habits specifically, previous research, including the PhenoAge validation analysis [17], reported that this age estimate was able to discriminate between smokers and nonsmokers [17,32]. In the PURE-SA-NW cohort, however, neither PhenoAge nor the clinical components of phenotypic age encapsulated the biological effects of smoking. The fact that the PURE-SA-NW study sample included only men should not have affected this association [14,32].

By contrast, the integral role of smoking in the GrimAge clock was confirmed by the associations observed between smoking status and multiple methylation components of the GrimAge clock in the PURE-SA-NW cohort. These findings reflect those observed by Zhao *et al.* in which smoking status associated not only with DNAm pack-years but also with other methylation components, such as adrenomedullin, B2M, GDF-15, cystatin C, PAI-1 and tissue inhibitor metalloproteinase [32]. The PURE-SA-NW data furthermore demonstrated that it is primarily smoking rather than combined smoking and alcohol use that is captured by GrimAge. The only methylation component that additional adjustment for alcohol consumption affected was DNAmPAI-1. This is in line with the finding of Zhao *et al.* that alcohol consumption has the strongest influence on DNAmPAI-1.

Strengths & limitations

Testing the DNAm clocks in an older population in apparently good health allowed for the largely disease-independent underestimation of biological age in older individuals to be investigated. Literature in this area is currently lacking. In addition, this manuscript contributes to the limited ethnic diversity represented in the epigenetic aging literature by reporting on a group of continental Africans. This investigation was, however, limited by the sample size, and the relatively narrow age range of the study sample. These limitations prohibits the extrapolation of reported findings to women or younger individuals or those with specific disease diagnoses. Furthermore, data on three of the ten clinical phenotypic age components were not avaialable for investigation,

and may have presented individual differences in smoking status groups. The fact that smoking-related differences were not observed in the aggregate PhenoAge marker does, however, suggest that this may be unlikely. Finally, the lack of longitudinal and mortality data limited the ability to test the PhenoAge and GrimAge clocks' accuracy in predicting age-related mortality risk and time to death, respectively. In a recent Swedish longitudinal study, Li *et al.* investigated a range of biological aging measures as predictors of mortality [38]. Methylation measures included were Horvath, Hannum, PhenoAge and GrimAge. Alongside frailty index, Horvath and GrimAge were independently predictive of mortality risk, confirming their value as measures of healthy aging and indicating that they may capture different components of health. Use of methylation-based measures combined with other measures of biological age (e.g., frailty index, as used by Li *et al.*) may, therefore, offer more optimized tools to identify individuals within populations who may benefit most from health interventions.

Conclusion

Epigenetic clocks provide unique possibilities for a better understanding of healthy versus accelerated aging and consequently could ultimately improve the quality of life of a global population with an ever-increasing life span. In the PURE-SA-NW cohort, although all the tested clocks underestimated the biological age of older individuals, the GrimAge clock was particularly useful, as it was the only clock that incorporated the damaging effects of smoking in biological aging. Our data illustrate that each clock provides a different fraction of information regarding the aging body. Although the differences in the physiological features these models capture strengthen the usefulness of specific clocks for particular investigations, they also make their accuracy sensitive to heterogeneity between communities. For this reason, replication in various cohorts is important to reduce the weight of these influences.

Future perspective

Research on the characterization of epigenetic age and aging is advancing at a rapid pace. Future research should focus on increased ethnic and environmental diversity in the data that underlie the algorithms designed for prediction to better predict morbidity and mortality in multiple population groups and the inclusion of methylation-based aging biomarkers (e.g., Brenner age and Pace of Aging measures) that provide estimates according to scale- rather than time-based risk. Because variation in lifestyle and ethnicity may influence the physiological features methylation clocks are built on, increased diversity will enable the developers of aging algorithms to improve their accuracy and ultimately reduce the weight of these influences on generally used clocks. Apart from optimizing the precision of methylation-based biomarkers, increased efforts should be made to integrate methylation-based and methylation-independent measures of biological age (e.g., frailty index) to broaden our understanding of the 'aging' phenotype and improve population-based risk stratification.

Summary points

- Accelerated biological versus chronological aging is associated with a variety of chronic diseases and mortality.
- DNA methylation clocks are widely used to estimate biological age, although only limited data are available on non-European ethnicities.
- This study characterizes three age predictor (Horvath, Hannum and skin and blood) and two mortality predictor (PhenoAge and GrimAge) DNA methylation clocks in black South African men aged 45–88 years.
- Of the age predictor clocks, skin and blood age had the strongest correlation with chronological age and the least variation in DNA methylation age acceleration.
- All tested clocks tended to underestimate the biological age of older individuals.
- The Horvath, Hannum and skin and blood clocks best estimated chronological age at 55, 29 and 44 years, respectively.
- The PhenoAge clock estimated a much lower age-related mortality risk than suggested by both chronological age and the time to death reflected by GrimAge.
- GrimAge provided superior characterization of age-related biological decline compared with PhenoAge.
- The superior performance of GrimAge versus PhenoAge in this population is likely due to its incorporation of smoking-related biological decline, as more than half of the study sample were current smokers.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/epi-2020-0333

Author contributions

H T Cronjé, M Pieters, C Nienaber-Rousseau and H R Elliott conceptualized the study. Funding was acquired by M Pieters and F R Green. H T Cronjé performed the data analysis and wrote the manuscript. M Pieters supervised the data analysis and interpreted the results with J L Min, H R Elliott and H T Cronjé. All authors contributed to the critical review and editing of the manuscript.

Acknowledgments

The authors thank all those who participated in the PURE-SA-NW study and those who made the PURE-SA-NW study possible, including the fieldworkers, researchers and staff of both PURE-SA-NW (Africa Unit for Transdisciplinary Health Research, Faculty of Health Sciences, North-West University, Potchefstroom, South Africa) and PURE International (S Yusuf and the PURE project office staff at the Population Health Research Institute, Hamilton Health Sciences and McMaster University, Ontario, Canada) teams. The authors also thank the staff of Bristol Bioresource Laboratories (Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK), who generated the DNA methylation data.

Financial & competing interests disclosure

The PURE-SA-NW study was funded by North-West University, South African National Research Foundation, Population Health Research Institute, South African Medical Research Council, North West Province Health Department and South African Netherlands Partnerships in Development. Grants from the South African National Research Foundation, Academy of Medical Sciences UK (Newton Fund Advanced Fellowship Grant [AMS-NAF1-Pieters to M Pieters and F R Green]) and South African Medical Research Council funded the DNA methylation analysis. H T Cronjé was supported by the South African National Research Foundation (SFH106264, MND121094) and is currently supported by the Novo Nordisk Fonden Challenge Programme: Harnessing the Power of Big Data to Address the Societal Challenge of Aging (NNF17OC0027812). H R Elliott works in the Medical Research Council Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council and the University of Bristol (MC_UU_00011/5). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Ethical approval for the 2015 data collection of the PURE-SA-NW study was granted by the Health Research Ethics Committee of North-West University (NWU-00016-10-A1, NWU-00119-17-A1). All participants provided written informed consent, including consent for genetic/epigenetic analysis.

Data sharing statement

The data that support the findings of this study are available upon reasonable request and with the permission of the Health Research Ethics Committee of North-West University and the principal investigator of the PURE-SA-NW study, I M Kruger (lanthe.kruger@nwu.ac.za), at North-West University's Africa Unit for Transdisciplinary Health Research.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Wang H, Naghavi M, Allen C *et al.* Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* 388(10053), 1459–1544 (2016).
2. Beard HPJR, Bloom DE. Towards a comprehensive public health response to population ageing. *Lancet* 385(9968), 658–661 (2015).
3. Jylhava J, Pedersen NL, Hagg S. Biological age predictors. *EBioMedicine* 21, 29–36 (2017).
4. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* 19(6), 371–384 (2018).
5. Martin GM. Epigenetic drift in aging identical twins. *Proc. Natl Acad. Sci. U. S. A.* 102(30), 10413–10414 (2005).
6. Wang Y, Karlsson R, Lampa E *et al.* Epigenetic influences on aging: a longitudinal genome-wide methylation study in old Swedish twins. *Epigenetics* 13(9), 975–987 (2018).

7. Teschendorff AE, Relton CL. Statistical and integrative system-level analysis of DNA methylation data. *Nat. Rev. Genet.* 19(3), 129–147 (2018).
8. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 14(10), (2013).

•• **Original Horvath clock conceptualization, development and application article.**

9. Marioni R, Suderman MJ, Chen BH *et al.* Tracking the epigenetic clock across the human life course: a meta-analysis of longitudinal cohort data. *J. Gerontol. A Biol. Sci. Med. Sci.* 74(1), 57–61 (2019).
10. Grant CD, Jafari N, Hou L *et al.* A longitudinal study of DNA methylation as a potential mediator of age-related diabetes risk. *Geroscience* 39(5–6), 475–489 (2017).
11. Perna L, Zhang Y, Mons U, Holleczek B, Saum K-U, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin. Epigenetics* 8, 64 (2016).
12. Lind L, Ingelsson E, Sundström J, Siegbahn A, Lampa E. Methylation-based estimated biological age and cardiovascular disease. *Eur. J. Clin. Invest.* 48(2), (2018).
13. Marioni RE, Shah S, Mcrae AF *et al.* DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* 16(1), 25 (2015).
14. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. *Clin. Epigenetics* 11(1), 62 (2019).
15. Hannum G, Guinney J, Zhao L *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* 49(2), 359–367 (2013).

•• **Original Hannum clock conceptualization, development and application article.**

16. Horvath S, Oshima J, Martin GM *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford progeria syndrome and ex vivo studies. *Aging (Albany NY)* 10(7), 1758–1775 (2018).

•• **Original skin and blood clock conceptualization, development and application article.**

17. Levine ME, Lu AT, Quach A *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging* 10(4), 573–591 (2018).

•• **Original PhenoAge clock conceptualization, development and application article.**

18. Lu AT, Quach A, Wilson JG *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* 11(2), 303–327 (2019).

•• **Original GrimAge clock conceptualization, development and application article.**

19. El Khoury LY, Gorrie-Stone T, Smart M *et al.* Systematic underestimation of the epigenetic clock and age acceleration in older subjects. *Genome Biol.* 20(1), 283 (2019).

• **Highlighted for the first time the systematic underestimation of biological ages when DNA methylation (DNAm) clocks are applied to older cohorts.**

20. Tajuddin SM, Hernandez DG, Chen BH *et al.* Novel age-associated DNA methylation changes and epigenetic age acceleration in middle-aged African Americans and whites. *Clin. Epigenetics* 11(1), 119 (2019).
21. Liu Z, Chen BH, Assimes TL, Ferrucci L, Horvath S, Levine ME. The role of epigenetic aging in education and racial/ethnic mortality disparities among older US women. *Psychoneuroendocrinology* 104, 18–24 (2019).
22. Philibert R, Beach SRH, Lei M-K *et al.* Array-based epigenetic aging indices may be racially biased. *Genes* 11(6), 685 (2020).

• **Emphasizes the importance of keeping the influence of genetic variation in the accuracy of methylation-derived aging algorithms in mind in epigenetic epidemiology.**

23. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A systematic review and meta-analysis of environmental, lifestyle, and health factors associated with DNA methylation age. *J. Gerontol. A Biol. Sci. Med. Sci.* 75(3), 481–494 (2020).
24. Teo K, Chow CK, Vaz M, Rangarajan S, Yusuf S. The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *Am. Heart J.* 158(1), 1–7.e1 (2009).
25. Cronjé HT, Elliott HR, Nienaber-Rousseau C, Pieters M. Replication and expansion of epigenome-wide association literature in a black South African population. *Clin. Epigenetics* 12(1), 6 (2020).
26. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* 34(23), 3983–3989 (2018).
27. R Core Team. R: a language and environment for statistical computing. (2017). https://www.R-project.org/
28. Salas LA, Koestler DC, Butler RA *et al.* An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray. *Genome Biol.* 19(1), 64–64 (2018).
29. Mcewen LM, Jones MJ, Lin DTS *et al.* Systematic evaluation of DNA methylation age estimation with common preprocessing methods and the Infinium MethylationEPIC BeadChip array. *Clin. Epigenetics* 10(1), 123 (2018).
30. Fiorito G, Mccrory C, Robinson O *et al.* Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. *Aging* 11(7), 2045–2070 (2019).

31. Philibert R, Dogan M, Beach SR, Mills JA, Long JD. AHRR methylation predicts smoking status and smoking intensity in both saliva and blood DNA. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 183(1), 51–60 (2020).

32. Zhao W, Ammous F, Ratliff S *et al.* Education and lifestyle factors are associated with DNA methylation clocks in older African Americans. *Int. J. Environ. Res. Public Health* 16(17), 3141 (2019).

33. Quach A, Levine ME, Tanaka T *et al.* Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging* 9(2), 419–446 (2017).

34. Azab B, Camacho-Rivera M, Taioli E. Average values and racial differences of neutrophil lymphocyte ratio among a nationally representative sample of United States subjects. *PLoS One* 9(11), e112361 (2014).

35. Forget P, Khalifa C, Defour J-P, Latinne D, Van Pel M-C, De Kock M. What is the normal value of the neutrophil-to-lymphocyte ratio? *BMC Res. Notes* 10(1), 12 (2017).

36. Kelsey KT, Wiencke JK. Immunomethylomics: a novel cancer risk prediction tool. *Ann. Am. Thorac. Soc.* 15(Suppl. 2), S76–S80 (2018).

37. Astle WJ, Elding H, Jiang T *et al.* The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell* 167(5), 1415–1429.e19 (2016).

38. Li X, Ploner A, Wang Y *et al.* Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *eLife* 9, e51507 (2020).

- **Highlights the opportunity for and usefulness of combining various biological age estimators (including those that derive from physical, telomere and methylation measurements) to better predict adverse outcomes and provide a more holistic understanding of aging.**