

9-2016

## DNA methylation-based measures of biological age: meta-analysis predicting time to death.

B. Chen

R. Marioni

E. Colicino

M. Peters

C. Ward-Caviness

*See next page for additional authors*

Follow this and additional works at: [http://hsrc.himmelfarb.gwu.edu/smhs\\_crl\\_facpubs](http://hsrc.himmelfarb.gwu.edu/smhs_crl_facpubs)

 Part of the [Genetics and Genomics Commons](#), and the [Health and Medical Administration Commons](#)

---

### APA Citation

Chen, B., Marioni, R., Colicino, E., Peters, M., Ward-Caviness, C., Carty, C. L., & +several additional authors (2016). DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging*, 8 (9). <http://dx.doi.org/10.18632/aging.101020>

This Journal Article is brought to you for free and open access by the Clinical Research and Leadership at Health Sciences Research Commons. It has been accepted for inclusion in Clinical Research and Leadership Faculty Publications by an authorized administrator of Health Sciences Research Commons. For more information, please contact [hsrc@gwu.edu](mailto:hsrc@gwu.edu).

---

**Authors**

B. Chen, R. Marioni, E. Colicino, M. Peters, C. Ward-Caviness, Cara L. Carty, and +several additional authors

## DNA methylation-based measures of biological age: meta-analysis predicting time to death

Brian H. Chen<sup>1,2,3\*</sup>, Riccardo E. Marioni<sup>4,5,6\*</sup>, Elena Colicino<sup>7\*</sup>, Marjolein J. Peters<sup>8</sup>, Cavin K. Ward-Caviness<sup>9</sup>, Pei-Chien Tsai<sup>10</sup>, Nicholas S. Roetker<sup>11</sup>, Allan C. Just<sup>7</sup>, Ellen W. Demerath<sup>11</sup>, Weihua Guan<sup>12</sup>, Jan Bressler<sup>13</sup>, Myriam Fornage<sup>13,14</sup>, Stephanie Studenski<sup>1</sup>, Amy R. Vandiver<sup>15</sup>, Ann Zenobia Moore<sup>1</sup>, Toshiko Tanaka<sup>1</sup>, Douglas P. Kiel<sup>16,17</sup>, Liming Liang<sup>18,19</sup>, Pantel Vokonas<sup>18</sup>, Joel Schwartz<sup>18</sup>, Kathryn L. Lunetta<sup>20,2</sup>, Joanne M. Murabito<sup>2,21</sup>, Stefania Bandinelli<sup>22</sup>, Dena G. Hernandez<sup>23</sup>, David Melzer<sup>24</sup>, Michael Nalls<sup>23</sup>, Luke C. Pilling<sup>24</sup>, Timothy R. Price<sup>23</sup>, Andrew B. Singleton<sup>23</sup>, Christian Gieger<sup>9,25</sup>, Rolf Holle<sup>26</sup>, Anja Kretschmer<sup>9,25</sup>, Florian Kronenberg<sup>27</sup>, Sonja Kunze<sup>9,25</sup>, Jakob Linseisen<sup>9</sup>, Christine Meisinger<sup>9</sup>, Wolfgang Rathmann<sup>28</sup>, Melanie Waldenberger<sup>9,25</sup>, Peter M. Visscher<sup>4,6,29</sup>, Sonia Shah<sup>6,29</sup>, Naomi R. Wray<sup>6</sup>, Allan F. McRae<sup>6,29</sup>, Oscar H. Franco<sup>30</sup>, Albert Hofman<sup>18,30</sup>, André G. Uitterlinden<sup>8,30</sup>, Devin Absher<sup>31</sup>, Themistocles Assimes<sup>32</sup>, Morgan E. Levine<sup>33</sup>, Ake T. Lu<sup>33</sup>, Philip S. Tsao<sup>32,34</sup>, Lifang Hou<sup>35,36</sup>, JoAnn E. Manson<sup>37</sup>, Cara L. Carty<sup>38</sup>, Andrea Z. LaCroix<sup>39</sup>, Alexander P. Reiner<sup>40,41</sup>, Tim D. Spector<sup>10</sup>, Andrew P. Feinberg<sup>15,42</sup>, Daniel Levy<sup>2,43\*</sup>, Andrea Baccarelli<sup>7,44\*</sup>, Joyce van Meurs<sup>8\*</sup>, Jordana T. Bell<sup>10\*</sup>, Annette Peters<sup>9\*</sup>, Ian J. Deary<sup>4,45\*</sup>, James S. Pankow<sup>11\*</sup>, Luigi Ferrucci<sup>1\*</sup>, Steve Horvath<sup>33,46\*</sup>

<sup>1</sup>Longitudinal Studies Section, Translational Gerontology Branch, Intramural Research Program, National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA

<sup>2</sup>The NHLBI's Framingham Heart Study, Framingham, MA 01702, USA

<sup>3</sup>Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 01702, USA

<sup>4</sup>Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK

<sup>5</sup>Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

<sup>6</sup>Queensland Brain Institute, University of Queensland, Brisbane, QLD, Australia

<sup>7</sup>Laboratory of Environmental Epigenetics, Departments of Environmental Health Sciences and Epidemiology, Columbia University Mailman School of Public Health, New York, NY 10032, USA

<sup>8</sup>Department of Internal Medicine, Erasmus University Medical Centre, Rotterdam, 3000 CA, The Netherlands

<sup>9</sup>Institute of Epidemiology II, Helmholtz Zentrum München, 85764 Neuherberg, Germany

<sup>10</sup>Department of Twin Research and Genetic Epidemiology, Kings College London, London SE1 7EH, UK

<sup>11</sup>Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, MN 55455, USA

<sup>12</sup>Division of Biostatistics, University of Minnesota School of Public Health, Minneapolis, MN, 55455, USA

<sup>13</sup>Human Genetics Center, School of Public Health, University of Texas Health Sciences Center at Houston, Houston, TX, USA

<sup>14</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA

<sup>15</sup>Center for Epigenetics, Johns Hopkins University, Baltimore, MD 21205, USA

<sup>16</sup>Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, USA

<sup>17</sup>Institute for Aging Research, Hebrew Senior Life, Boston, MA 02215, USA

- <sup>18</sup>Department of Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
- <sup>19</sup>Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA
- <sup>20</sup>Department of Biostatistics, Boston University School of Public Health, Boston, MA 02118, USA
- <sup>21</sup>Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA 02118, USA
- <sup>22</sup>Geriatric Unit, Usl Centro Toscana, Florence, Italy
- <sup>23</sup>Laboratory of Neurogenetics, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda, MD 20814, USA
- <sup>24</sup>Epidemiology and Public Health, Medical School, University of Exeter, RILD, Exeter EX2 5DW, UK
- <sup>25</sup>Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, 85764 Neuherberg, Germany
- <sup>26</sup>Institute of Health Economics and Health Care Management, Helmholtz Zentrum München, 85764 Neuherberg, Germany
- <sup>27</sup>Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Innsbruck Medical University, Innsbruck 6020, Austria
- <sup>28</sup>Institute for Biometrics and Epidemiology, German Diabetes Center, Leibniz Center for Diabetes Research at Heinrich Heine University, 40225 Düsseldorf, Germany
- <sup>29</sup>University of Queensland Diamantina Institute, University of Queensland, Brisbane, Queensland, Australia
- <sup>30</sup>Department of Epidemiology, Erasmus University Medical Centre, Rotterdam, 3015 CN, The Netherlands
- <sup>31</sup>HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA
- <sup>32</sup>Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA;
- <sup>33</sup>Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA
- <sup>34</sup>VA Palo Alto Health Care System, Palo Alto CA 94304, USA
- <sup>35</sup>Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University Chicago, IL 60611, USA
- <sup>36</sup>Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University Chicago, IL 60611, USA
- <sup>37</sup>Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, and the Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02215, USA
- <sup>38</sup>Center for Translational Science Children's National Medical Center, George Washington University Washington, DC 20010, USA
- <sup>39</sup>Department of Family Medicine and Public Health, University of California-San Diego, La Jolla, CA 92093-0725, USA
- <sup>40</sup>Department of Epidemiology, University of Washington School of Public Health, Seattle, WA 98195, USA
- <sup>41</sup>Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
- <sup>42</sup>Departments of Medicine, Molecular Biology/Genetics, Oncology, and Biostatistics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA
- <sup>43</sup>Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD 01702, USA
- <sup>44</sup>Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
- <sup>45</sup>Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, UK
- <sup>46</sup>Department of Biostatistics, School of Public Health, University of California Los Angeles, Los Angeles, CA 90095, USA

\*Joint first or senior authors

**Correspondence to:** Steve Horvath; **email:** [shorvath@mednet.ucla.edu](mailto:shorvath@mednet.ucla.edu)

**Keywords:** all-cause mortality; lifespan; epigenetics; epigenetic clock; DNA methylation; mortality

**Received:** July 01, 2016

**Accepted:** August 18, 2016

**Published:** September 28, 2016

## ABSTRACT

Estimates of biological age based on DNA methylation patterns, often referred to as "epigenetic age", "DNAm age", have been shown to be robust biomarkers of age in humans. We previously demonstrated that independent of chronological age, epigenetic age assessed in blood predicted all-cause mortality in four human cohorts. Here, we expanded our original observation to 13 different cohorts for a total sample size of 13,089 individuals, including three racial/ethnic groups. In addition, we examined whether incorporating information on blood cell composition into the epigenetic age metrics improves their predictive power for mortality. All considered measures of epigenetic age acceleration were predictive of mortality ( $p \leq 8.2 \times 10^{-9}$ ), independent of chronological age, even after adjusting for additional risk factors ( $p < 5.4 \times 10^{-4}$ ), and within the racial/ethnic groups that we examined (non-Hispanic whites, Hispanics, African Americans). Epigenetic age estimates that incorporated information on blood cell composition led to the smallest p-values for time to death ( $p = 7.5 \times 10^{-43}$ ). Overall, this study a) strengthens the evidence that epigenetic age predicts all-cause mortality above and beyond chronological age and traditional risk factors, and b) demonstrates that epigenetic age estimates that incorporate information on blood cell counts lead to highly significant associations with all-cause mortality.

## INTRODUCTION

DNA methylation-based biomarkers, often referred to as "epigenetic age" or "epigenetic clock", are robust estimators of chronological age of an individual [1–4]. For example, a measure of epigenetic age based on levels of methylation in 353 CpG dinucleotide markers (cytosine linked to guanine by a phosphate group) allow the estimation of the age of an individual. This estimate is consistent across most types of biological specimens, including whole blood, brain, breast, kidney, liver, lung, and saliva and cell types, including CD4+ T cells, monocytes, B cells, glial cells, and neurons [3].

Recent studies suggested that epigenetic age is associated with age-related health outcomes above and beyond chronological age. For example, we and others have shown that individuals whose epigenetic age was greater than their chronological age (i.e., individuals exhibiting epigenetic "age acceleration") were at an increased risk for death from all causes, even after accounting for known risk factors [5–7]. Further, we recently showed that the offspring of semi-supercentenarians (subjects who reached an age of 105–109 years) have a lower epigenetic age than age-matched controls [8]. Based on these findings, it has been hypothesized that epigenetic age captures some aspect of biological age and the resulting susceptibility to disease and multiple health outcomes. A first step in testing this hypothesis is to test whether epigenetic age predicts longevity in multiple populations and across ethnic groups.

changes might be independent predictors of mortality [9–12]. Thus, it is of interest to understand whether considering information on blood cell composition in measures of epigenetic age improves their predictive power for mortality.

Here, we evaluated the ability to predict time to death for blood-based epigenetic age measures, both published and novel measures that incorporate information on blood cell composition. Due to the well documented age-related changes in blood cell composition, we distinguished epigenetic measures of age that were independent of changes in blood cell composition (cell-intrinsic measures), and measures that incorporated age-related changes in blood cell composition ("extrinsic" measures). By increasing the number of independent cohort studies, we more than doubled the number of mortality events available for analysis, which allowed for detailed subgroup analyses including those based on race/ethnicity.

## RESULTS

### Cohort studies

Our meta-analysis included 13 population-based cohorts. An overview of the cohorts is provided in Table 1. Our study involved 3 racial/ethnic groups: non-Hispanic whites ( $n=9,215$ ), Hispanics ( $n=431$ ), and Blacks ( $n=3,443$ ). Detailed descriptions of each cohort can be found in the Supplemental Materials.

## Epigenetic age estimation

We used two methods for estimating the epigenetic age of each blood sample (Table 2). First, we used the approach by Horvath (2013) based on 353 CpGs, as described in [3] and Methods. Second, we used the approach by Hannum et al. (2013) based on 71 CpGs [2]. Both epigenetic age estimates were correlated with chronological age at the time of blood draw (Table 1) with biweight midcorrelation coefficients ranging from 0.65 to 0.89. But birth cohorts were excluded from this correlation analysis because it is not meaningful to calculate correlations with chronological age in this situation. The Horvath and Hannum estimates were also highly correlated with each other ( $r=0.76$ ) even though the underlying sets of CpGs share only 6 CpGs in common. (Supplementary Table 1).

## Estimated blood cell counts that relate to chronological age

We estimated the abundance of ten blood cell types based on observed DNA methylation patterns (Methods) – exhausted/senescent CD8<sup>+</sup> T cells (CD8+CD28-CD45RA-), CD8<sup>+</sup> naïve, CD8<sup>+</sup> total, CD4<sup>+</sup> naïve, CD4<sup>+</sup> total, natural killer cells, B cells, monocytes, granulocytes, and plasmablasts. To study age-related changes in blood cell composition, we correlated these estimated blood cell counts with chronological age in all of the cohort studies (Supplementary Table 2). Our results are congruent with findings from flow cytometric studies that demonstrate that the abundance of naïve CD8<sup>+</sup> T cells decreases with age (reflecting thymic involution), whereas exhausted/senescent CD8<sup>+</sup> T cells increase with age [9–12].

**Table 1. Baseline characteristics of participating cohorts.**

Cohort	N	N <sub>deaths</sub> (%)	Follow-up duration (years)*	Age (years)*	r <sub>Horvath</sub> <sup>†</sup>	r <sub>Hannum</sub> <sup>‡</sup>
1. WHI (White)	995	309 (31%)	15.4 (14.0-16.4)	68 (65-72)	0.67 (p=5.1x10 <sup>-131</sup> )	0.73 (p=8.0x10 <sup>-167</sup> )
2. WHI (Black)	675	176 (26%)	15.4 (13.7-16.5)	62 (57-67)	0.70 (p=1.2x10 <sup>-100</sup> )	0.76 (p=3.0x10 <sup>-128</sup> )
3. WHI (Hispanic)	431	78 (18%)	15.2 (14.1-16.3)	61 (56-67)	0.78 (p=8.9x10 <sup>-90</sup> )	0.79 (p=1.3x10 <sup>-93</sup> )
4. LBC 1921	445	312 (70%)	10.2 (6.2-12.9)	79 (78-79)	0.15 (p=1.5x10 <sup>-3</sup> )	0.13 (p=6.0x10 <sup>-3</sup> )
5. LBC 1936	919	106 (12%)	7.5 (6.9-8.4)	69 (68-70)	0.15 (p=4.9x10 <sup>-6</sup> )	0.16 (p=1.1x10 <sup>-6</sup> )
6. NAS	647	221 (34%)	11.6 (8.6-12.9)	72 (68-77)	0.69 (p=1.3x10 <sup>-92</sup> )	0.76 (p=8.2x10 <sup>-123</sup> )
7. ARIC (Black)	2,768	1,075 (39%)	20.3 (14.3-21.4)	57 (52-62)	0.65 (p<1x10 <sup>-200</sup> )	0.71 (p<1x10 <sup>-200</sup> )
8. FHS	2,614	236 (9%)	6.2 (5.6-6.9)	66 (60-73)	0.84 (p<1x10 <sup>-200</sup> )	0.86 (p<1x10 <sup>-200</sup> )
9. KORA	1,257	42 (3%)	4.4 (4.0-4.8)	61 (54-68)	0.84 (p<1x10 <sup>-200</sup> )	0.88 (p<1x10 <sup>-200</sup> )
10. InCHIANTI	506	91 (18%)	15.0 (14.6-15.5)	67 (57-73)	0.82 (p=3.2x10 <sup>-124</sup> )	0.85 (p=2.1x10 <sup>-142</sup> )
11. Rotterdam	710	32 (5%)	5.6 (5.3-5.8)	58 (54-62)	0.72 (p=1.9x10 <sup>-114</sup> )	0.76 (p=1.3x10 <sup>-134</sup> )
12. Twins UK	805	30 (4%)	8.5 (7.5-8.5)	58 (51-65)	0.87 (p<1x10 <sup>-200</sup> )	0.89 (p<1x10 <sup>-200</sup> )
13. BLSA (white)	317	26 (8%)	5.3 (4.0-6.6)	66 (58-73)	0.85 (p=1.1x10 <sup>-89</sup> )	0.88 (p=7.2x10 <sup>-104</sup> )
<b>Total</b>	<b>13,089</b>	<b>2734 (21%)</b>				

The last 3 columns report robust correlation coefficients (biweight midcorrelation) between chronological age and two epigenetic age estimates (Horvath and Hannum).

\* Median (25<sup>th</sup> percentile - 75<sup>th</sup> percentile)

† Biweight midcorrelation coefficient of chronological age with epigenetic age using the Horvath method.

‡ Biweight midcorrelation coefficient of chronological age with epigenetic age using the Hannum method.

## Measures of epigenetic age acceleration

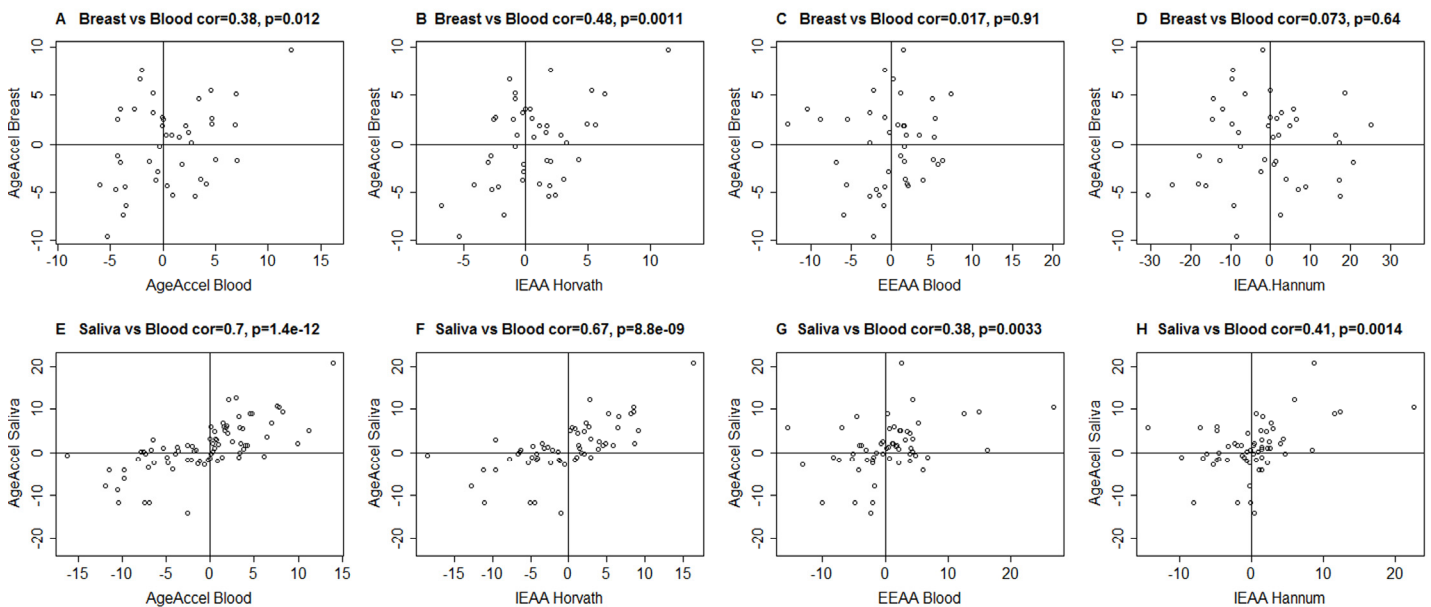
Despite high correlations, epigenetic age can deviate substantially from chronological age at the individual level. The difference between epigenetic age and chro-

nological age can be used to define "delta age" but the resulting measure exhibits a negative correlation with chronological age. By contrast, all of our measures of epigenetic age acceleration are defined such that they are uncorrelated with chronological age.

**Table 2. Overview of various measures of epigenetic age acceleration.**

Measure of age acceleration	Short name of measure	Epigenetic age estimate	Uses blood counts	Correlation with blood counts	Conserved in breast tissue
(Universal) epigenetic age acceleration	<i>AgeAccel</i> <sub>Horvath</sub> ( <i>AgeAccel</i> )	Horvath: 353 CpGs	no	weak	yes
Intrinsic epigenetic age acceleration (Horvath)	<i>IEAA</i> .Horvath ( <i>IEAA</i> )	Horvath: 353 CpGs	yes	very weak	yes
Age acceleration based on Hannum	<i>AgeAccel</i> <sub>Hannum</sub>	Hannum: 71 CpGs	no	moderate	no
Intrinsic epigenetic age acceleration (Hannum)	<i>IEAA</i> .Hannum	Hannum: 71 CpGs	yes	very weak	no
Extrinsic epigenetic age acceleration	<i>EEAA</i>	Enhanced Hannum	yes	strong	no

Description of the differences between epigenetic age and age acceleration measures. Column "Correlation with blood counts" relates to Supplementary Table 4. Column "Conserved in breast tissue" relates to Figure 1.



**Figure 1. Epigenetic age acceleration in blood versus that in breast or saliva.** (A-D) Epigenetic age acceleration in healthy female breast tissue (y-axis) versus various measures of epigenetic age acceleration in blood: (A) universal measure of age acceleration in blood, (B) intrinsic epigenetic age acceleration based on the Horvath estimate of epigenetic age, (C) extrinsic epigenetic age acceleration, (D) intrinsic epigenetic age acceleration based on the Hannum estimate of epigenetic age. (E-H) analogous plots for epigenetic age acceleration in saliva (y-axis) and (E) *AgeAccel*, (F) *IEAA* based on Horvath, (G) *EEAA*, (H) *IEAA* based on the Hannum estimate. The y-axis of each plot represents the universal measure of age acceleration defined as the raw residual resulting from regressing epigenetic age (based on Horvath) on chronological age.

An overview of several measures of epigenetic age acceleration is presented in Table 2. One such measure (denoted as *AgeAccel*) is defined as the residual that results from regressing epigenetic age on chronological age. Thus, a positive value of *AgeAccel* indicates that the epigenetic age is higher than expected, based on chronological age. These Horvath and Hannum based measures of age acceleration are denoted by *AgeAccel<sub>Horvath</sub>* and *AgeAccel<sub>Hannum</sub>*, respectively. For the sake of brevity and consistency with other publications from our group, we abbreviate *AgeAccel<sub>Horvath</sub>* as *AgeAccel*.

*AgeAccel<sub>Hannum</sub>* and to a lesser extent *AgeAccel* were previously shown to correlate with blood cell counts [5]. Thus, we distinguished two broad categories of measures of epigenetic age acceleration when dealing with DNA methylation from blood or peripheral blood mononuclear cells (PBMCs): intrinsic and extrinsic epigenetic measures, which are independent of, or enhanced by blood cell count information, respectively. We define *intrinsic* epigenetic age acceleration (*IEAA*) as the residual resulting from regressing epigenetic age on chronological age and measures of blood cell counts (Methods). By definition, *IEAA* is not correlated with chronological age and is weakly correlated with estimated measures of blood cell counts (Supplementary Table 4). *IEAA* is meant to capture cell-intrinsic properties of the aging process that exhibit some pres-

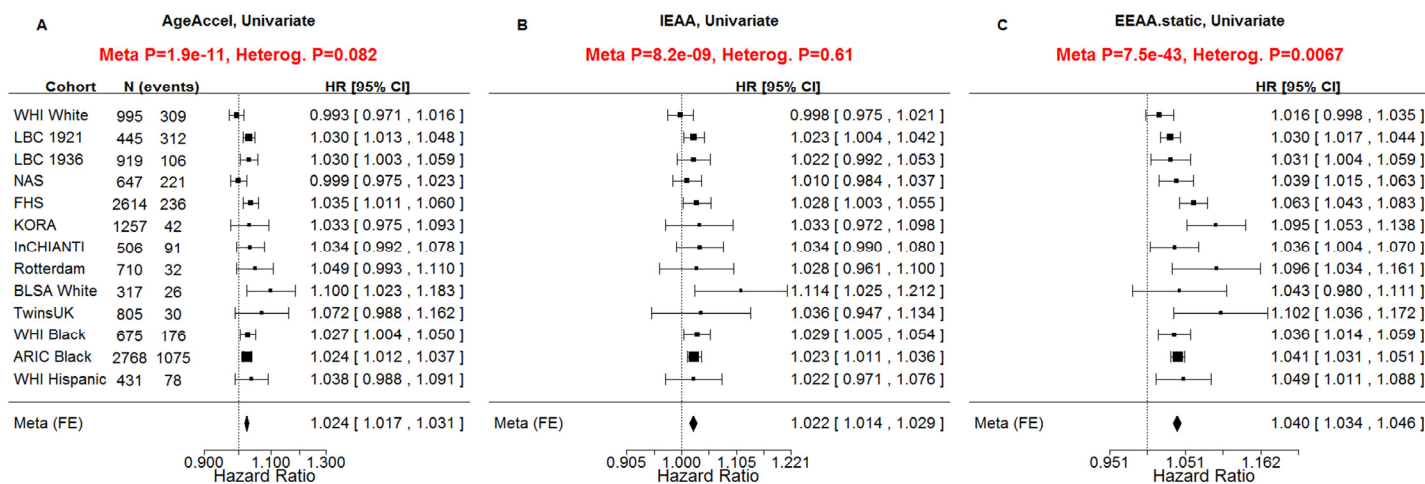
ervation across various cell types and organs. Compared to our other measures of age acceleration, *IEAA*, adapted from the Horvath measure of epigenetic age, exhibited significant correlations with epigenetic age acceleration in breast tissue ( $r=0.48$ ,  $p=0.0011$ , Figure 1B) and saliva ( $r=0.67$ ,  $p=8.8 \times 10^{-9}$ , Figure 1F). By contrast, an analogous measure of *IEAA* based on the Hannum measure showed much weaker correlations ( $r=0.073$  in breast and  $r=0.41$  in saliva Figure 1D, 1H). For this reason, we focused on the Horvath measure of *IEAA*.

The age-related changes to blood cell composition (Supplementary Table 4) can be leveraged to capture aspects of immunosenescence. Using these measures, we derived a novel extrinsic epigenetic age acceleration (*EEAA*) measure by up-weighting the blood cell count contributions of *AgeAccel<sub>Hannum</sub>* (Methods and Supplementary Table 4).

Descriptive statistics (minimum, maximum, median) of the measures of epigenetic age acceleration can be found in Supplementary Table 3.

### Cox regression models of all-cause mortality

We used Cox regression models to assess the predictive value of our measures of epigenetic age acceleration for all-cause mortality. All of our Cox models were adjusted

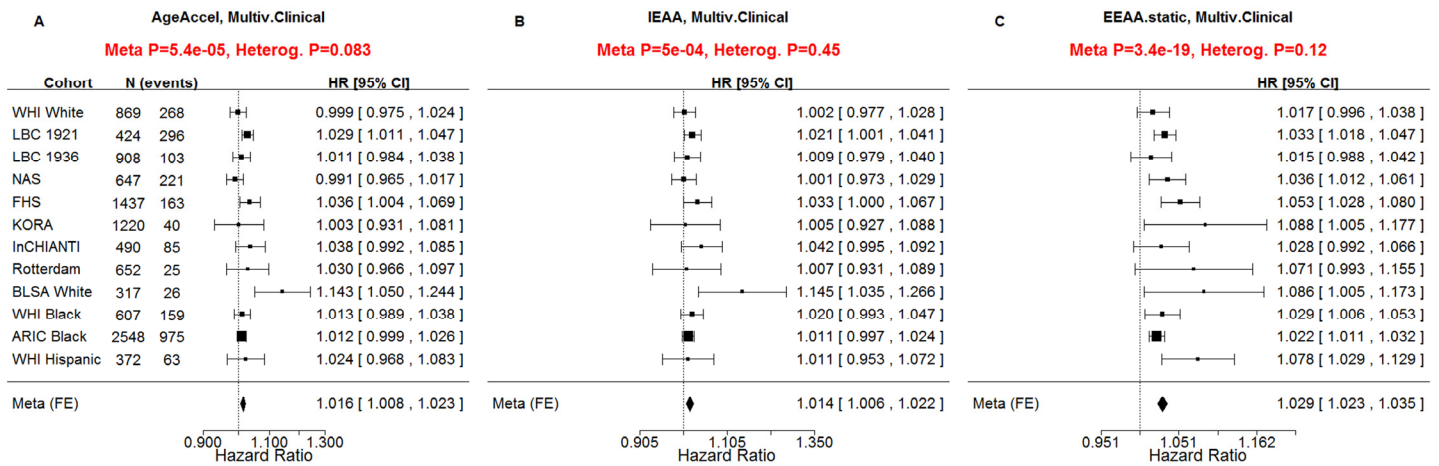


**Figure 2. Univariate Cox regression meta-analysis of all-cause mortality.** A univariate Cox regression model was used to relate the censored survival time (time to all-cause mortality) to (A) the universal measure of age acceleration (*AgeAccel*), (B) intrinsic epigenetic age acceleration (*IEAA*), (C) extrinsic epigenetic age acceleration (*EEAA*). The rows correspond to the different cohorts. Each row depicts the hazard ratio and a 95% confidence interval. The coefficient estimates from the respective studies were meta-analyzed using a fixed-effect model weighted by inverse variance (implemented in the *metafor* R package [34]). It is not appropriate to compare the hazard ratios and confidence intervals of the different measures directly because the measures have different scales/distributions. However, it is appropriate to compare the meta analysis p values (red sub-title of each plot). The p-value of the heterogeneity test (Cochran's Q-test) is significant if the cohort-specific estimates differed substantially.

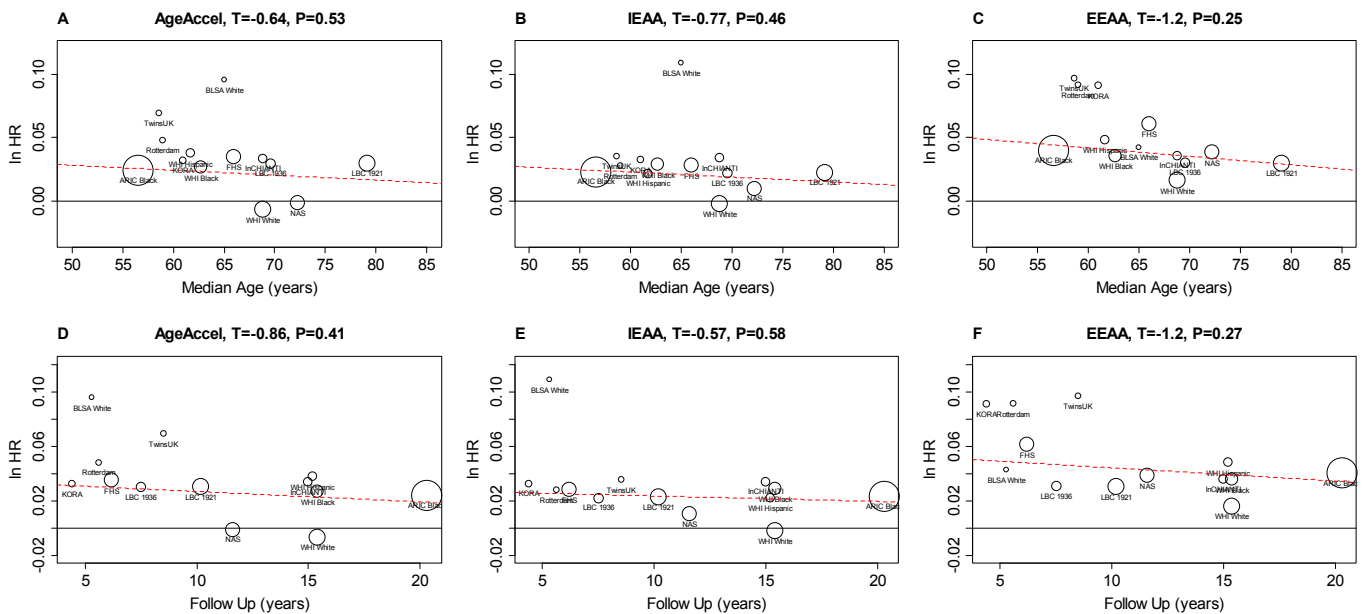


for the age at baseline (blood draw). Additional multivariate models further adjusted for covariates assessed at baseline (chronological age, body mass index,

educational level, alcohol intake, smoking pack-years, prior history of diabetes, prior history of cancer, hypertension status, self-reported recreational physical activity).



**Figure 3. Multivariate Cox regression meta-analysis adjusted for clinical covariates.** A multivariate Cox regression model was used to relate the censored survival time (time to all-cause mortality) to (A) the universal measure of age acceleration (*AgeAccel*), (B) intrinsic epigenetic age acceleration (*IEAA*), (C) extrinsic epigenetic age acceleration (*EEAA*). The multivariate Cox regression model included the following additional covariates: chronological age, body mass index (category), educational level (category), alcohol intake, smoking pack years, prior history of diabetes, prior history of cancer, hypertension status, recreational physical activity (category). The rows correspond to the different cohorts. Each row depicts the hazard ratio and a 95% confidence interval. The coefficient estimates from the respective studies were meta-analyzed using a fixed-effect model weighted by inverse variance (implemented in the *metafor* R package [34]). The sub-title of each plot reports the meta-analysis p-value and a heterogeneity test p-value (Cochran's Q-test).



**Figure 4. Hazard ratio of death versus cohort characteristics.** Each circle corresponds to a cohort (data set). Circle sizes correspond to the square root of the number of observed deaths, because the statistical power of a Cox model is determined by the number of observed deaths. (A-C) The y-axis of each panel corresponds to the natural log of the hazard ratio (ln HR) of a univariate Cox regression model for all-cause mortality. Each panel corresponds to a different measure of epigenetic age acceleration (A) universal age acceleration, (B) intrinsic age acceleration, (C) extrinsic age acceleration. Panels (D-F) are analogous to those in A-C but the x-axis corresponds to the median age of the subjects at baseline (Table 1). The title of each panel reports the Wald test statistic (T) and corresponding p-value resulting from a weighted linear regression model (y regressed on x) where each point (data set) is weighted by the square root of the number of observed deaths. The dotted red line represents the regression line. The black solid line represents the line of identity (i.e., no association).

**Table 3. Subgroup analysis by demographic factors.**

Subgroup	Age-adjusted		Full model	
	HR	p-value	HR	p-value
<b>Race</b>				
White	1.05	3.0x10 <sup>-26</sup>	1.03	1.3x10 <sup>-5</sup>
Black	1.04	7.8x10 <sup>-20</sup>	1.02	7.6x10 <sup>-3</sup>
Hispanic	1.05	1.1x10 <sup>-2</sup>	1.06	5.3x10 <sup>-2</sup>
<i>P</i> interaction		0.62		0.14
<b>Sex</b>				
Men	1.04	7.1x10 <sup>-15</sup>	1.03	1.9x10 <sup>-2</sup>
Women	1.04	3.7x10 <sup>-10</sup>	1.03	1.9x10 <sup>-5</sup>
<i>P</i> interaction		0.63		0.95
<b>Follow-up duration</b>				
<5 years	1.02	0.20	0.98	0.79
5-10 years	1.02	1.8x10 <sup>-3</sup>	1.02	0.17
>10 years	1.03	4.5x10 <sup>-9</sup>	1.02	4.1x10 <sup>-2</sup>
<i>P</i> interaction		0.67		0.84
<b>BMI categories</b>				
Underweight	1.11	9.4x10 <sup>-3</sup>	1.04	8.9x10 <sup>-3</sup>
Normal	1.06	6.1x10 <sup>-19</sup>	1.04	2.3x10 <sup>-2</sup>
Overweight	1.04	1.46x10 <sup>-8</sup>	1.03	5.0x10 <sup>-2</sup>
Obese	1.04	2.2x10 <sup>-11</sup>	1.02	7.1x10 <sup>-2</sup>
<i>P</i> interaction		0.05		0.75
<b>Smoking status</b>				
Never	1.03	6.9x10 <sup>-6</sup>	1.04	4.8x10 <sup>-3</sup>
Former	1.05	4.2x10 <sup>-22</sup>	1.03	6.3x10 <sup>-4</sup>
Current	1.06	2.1x10 <sup>-4</sup>	1.01	0.47
<i>P</i> interaction		0.05		0.20
<b>Physical activity status</b>				
Yes	1.05	3.8x10 <sup>-6</sup>	1.02	1.9x10 <sup>-3</sup>
No	1.03	2.5x10 <sup>-2</sup>	1.03	2.2x10 <sup>-2</sup>
<i>P</i> interaction		0.23		0.65

Age-adjusted and fully adjusted associations for EEAA to all-cause mortality by subgroup (rows). The fully adjusted model includes the following covariates: body mass index, educational level, alcohol intake, smoking pack-years, prior history of diabetes, prior history of cancer, hypertension status, self-reported recreational physical activity.

Our novel measure of extrinsic age acceleration *EEAA* led to smaller p-values for the associations with all-cause mortality than the original measure *AgeAccel<sub>Hannum</sub>* in univariate Cox models ( $p_{EEAA}=7.5 \times 10^{-43}$ ,  $p_{AgeAccelHannum}=1.4 \times 10^{-34}$ , Supplementary Figure 1) and in multivariate Cox models ( $p_{EEAA}=3.4 \times 10^{-19}$ ,  $p_{AgeAccelHannum}=6 \times 10^{-15}$ , Supplementary Figure 2). Further, when both *EEAA* and *AgeAccel<sub>Hannum</sub>* were included in the same Cox model, only *EEAA* remained significant in the WHI data and FHS univariate models. Since these results indicate that *EEAA* outperforms the closely related measure *AgeAccel<sub>Hannum</sub>* when it comes to mortality prediction, we removed the latter from subsequent analyses.

All considered measures of epigenetic age acceleration were predictive of time to death in univariate Cox models ( $p_{AgeAccel}=1.9 \times 10^{-11}$ ,  $p_{IEAA}=8.2 \times 10^{-9}$ ,  $p_{EEAA}=7.5 \times 10^{-43}$ , Figure 2) and multivariate Cox models adjusting for risk factors and pre-existing disease status ( $p_{AgeAccel}=5.4 \times 10^{-5}$ ,  $p_{IEAA}=5.0 \times 10^{-4}$ ,  $p_{EEAA}=3.4 \times 10^{-19}$ , Figure 3).

### Interpreting effect sizes and variance of epigenetic age acceleration

Subjects differed substantially in terms of their measures of epigenetic age acceleration, e.g. *EEAA* ranged from -28 to 28 years in the WHI (standard deviation =6.4 years, Supplementary Table 3).

About five percent of the participants of the WHI exhibited an *EEAA* value larger than 10, which is associated with a 48% increased hazard of death as can be seen from the following calculation. The HR of *EEAA* is 1.040 if *EEAA*=1 (Figure 2c) but it is  $HR=1.48=(1.040)^{10}$  if *EEAA*=10. Negative values of age acceleration were associated with a lower hazard of mortality. For example, 20% of subjects had an *EEAA* value less than -5, which is associated with an 18% decrease in the hazard of death ( $HR=0.82=1.04^{-5}$ ).

### Subgroup analysis

With few exceptions, we found that the associations between *EEAA* and time to death remained highly signi-

**Table 4. Subgroup analysis by prevalent disease status.**

Subgroup	Age-adjusted		Full model	
	HR	p-value	HR	p-value
<b>Cancer status</b>				
Yes	1.05	$2.5 \times 10^{-10}$	1.02	0.18
No	1.05	$2.3 \times 10^{-13}$	1.03	$1.7 \times 10^{-4}$
<i>P</i> interaction		0.92		0.73
<b>Coronary artery disease status</b>				
Yes	1.04	$2.4 \times 10^{-5}$	1.01	0.60
No	1.04	$1.5 \times 10^{-12}$	1.02	$1.5 \times 10^{-4}$
<i>P</i> interaction		0.43		0.99
<b>Hypertension status</b>				
Yes	1.04	$7.4 \times 10^{-17}$	1.03	$2.9 \times 10^{-3}$
No	1.05	$7.1 \times 10^{-6}$	1.02	$8.6 \times 10^{-3}$
<i>P</i> interaction		0.41		0.45
<b>Type 2 diabetes status</b>				
Yes	1.04	$8.6 \times 10^{-13}$	1.03	$1.7 \times 10^{-3}$
No	1.04	$1.2 \times 10^{-10}$	1.02	$9.3 \times 10^{-3}$
<i>P</i> interaction		0.70		0.25

Age-adjusted and fully adjusted associations for *EEAA* to all-cause mortality in different subgroups (rows). The fully adjusted model includes the following covariates: body mass index, educational level, alcohol intake, smoking pack-years, prior history of diabetes, prior history of cancer, hypertension status, self-reported recreational physical activity.

ficant in subgroups stratified by race, sex, follow-up duration, body mass index, smoking status, physical activity (Table 3) and in subgroups stratified by prevalent disease at baseline such as cancer, coronary artery disease, hypertension and type 2 diabetes (Table 4). Only one subgroup led to an insignificant finding ( $p>0.05$ ) in our univariate model analysis: namely subjects with less than 5 years of follow up (Table 3). For multivariate models, we failed to observe significant associations for the following subgroups: i) less than 5 years of follow up, ii) between 5 and 10 years of follow up, iii) current smokers, iv) obese individuals, v) Hispanics, vi) individuals with cancer, and vii) subjects with coronary artery disease. The insignificant results in multivariate models in cancer patients or CAD patients might reflect the relatively low sample sizes or that epigenetic age acceleration is dwarfed by other predictors of mortality in subjects with severe diseases. Hazard ratio estimates remained highly consistent across all subgroups examined.

We did not observe significant differences in the estimated hazard ratios across any subgroup (Tables 3 and 4). Specifically, racial/ethnic differences in HR were not observed (interaction  $p=0.62$  in age-adjustment models and  $p=0.14$  in full models). Overall, these subgroup analysis results confirm that epigenetic age acceleration is an independent predictor of earlier mortality even after adjusting for possible confounders and within major subgroups of the population.

### **Hazard ratio of death versus follow up time and median age**

The large number of cohorts allowed us to relate cohort characteristics (such as median age or median follow up removing time) to strength of association with mortality. We did not find a statistically significant relationship between the hazard ratio of death for the median age of the cohort or the follow up time (Figure 4).

### **Robustness analysis**

To assess the robustness of our findings, we also carried out a leave-one-out analysis by re-running the metaanalysis after removing data from individual cohorts. The resulting  $p$ -values are highly robust with respect to a single data set from the analysis (Supplementary Table 5). In our study, we used a fixed effects meta-analysis method for the sake of consistency with previous analyses [5]. However, our results remain qualitatively the same after using a random effects meta-analysis method (Supplementary Figure 4).

## **DISCUSSION**

The current study corroborates previous findings regarding the predictive power of DNA methylation-based biomarkers of age for mortality [5,6,8]. We further examined novel variants of these measures that are either independent of blood cell counts or are enhanced by changes in blood cell sub-populations. We showed that the extrinsic measure *EEAA* out-performs previous measures of age acceleration when it comes to predicting all-cause mortality. Furthermore, the associations between epigenetic age acceleration and mortality did not differ significantly across subgroups of race/ethnicity, sex, BMI, smoking status, physical activity status, or major chronic diseases. The consistency of the associations across multiple subgroups lends support to the notion that epigenetic age acceleration captures some aspect of biological aging over and above chronological age and other risk factors.

The development of suitable measures of biological age has been a key goal in the field of aging research [13]. Many biomarkers of age have been posited including epigenetic alterations of the DNA (e.g., DNA methylation), transcriptomic changes in blood [14], telomere length [15], whole-body function such as gait speed (reviewed in [16]). The current study does not aim to replace existing blood based biomarkers, but rather, we aimed to demonstrate that it complements existing markers. Above all, this study shows that epigenetic age captures an aspect of biological age, as assessed through lifespan, above and beyond chronological age, blood cell composition, and a host of traditional risk factors of mortality.

The measures of epigenetic age acceleration are attractive because they are highly robust and because their measurement only involve DNA methylation data. While actual flow cytometry data will always be preferable to imputed blood cell count data (based on DNA methylation data), the measures of age acceleration do not require the measurement of flow data. Rather, measures of intrinsic and extrinsic epigenetic age used blood cell count estimates resulting from DNA methylation data. The measure of extrinsic age acceleration *EEAA* reflects aspects of immunosenescence because, by construction, it correlates with age-related changes in blood cell composition, such as T lymphocyte populations, which underlie much of the age-related decline in the protective immune response [9–12]. Thus, the high predictive significance of *EEAA* for all-cause mortality probably reflects the fact that it assesses multiple aspects of the biological age of the immune system including both changes in blood cell

composition and cell-intrinsic epigenetic changes. It has been known for decades that poor T cell functioning is predictive of mortality [17].

The findings surrounding the predictive utility of intrinsic epigenetic age acceleration are biologically compelling and point to a new frontier in aging research. Our study strongly suggests *IEAA* is reflective of an intrinsic epigenetic clock that is associated with mortality independent of chronological age, changes in blood cell composition, and traditional risk factors of mortality. *IEAA* probably captures a cell-type independent component of the aging process for the following reasons. First, *IEAA* is moderately preserved across different tissues and cell types collected from the same subject (Figure 1). Second, *IEAA* but not *EEAA* is predictive of lung cancer [18]. Third, only *IEAA* and *AgeAccel* relate to centenarian status [8].

Overall, our results inform the ongoing debate about whether epigenetic biomarkers of age capture an aspect of biological age. While epigenetic processes are unlikely to be the only mediators of chronological age on mortality—in fact, multiple risk factors have stronger effects on mortality—our results suggest that at least one of the mediating processes relates to the epigenetic age of blood tissue and that this process is independent of age-dependent changes in blood cell composition. Future studies will be useful for gaining a mechanistic understanding of this intrinsic epigenetic aging process.

## MATERIALS AND METHODS

### Measures of epigenetic age

We used an epigenetic biomarker of age based on 353 CpG markers as one measure of epigenetic age because: a) it is an accurate measurement of age across multiple tissues [3]; b) we previously showed that it is predictive of all-cause mortality [5]; c) it correlated with measures of cognitive/physical fitness and neuro-pathology in the elderly [19,20]; and d) it was associated with conditions that are of interest in aging research including Down's syndrome [21], Huntington's disease [22], Parkinson's disease [23], obesity [24], HIV infection [25], menopause [26], centenarian status [27], ethnicity and sex [28], and cellular senescence [3,29]. This epigenetic age estimator not only lends itself to measuring aging effects in elderly subjects; but also applies to prenatal brain samples [30] and blood samples from minors [31]. Epigenetic age is defined as the predicted value of age based on the DNA methylation levels of 353 CpGs. Mathematical details and software tutorials for estimating epigenetic age can be found in the additional

files of [3]. All of the described epigenetic measures of aging and age acceleration are implemented in our freely available software (<https://dnamage.genetics.ucla.edu>) [3].

### DNA methylation age estimate by Hannum et al (2013)

We also used an alternative measure of epigenetic age developed by Hannum et al (2013) [2]. The resulting age estimate is based on the 71 CpGs and coefficient values from the third supplementary table [2]. The authors developed this age prediction method by using an elastic net regression model for predicting chronological age based on DNA methylation levels from whole blood.

### Measures of epigenetic age acceleration

Table 2 provides an overview of our measures of epigenetic age acceleration. The universal measure of age acceleration (*AgeAccel*), which is valid for a wide range of tissue types, is defined as the residual resulting from a linear regression model that regresses the Horvath estimate of epigenetic age on chronological age. Thus, a positive value for *AgeAccel* indicates that the observed epigenetic age is higher than that predicted, based on chronological age. *AgeAccel* has a relatively weak correlation with blood cell counts [25], but it still relates to estimated blood cell counts, as seen in Supplementary Table 4.

To estimate "pure" epigenetic aging effects that are not influenced by differences in blood cell counts ("intrinsic" epigenetic age acceleration, *IEAA*), we obtained the residual resulting from a multivariate regression model of epigenetic age on chronological age and various blood immune cell counts (naïve CD8+ T cells, exhausted CD8+ T cells, plasmablasts, CD4+ T cells, natural killer cells, monocytes, and granulocytes) imputed from methylation data.

Extrinsic epigenetic age acceleration measures capture both cell intrinsic methylation changes and extracellular changes in blood cell composition. Our measure of *EEAA* is defined using the following three steps. First, we calculated the epigenetic age measure from Hannum et al [2], which already correlated with certain blood cell types [5]. Second, we increased the contribution of immune blood cell types to the age estimate by forming a weighted average of Hannum's estimate with 3 cell types that are known to change with age: naïve (CD45RA+CCR7+) cytotoxic T cells, exhausted (CD28-CD45RA-) cytotoxic T cells, and plasmablasts using the Klemmer-Doubal approach [32]. The weights

used in the weighted average are determined by the correlation between the respective variable and chronological age [32]. The weights were chosen on the basis of the WHI data. Thus, the same (static) weights were used for all data sets. *EEAA* was defined as the residual variation resulting from a univariate model regressing the resulting age estimate on chronological age. By construction, *EEAA* is positively correlated with the estimated abundance of exhausted CD8+ T cells, plasmablast cells, and a negative correlated with naive CD8+ T cells. Blood cell counts were estimated based on DNA methylation data as described in the next section. By construction, the measures of *EEAA* track both age related changes in blood cell composition and intrinsic epigenetic changes. None of our four measures of epigenetic age acceleration are correlated with chronological age.

### Estimating blood cell counts based on DNA methylation levels

We estimate blood cell proportions using two different software tools. Houseman's estimation method [33], which is based on DNA methylation signatures from purified leukocyte samples, was used to estimate the proportions of cytotoxic (CD8+) T cells, helper (CD4+) T, natural killer, B cells, and granulocytes. The software does not allow us to identify the type of granulocytes in blood (neutrophil, eosinophil, or basophil) but we note that neutrophils tend to be the most abundant granulocyte (~60% of all blood cells compared with 0.5-2.5% for eosinophils and basophils). To estimate the percentage of exhausted CD8+ T cells (defined as CD28-CD45RA-), plasmablasts, and the number (count) of naïve CD8+ T cells (defined as CD45RA+CCR7+), we used the "Horvath method" [25], which is implemented in the advanced analysis option of the epigenetic age calculator software [3]. We and others have shown that imputed blood cell counts have moderately high correlations with corresponding flow cytometric data, e.g.  $r=0.86$  for naïve CD4+ T cells,  $r=0.68$  for naïve CD8+T, and  $r=0.49$  for exhausted CD8+ T cells [28].

### Cox regression models and meta-analysis

Here, we used Cox models for analyzing the censored survival time data (from the age at blood draw until age at death or last follow-up). We regressed the censored survival times on covariates using Cox regression models implemented in the R function *coxph* in the *survival* package. The resulting coefficient values (interpreted as log hazard ratios) and standard errors were combined using the R software package *metafor* [34]. The meta-analysis was carried out with the R

command *rma* (with arguments *method="FE"* to get fixed effects estimates). The forest plots were created using the R function *forest* (with argument *atransf=exp* to exponentiate the estimate of the log hazard ratios).

### Sample exclusions

In addition to cohort-specific quality checks, we further excluded individuals who had ever been diagnosed with leukemia (ICD-9: 203-208), reported receiving chemotherapy, and whose methylation beta value distributions deviated substantially from a gold standard (according to the quality statistic *corSampleVSgold standard*<0.80 from the online age calculator [35–37]).

## ACKNOWLEDGEMENTS

### Acknowledgement ARIC

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C). The authors thank the staff and participants of the ARIC study for their important contributions. Methylation profiling of the ARIC samples was also supported by the National Institutes of Health (NIH) American Recovery and Reinvestment Act of 2009 (ARRA) Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium (5RC2HL102419) (PI: E. Boerwinkle).

### Acknowledgement Rotterdam

This study was funded by The Netherlands Society for Scientific Research (NWO) VIDI Grant 917103521. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

The generation and management of the Illumina 450K methylation array data for the Rotterdam Study were executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. The methylation data was funded by the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC,

and by the Netherlands Organization for Scientific Research (NWO; project number 184021007) and made available as a Rainbow Project (RP3; BIOS) of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL).

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Mr. Michael Verbiest, Ms. Mila Jhamai, Ms. Sarah Higgins, Mr. Marijn Verkerk, for their help in creating the methylation database.

#### Acknowledgement KORA

The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF), by the State of Bavaria, the Federal Ministry of Health (Berlin, Germany), the Ministry of Innovation, Science, Research and Technology of the state North Rhine-Westphalia (Düsseldorf, Germany), and the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ. The work is partly supported by grants from the European Union's Seventh Framework Program (FP7-Health) under grant agreement no. 305280 (MIMomics) and by the BMBF: e:Med project: e:AtheroSysMed - Systems medicine of myocardial infarction and stroke, as well as a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development.

#### Acknowledgements Lothian Birth Cohorts

We thank the cohort participants and team members who contributed to these studies. This work was supported by multiple sources. Phenotype collection in the Lothian Birth Cohort 1921 was supported by the UK's Biotechnology and Biological Sciences Research Council (BBSRC), The Royal Society and The Chief Scientist Office of the Scottish Government. Phenotype collection in the Lothian Birth Cohort 1936 was supported by Age UK (The Disconnected Mind project). Methylation typing was supported by Centre for Cognitive Ageing and Cognitive Epidemiology (Pilot Fund award), Age UK, The Wellcome Trust Institutional Strategic Support Fund, The University of Edinburgh, and The University of Queensland. REM, IJD and PMV are members of the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE). CCACE is supported by funding from the BBSRC, the Medical Research Council (MRC), and the University of Edinburgh as part of the cross-council Lifelong Health and Wellbeing initiative (MR/K026992/1).

#### Acknowledgement Framingham Heart Study

The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195 and HHSN268201500001I. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, National Institutes of Health, Bethesda, MD. JMM and KLL were supported by R01AG029451. This work utilized the computational resources of the NIH HPC Biowulf cluster. (<http://hpc.nih.gov>). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services.

#### Acknowledgement WHI

The generation of the WHI was supported by NIH/NHLBI 60442456 BAA23 (Assimes, Absher, Horvath). SH, ML, ATL were supported by NIH/NIA 5R01AG042511-02 (Horvath, Levine) and NIH/NIA 1U34AG051425-01 (Horvath). The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C.

We would like to acknowledge the following WHI investigators. Program Office: (National Heart, Lung, and Blood Institute, Bethesda, Maryland) Jacques Rossouw, Shari Ludlam, Dale Burwen, Joan McGowan, Leslie Ford, and Nancy Geller. Clinical Coordinating Center: Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, WA) Garnet Anderson, Ross Prentice, Andrea LaCroix, and Charles Kooperberg Investigators and Academic Centers: (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) Barbara V. Howard; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Arizona, Tucson/Phoenix, AZ) Cynthia A. Thomson; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker. Women's Health Initiative Memory Study: (Wake

Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker.

### Acknowledgement TwinsUK

We would like to thank all the twins and family members in the TwinsUK cohort. Support for this work was obtained from the European Research Council (ERC 250157) and in part from the TwinsUK resource, which is funded by the Wellcome Trust; the European Community's Seventh Framework Programme (FP7/2007–2013); and the National Institute for Health Research (NIHR) BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

### **AUTHOR CONTRIBUTIONS**

All co-authors contributed data or analyzed the data or helped interpret the findings or helped write the article. BHC coordinated the analysis of the cohort studies and created the software code. SH, D. Levy, A. Baccarelli, J van Meurs, J. Bell, A. Peters, I. Deary, J. Pankow, L. Ferrucci, led cohort studies of this effort. SH developed and implemented the novel measures of epigenetic age acceleration and wrote the first draft of the article.

### **FUNDING**

This work was supported by NIH/NIA 1U34AG051425-01 and by grants mentioned in the acknowledgement section. The United States Department of Veterans Affairs (VA) Normative Aging Study (NAS) is supported by the Cooperative Studies Program/ERIC and is a research component of the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), Boston Massachusetts.

### **CONFLICTS OF INTEREST**

The Regents of the University of California is the sole owner of a provisional patent application directed at the invention of measures of epigenetic age acceleration for which SH is a named inventor. The other authors declare no conflicts of interest.

### **REFERENCES**

1. Bocklandt S, Lin W, Sehl ME, Sánchez FJ, Sinsheimer JS, Horvath S, Vilain E. Epigenetic predictor of age. *PLoS One*. 2011; 6:e14821. doi.org/10.1371/journal.pone.0014821
2. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013; 49:359–67. doi.org/10.1016/j.molcel.2012.10.016
3. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013; 14:R115. doi.org/10.1186/gb-2013-14-10-r115
4. Lin Q, Weidner CI, Costa IG, Marioni RE, Ferreira MR, Deary IJ, Wagner W. DNA methylation levels at individual age-associated CpG sites can be indicative for life expectancy. *Aging (Albany NY)*. 2016; 8:394–401. doi.org/10.18632/aging.100908
5. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, et al. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol*. 2015; 16:25. doi.org/10.1186/s13059-015-0584-6
6. Christiansen L, Lenart A, Tan Q, Vaupel JW, Aviv A, McGue M, Christensen K. DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging Cell*. 2016; 15:149–54. doi.org/10.1111/acel.12421
7. Perna L, Zhang Y, Mons U, Holleczeck B, Saum K-U, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics*. 2016; 8:64. doi.org/10.1186/s13148-016-0228-z
8. Horvath S, Pirazzini C, Bacalini MG, Gentilini D, Di Blasio AM, Delledonne M, Mari D, Arosio B, Monti D, Passarino G, De Rango F, D'Aquila P, Giuliani C, et al. Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. *Aging (Albany NY)*. 2015; 7:1159–70. doi.org/10.18632/aging.100861
9. Fagnoni FF, Vescovini R, Passeri G, Bologna G, Pedrazzoni M, Lavagetto G, Casti A, Franceschi C, Passeri M, Sansoni P. Shortage of circulating naive CD8(+) T cells provides new insights on immunodeficiency in aging. *Blood*. 2000; 95:2860–68.
10. Franceschi C. Inflammaging as a major characteristic of old people: can it be prevented or cured? *Nutr Rev*. 2007; 65:S173–76. doi.org/10.1301/nr.2007.dec.S173-S176
11. Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, De Benedictis G. Inflamm-aging. An evolutionary perspective on immunosenescence. *Ann N Y Acad Sci*. 2000; 908:244–54. doi.org/10.1111/j.1749-6632.2000.tb06651.x
12. Miller RA. The aging immune system: primer and prospectus. *Science*. 1996; 273:70–74. doi.org/10.1126/science.273.5271.70



13. Baker GT 3rd, Sprott RL. Biomarkers of aging. *Exp Gerontol.* 1988; 23:223–39. doi.org/10.1016/0531-5565(88)90025-3
14. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, Reinmaa E, Sutphin GL, Zhernakova A, Schramm K, Wilson YA, Kobes S, Tukiainen T, et al, and NABEC/UKBEC Consortium. The transcriptional landscape of age in human peripheral blood. *Nat Commun.* 2015; 6:8570. doi.org/10.1038/ncomms9570
15. Sanders JL, Newman AB. Telomere length in epidemiology: a biomarker of aging, age-related disease, both, or neither? *Epidemiol Rev.* 2013; 35:112–31. doi.org/10.1093/epirev/mxs008
16. Sanders J, Boudreau R, Newman A. 2012. *Understanding the Aging Process Using Epidemiologic Approaches.* (Dordrecht Heidelberg New York: Springer).
17. Roberts-Thomson IC, Whittingham S, Youngchaiyud U, Mackay IR. Ageing, immune response, and mortality. *Lancet.* 1974; 2:368–70. doi.org/10.1016/S0140-6736(74)91755-3
18. Levine ME, Hosgood HD, Chen B, Absher D, Assimes T, Horvath S. DNA methylation age of blood predicts future onset of lung cancer in the women’s health initiative. *Aging (Albany NY).* 2015; 7:690–700. doi.org/10.18632/aging.100809
19. Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, Gibson J, Redmond P, Cox SR, Pattie A, Corley J, Taylor A, Murphy L, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. *Int J Epidemiol.* 2015; 44:1388–96. doi.org/10.1093/ije/dyu277
20. Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer’s disease related cognitive functioning. *Aging (Albany NY).* 2015; 7:1198–211. doi.org/10.18632/aging.100864
21. Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio AM, Giuliani C, Tung S, Vinters HV, Franceschi C. Accelerated epigenetic aging in Down syndrome. *Aging Cell.* 2015; 14:491–95. doi.org/10.1111/ace1.12325
22. Horvath S, Langfelder P, Kwak S, Aaronson J, Rosinski J, Vogt TF, Eszes M, Faull RL, Curtis MA, Waldvogel HJ, Choi OW, Tung S, Vinters HV, et al. Huntington’s disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. *Aging (Albany NY).* 2016; 8:1485–512. doi.org/10.18632/aging.101005
23. Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson’s disease patients. *Aging (Albany NY).* 2015; 7:1130–42. doi.org/10.18632/aging.100859
24. Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, Heits N, Bell JT, Tsai P-C, Spector TD, Deloukas P, Siebert R, Sipos B, et al. Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci USA.* 2014; 111:15538–43. doi.org/10.1073/pnas.1412759111
25. Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. *J Infect Dis.* 2015; 212:1563–73. doi.org/10.1093/infdis/jiv277
26. Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, Bandinelli S, Salfati E, Manson JE, Quach A, Kusters CD, Kuh D, Wong A, et al. Menopause accelerates biological aging. *Proc Natl Acad Sci USA.* 2016; 113:9327–32. doi.org/10.1073/pnas.1604558113
27. Horvath S, Mah V, Lu AT, Woo JS, Choi OW, Jasinska AJ, Riancho JA, Tung S, Coles NS, Braun J, Vinters HV, Coles LS. The cerebellum ages slowly according to the epigenetic clock. *Aging (Albany NY).* 2015; 7:294–306. doi.org/10.18632/aging.100742
28. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, Ritz BR, Chen B, Lu AT, Rickabaugh TM, Jamieson BD, Sun D, Li S, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* 2016; 17:171. doi.org/10.1186/s13059-016-1030-0
29. Lowe D, Horvath S, Raj K. Epigenetic clock analyses of cellular senescence and ageing. *Oncotarget.* 2016; 7:8524–31. doi:10.18632/oncotarget.7383.
30. Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CC, O’Donovan MC, Bray NJ, Mill J. Methyloomic trajectories across human fetal brain development. *Genome Res.* 2015; 25:338–52. doi.org/10.1101/gr.180273.114
31. Walker RF, Liu JS, Peters BA, Ritz BR, Wu T, Ophoff RA, Horvath S. Epigenetic age analysis of children who seem to evade aging. *Aging (Albany NY).* 2015; 7:334–39. doi.org/10.18632/aging.100744
32. Klemmera P, Doubal S. A new approach to the concept and computation of biological age. *Mech Ageing Dev.* 2006; 127:240–48. doi.org/10.1016/j.mad.2005.10.004
33. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate

measures of cell mixture distribution. *BMC Bioinformatics*. 2012; 13:86. doi.org/10.1186/1471-2105-13-86

34. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *J Stat Softw*. 2010; 36:1–48. doi.org/10.18637/jss.v036.i03.
35. Curb JD, McTiernan A, Heckbert SR, Kooperberg C, Stanford J, Nevitt M, Johnson KC, Proulx-Burns L, Pastore L, Criqui M, Daugherty S, and WHI Morbidity and Mortality Committee. Outcomes ascertainment and adjudication methods in the Women’s Health Initiative. *Ann Epidemiol*. 2003 (Suppl ); 13:S122–28. doi.org/10.1016/S1047-2797(03)00048-6
36. Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: the Lothian Birth Cohorts of 1921 and 1936. *Int J Epidemiol*. 2012; 41:1576–84. doi.org/10.1093/ije/dyr197
37. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, Guralnik JM. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J Am Geriatr Soc*. 2000; 48:1618–25. doi.org/10.1111/j.1532-5415.2000.tb03873.x