



A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge

Dayoon Kwon · Daniel W. Belsky

Received: 4 September 2021 / Accepted: 22 October 2021 / Published online: 2 November 2021
© The Author(s), under exclusive licence to American Aging Association 2021

Abstract Methods to quantify biological aging are emerging as new measurement tools for epidemiology and population science and have been proposed as surrogate measures for healthy lifespan extension in geroscience clinical trials. Publicly available software packages to compute biological aging measurements from DNA methylation data have accelerated dissemination of these measures and generated rapid gains in knowledge about how different measures perform in a range of datasets. Biological age measures derived from blood chemistry data were introduced at the same time as the DNA methylation measures and, in multiple studies, demonstrate superior performance to these measures in prediction of healthy lifespan. However, their dissemination has been

slow by comparison, resulting in a significant gap in knowledge. We developed a software package to help address this knowledge gap. The BioAge R package, available for download at GitHub (<http://github.com/dayoonkwon/BioAge>), implements three published methods to quantify biological aging based on analysis of chronological age and mortality risk: Klemera-Doubal biological age, PhenoAge, and homeostatic dysregulation. The package allows users to parametrize measurement algorithms using custom sets of biomarkers, to compare the resulting measurements to published versions of the Klemera-Doubal method and PhenoAge algorithms, and to score the measurements in new datasets. We applied BioAge to safety lab data from the CALERIE™ randomized controlled trial, the first-ever human trial of long-term calorie restriction in healthy, non-obese adults, to test effects of intervention on biological aging. Results contribute evidence that CALERIE intervention slowed biological aging. BioAge is a toolkit to facilitate measurement of biological age for geroscience.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11357-021-00480-5>.

D. Kwon
Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles, CA 90095, USA

D. Kwon · D. W. Belsky (✉)
Robert N. Butler Columbia Aging Center, Mailman School of Public Health, Columbia University, New York, NY 10032, USA
e-mail: db3275@cumc.columbia.edu

D. W. Belsky
Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032, USA

Keywords Biological age · Geroscience · CALERIE · Biomarkers · Aging · Healthspan

Introduction

Biological aging is the gradual and progressive decline in system integrity that occurs with advancing chronological age [1]. Processes of biological aging

begin with accumulation of cellular level changes that increase vulnerability of tissues and organs to loss of function, ultimately causing disease, disability, and death [2, 3]. Experiments with animals show these processes can be modified, extending healthy lifespan for worms, flies, and mice [4, 5]. The emerging field of geroscience is focused on translating these therapies to extend healthy lifespan in humans [6, 7]. Key to these translational efforts is the development of biomarkers that can detect effects of treatments that slow or reverse biological aging.

Development of biomarkers of aging has a long history and remains a work in progress [8]. For geroscience, aging biomarkers are needed because it will take decades to establish whether treatments extend healthy lifespan in humans [9, 10]. Biomarker measurements, by contrast, have potential to enable early tests of treatment effectiveness over timescales of months or years.

There is still no gold standard biomarker of aging. Among those showing most promise are a family of algorithms applied to DNA methylation data that estimate a person's biological age or mortality risk [11–18]. These biomarkers first emerged early in the last decade and since then have undergone rapid refinement, increasing their reliability and predictive power [19]. Many clinical and cohort studies are now conducting DNA methylation analysis of stored biospecimens. Using publicly available tools, e.g., <https://horvath.genetics.ucla.edu/html/dnamage/>, research teams around the world are using these datasets to compute DNA methylation measures of aging and advance the science.

A second set of promising aging biomarkers consists of algorithms derived from blood chemistries and other clinical data. Although these algorithms based on clinical parameters are as or more predictive of disease, disability, and mortality as compared to DNA methylation measures [20–22] and show evidence of sensitivity to a range of causes hypothesized to accelerate aging [23–26], they have received much less research attention. One barrier to wider integration of these clinical data algorithms into aging research is a lack of software for computing these measures in new datasets. A further barrier is that many studies will include several but not all of the clinical markers included in a particular algorithm. Unlike DNA methylation datasets, which are generated from a single multiplex array and include the

same measurements across studies, datasets of clinical markers are assembled from multiple assays. As a result, a study may be missing one or another of the markers included in an algorithm.

To address these barriers, we present a novel R package, “BioAge,” which is currently programmed to implement three methods to quantify biological aging: Klemara-Dougal method (KDM) biological age [27], PhenoAge [13], and homeostatic dysregulation [28]. The package has two sets of functions. One set of functions allows the user to develop new versions of the KDM biological age, PhenoAge, and homeostatic dysregulation measures using biomarker sets of their own choosing and data from the US Health and Nutrition Examination Surveys (NHANES). This set of functions (1) trains new algorithms using data from NHANES III and (2) compares the new algorithms to published versions using data from NHANES IV. The second set of functions allows the user to train new versions of the algorithms and then apply them to test data of the user's choosing. Together, these functions enable users to develop new versions of published algorithms that are customized to the biomarkers available within their own datasets.

BioAge was designed as an easy-to-use package that only requires a set of available clinical markers as inputs, enabling researchers to compute the biological aging measures and evaluate the performance of the biological age algorithms. Here, we provide an example of BioAge implementation using data from a randomized controlled trial, CALERIE. The CALERIE trial tested the effects of 2 years of caloric restriction in a sample of healthy, non-obese adults. We use the BioAge package to compute measures of biological aging at pre-intervention baseline and at 12- and 24-month follow-up assessments. We then use the computed measures to evaluate the effect of CALERIE intervention on biological aging.

Methods

The BioAge package develops algorithms to measure biological aging from a user-specified list of biomarkers based on three published methods: the Klemara-Dougal method biological age (KDM BA) [27], the PhenoAge [13], and homeostatic dysregulation (HD) [28]. The package utilizes datasets for training and

testing algorithms from the US Health and Nutrition Examination Surveys (<https://www.cdc.gov/nchs/nhanes/Default.aspx>). The package is available on GitHub (<http://github.com/dayoonkwon/BioAge>) and is licensed under the GNU General Public License v3.0.

The package contains two sets of functions. The first set of functions (i) applies published methods to generate biological age algorithms based on a user-specified list of biomarkers using the National Health and Nutrition Examination Surveys (NHANES) III dataset as a training sample and (ii) compares these new algorithms to one another and to published algorithms using NHANES IV dataset as a test sample. These functions are labeled with the suffix “_nhanes” (kdm_nhanes, phenoage_nhanes, hd_nhanes). The second set of functions applies published methods to generate biological age algorithms based on a user-specified list of biomarkers using a user-specified training dataset and projects these new algorithms onto a user-specified test dataset. These functions make it possible to train new algorithms using the NHANES datasets and then to project these algorithms onto new test datasets. These functions are labeled with the suffix “_calc” (kdm_calc, phenoage_calc, hd_calc).

The following sections introduce the three methods of calculating biological age, the NHANES data, and the comparative analysis performed in the _nhanes functions.

Biological aging measures

The BioAge package calculates biological aging measures using three methods: KDM BA, PhenoAge, and HD [13, 27, 28]. These biological aging measures are patient-level measures that combine information from multiple clinical biomarkers to quantify aging-related deficits in system integrity [27–29]. We selected these three methods based on previous literature and published evidence for links with morbidity, mortality, and indicators of healthspan in young and older populations [13, 27, 29–33].

KDM biological age

An individual’s KDM BA prediction corresponds to the chronological age at which her/his physiology

would be approximately normal. KDM BA older than chronological age indicates an advanced state of biological aging and increased risk for disease, disability, and mortality. KDM BA younger than biological age indicates delayed biological aging and reduced risk for disease, disability, and mortality.

The KDM BA algorithm [27] is derived from a series of regressions of individual biomarkers on chronological age in a reference population. The equation takes information from n number of regression lines of chronological age regressed on n biomarkers:

$$BA_{EC} = \frac{\sum_{i=1}^n (x_i - q_i) \frac{k_i}{s_i^2} + \frac{CA}{s_{BA}^2}}{\sum_{i=1}^n \left(\frac{k_i}{s_i}\right)^2 + \frac{1}{s_{BA}^2}} \quad (1.1)$$

where x is the value of biomarker i measured for an individual. For each biomarker i , the parameters k , q , and s are estimated from a regression of chronological age on the biomarker in the reference sample. k , q , and s are the regression intercept, slope, and root mean squared error, respectively. s_{BA} is a scaling factor equal to the square root of the variance in chronological age explained by the biomarker set in the reference sample. CA is chronological age. In the kdm_nhanes function in BioAge package, the reference sample is NHANES III nonpregnant participants aged 30–75 years. Algorithm parameters are estimated separately for men and women.

PhenoAge

An individual’s PhenoAge prediction corresponds to the chronological age at which their mortality risk would be approximately normal in a reference population. PhenoAge older than chronological age indicates an advanced state of biological aging and increased risk for disease, disability, and mortality. PhenoAge younger than biological age indicates delayed biological aging and reduced risk for disease, disability, and mortality.

The PhenoAge algorithm is derived from multivariate analysis of mortality hazards [32, 34]. The original PhenoAge algorithm was constructed from elastic-net Gompertz regression of mortality on 42 biomarkers in the NHANES III [13]. This analysis selected nine biomarkers and chronological age as a

parsimonious model. This model is used to compute a mortality prediction score (referred to as the “Mortality Score”; see Eq. 2.1).

$$\text{Mortality score}_i = \text{CDF}(t = 120, x_i) = 1 - e^{\frac{-(e^{\gamma t} - 1)e^{xb}}{\gamma}} \quad (2.1)$$

where xb represents the linear combination of biomarkers from the fitted model. γ is an ancillary parameter to be estimated from the data and t denotes time (here in units of months). Thus, $\text{CDF}(t = 120, x_i)$ denotes the probability that the i th individual will die within the next 120 months.

The mortality prediction score is then converted into a biological age value by matching the elastic-net model predicted score with mortality scores from a univariate Gompertz regression including only chronological age as a predictor:

$$\text{CDF.univariate}(t = 120, \text{age}_i) = 1 - e^{\frac{-(e^{\gamma t} - 1)(e^{XB_{\text{age}}})}{\gamma}} \quad (2.2)$$

estimates the probability that the i th individual will die within the next 120 months as follows $\text{CDF.univariate}(120, \text{age}_i)$ where age_i is the chronological age of the i th individual.

Finally, solving the equation

$$\text{CDF}(120, x_i) = \text{CDF.univariate}(120, \text{age}_i) \quad (2.3)$$

for the variable age_i provides the PhenoAge value. In other words, the PhenoAge is the chronological age at which a participant’s predicted mortality score is equal to the average mortality hazard in the NHANES reference sample. In the “phenoage_nhanes” function, a set of biomarkers specified by the users is used to form the mortality prediction score in place of the original elastic-net model.

Homeostatic dysregulation

An individual’s HD value corresponds to how different their physiology is from a healthy reference. Higher values of HD indicate an advanced state of biological aging and increased risk for disease, disability, and mortality. Lower values of homeostatic dysregulation indicate delayed biological aging and reduced risk for disease, disability, and mortality.

HD is computed as the Mahalanobis distance [35, 36] for a set of biomarkers relative to a reference

sample. The Mahalanobis distance equation [35, 36] takes the form:

$$D_M(\vec{x}) = \sqrt{(\vec{x} - \vec{\mu})^T S^{-1} (\vec{x} - \vec{\mu})} \quad (3.1)$$

where x is a multivariate observation (all the biomarker values for an individual) and μ is the equivalent length vector of reference sample means for each variable. S is the reference sample variance–covariance matrix for the variables. If all variables are uncorrelated, then this is equivalent to scaling each biomarker by its variance and then summing the squared deviance for an observation:

$$D_M(\vec{x}) = \sqrt{\sum_{i=1}^n \frac{(x_i - \mu_i)^2}{\sigma^2(x_i)}} \quad (3.2)$$

where n is the number of biomarkers and $\sigma^2(x_i)$ is the variance in the i th biomarker. In the `hd_nhanes` function, we specify the reference sample to be NHANES III nonpregnant participants aged 20–30 years for whom all user-selected biomarkers fall within the clinically normal range. For analysis, all biomarkers are standardized to have mean=0, SD=1 separately for men and women based on this reference sample. This approach computes homeostatic dysregulation relative to a young, healthy sample, following the approach we have used previously [24, 31, 37].

NHANES

NHANES is an ongoing nationally representative, cross-sectional survey conducted by the US Centers for Disease Control and Prevention. NHANES administers questionnaires during in-home interviews and conducts health examinations, including blood draws, in a mobile examination center. Details of recruitment procedures and study design are available from the Center for Disease Control and Prevention (<https://www.cdc.gov/nchs/nhanes/index.htm>). We compiled demographic, socioeconomic, and functional performance; biomarker; and mortality data from adults aged 20–90 years participating in the NHANES III (1988–1994) and IV (1999–2018).

For analysis, we excluded biomarker outliers by computing sex-specific mean of standard deviations and dropping values more than five standard deviations from the sex-specific mean. Biomarkers with

skewed distributions were log-transformed. Details on biomarker measurements are available from the NHANES website (<https://www.cdc.gov/nchs/nhanes/index.htm>).

For several biomarkers, measurement methods changed during the 1999–2016 interval when NHANES IV data used by the package were collected. The package uses data normalized to account for these changes in methodology: Creatinine values from the 1999–2000 and 2005–2006 NHANES were corrected according to the analytical notes posted by NHANES (<https://www.cdc.gov/nchs/nhanes/index.htm>). High sensitivity C-reactive protein assays from the 2015–2016 NHANES were posted in units mg/L and were divided by 10 to match units in previous waves. Bone alkaline phosphatase values from the 1999–2000 NHANES were adjusted according to published equations. Measurement methods for plasma fasting glucose changed in the 2005–2006, 2007–2008, and 2015–2016 NHANES. Values were adjusted to be comparable across years using multiple regression equations (<https://www.cdc.gov/nchs/nhanes/index.htm>). Measurement methods for insulin changed in the 1999–2000, 2003–2004, 2005–2006, 2011–2012, and 2013–2014 NHANES. Values were adjusted to be comparable across years using multiple regression equations (<https://www.cdc.gov/nchs/nhanes/index.htm>).

In the BioAge package, data from NHANES III and NHANES IV are loaded as the datasets NHANES3 and NHANES4. The NHANES4 dataset also contains computed values of KDM biological age and PhenoAge based on the original versions of those algorithms published by Levine and colleagues [13, 38].

Comparison of biological aging measures

The package's `_nhanes` functions conduct a series of analyses to compare KDM BA, PhenoAge, and HD algorithms generated with user-specified sets of biomarkers with one another and with versions of KDM BA and PhenoAge algorithms published previously. All algorithms are trained in NHANES III data. Comparative analysis is conducted using NHANES IV data. Thus, training and test samples are distinct from one another. Analyses proceed in four steps. First, biological aging measures are

correlated with chronological age. Second, biological aging measures are correlated with one another. Third, biological aging measures are tested for association with healthspan-related characteristics: mortality, disability, physical function, and self-rated health. Finally, a set of analyses tests socio-economic patterning of biological aging algorithms.

Below, the measures included in the validation analysis are described briefly, followed by details of the analysis. Complete details on all measurements are available from the NHANES website (<https://www.cdc.gov/nchs/nhanes/index.htm>).

Mortality

NHANES III and NHANES 1999–2014 are independent cohorts with different lengths of follow-up for mortality. Participants' survival status and cause of death were determined through probabilistic matching to the death certificates from the National Death Index recorded through 2015 [39]. For analysis, we used information on aging-related mortality from diseases of the heart, malignant neoplasms, chronic lower respiratory disease, cerebrovascular disease, Alzheimer's disease, diabetes mellitus, nephritis, nephrotic syndrome, and nephrosis.

Disability, physical functioning, and self-rated health

We analyzed associations of biological aging measures with counts of limitations to activities of daily living (ADLs), walk speed, grip strength, and self-rated health. ADLs were measured as a count of functional impairments across 19 activities. Walk speed was measured in NHANES 1999–2002 from participants aged 50 and older. Measures were taken as time in seconds to walk 20 ft. Values were log-transformed for analysis. Grip strength was measured in NHANES 2011–2014 using a hand dynamometer. Values were averaged across three trials. We analyzed values for the dominant hand. To account for differences in distributions between men and women, values were transformed to have $M=0$, $SD=1$ within sex. Self-rated general health was assessed from a survey item with five response categories ranging from excellent to poor.

Socioeconomic circumstances

Socioeconomic circumstances measures included education, annual family income, and poverty income ratio. Education was categorized into five categories: less than 9th grade, 9–11th grade, high school graduate/GED, some college, and college graduate or higher. Annual family income was self-reported as one of 11 income categories ranging from less than \$5,000 to over \$75,000. Poverty income ratio was calculated based on family income and the poverty threshold based on family size.

Analysis of validation measures

For analysis, KDM BA and PhenoAge values were differenced from chronological age values and then standardized to have $M=0$, $SD=1$ separately for men and women. HD values were log-transformed and then standardized to have $M=0$, $SD=1$ separately for men and women. The package tests associations of biological aging measures with mortality (survival time) using Cox proportional hazards models to estimate hazard ratios (HR). The package tests associations of biological aging measures with counts of ADLs, log-walk-test time, grip strength, and self-rated health using linear regression models. For linear regression analysis, outcome variables are standardized to have $M=0$, $SD=1$. Coefficients reported are interpretable on a Pearson's r scale. The package tests associations of socioeconomic circumstances measures with biological aging measures using linear regression. In these models, measures of socioeconomic circumstances are specified as independent variables, and biological aging measures are specified as dependent variables. Socioeconomic circumstance measures are standardized to $M=0$, $SD=1$ for analysis so that effect sizes are denominated in terms of a 1 SD unit improvement in socioeconomic circumstances. Analysis includes NHANES IV participants for whom biological aging measures can be calculated and for whom any validation data are available. Models are reported for the full analysis sample with covariate adjustment for chronological age and sex and for samples stratified by sex, race, and chronological age (under/over 65 years).

Analysis

We compared original KDM BA and PhenoAge algorithms with algorithms composed with the new biomarker set in the NHANES IV dataset. First, two sets of “plot_” functions create comparative scatter plots using Pearson correlations. “plot_ba” function tests associations of chronological age with biological aging measures. “plot_baa” function tests associations among biological aging measures. In this function, KDM BA and PhenoAge were computed as the difference between biological age and chronological age. These biological age advancement (BAA) values were then standardized to have $mean=0$, $SD=1$ separately for men and women within the analysis sample.

Three sets of “table_” functions create regression tables with full sample and are stratified by gender, race, and age groups. “table_surv” function tests associations of biological aging measures with mortality. This function uses Cox proportional hazard models to estimate hazard ratios. “table_health” function tests associations of biological aging measures with functional performance measures. Biological aging measures were independent variables. Functional performance measures were dependent variables and standardized to have $mean=0$, $SD=1$ for analysis. “table_ses” function tests association of socioeconomic circumstance measures with biological aging measures. Socioeconomic circumstance measures were independent variables and standardized to have $mean=0$, $SD=1$ for analysis. Biological aging measures were dependent variables. In table_health and table_ses functions, linear regression is used to compute standardized beta coefficients (interpretable as Pearson's r).

Results

Part 1. Parameterization of KDM, PhenoAge, and HD biological aging measures for the CALERIE trial dataset using NHANES III

We previously analyzed KDM and HD biological aging measures in CALERIE using a biomarker set based on the original KDM algorithm published by Levine [29, 31]. We used the _nhanes functions of the BioAge package to train new KDM, PhenoAge, and

HD algorithms in the NHANES III data and then used the _calc functions to project the algorithms onto the CALERIE data. We conducted three sets of analyses: (i) using the same biomarker set included in our original CALERIE analysis (hereafter, “CALERIE Original”); (ii) using a biomarker set based on the original PhenoAge algorithm published by Levine [13] (hereafter “V1”); and (iii) using a biomarker set composed of those included in the CALERIE Original and V1 sets (hereafter “V2”). Biomarkers included in the three sets of measures and their correlations with chronological age in the NHANES IV test data and the CALERIE sample at baseline are reported in Table S2.1.

Part 2: Validation of new KDM, PhenoAge, and HD algorithms in NHANES IV data

The several biological aging measures were correlated with one another in the NHANES IV test data (Fig. 1). To analyze correlations among measures, we first differenced KDM and PhenoAge measures from chronological age to calculate age advancement values. Correlations among the biological aging measures in NHANES IV ranged from 0.39 to 0.96 (Fig. 1, Fig. S1.2).

The several biological aging measures were associated with mortality risk in the NHANES IV test sample. Compared to the KDM and PhenoAge algorithms first proposed by Levine, the CALERIE Original and V1 algorithms were somewhat less predictive of mortality (Table S1.1) and more weakly correlated with self-rated health and measures of physical functioning (Table S1.2); the V2 algorithms, which included biomarkers from both the CALERIE Original and V1 sets, performed better, although still not as well as the original PhenoAge. Similar results were found when comparing three algorithms to mortality, self-rated health and physical functioning, and socioeconomic factors across different gender, race, and age groups. For example, V2 PhenoAge produced the most robust results for individual Cox proportional hazard models ($HR=1.43$ 95% CI=[1.37–1.49], Table S1.1). Complete results from analysis of the NHANES IV data produced by the plot_ and table_ functions of the BioAge package are compiled in SI Section I. These include analysis of association with chronological age; intercorrelation with other measures of aging, including the original KDM and PhenoAge

algorithms proposed by Levine [13, 29]; and additional association analysis of self-rated health, and physical functioning, and socioeconomic factors. The V2 algorithm consistently performed better in these criterion validity analyses relative to the CALERIE Original and V1 algorithms. Therefore, we retained the V2 algorithm for CALERIE analysis. (For completeness of documentation, results for CALERIE analysis of the CALERIE Original and V1 algorithms are reported in SI Section II.)

Part 3. CALERIE Analysis

The CALERIE trial randomized $n=220$ non-obese midlife adults to 2 years of 25% caloric restriction (CR) or ad libitum (AL) diet, the control condition. Most participants in the CR intervention group did not achieve the prescribed dose of caloric restriction; the average percent CR over the 2 years was about half the prescribed dose [40]. Nevertheless, participants in the CR group lost significant weight over the first 12 months of the trial and maintained this weight loss over the second 12-month interval. They also experienced a range of physiological changes indicating improved cardiometabolic health [41]. We previously reported that the CR group demonstrated slower biological aging as compared to the AL group based on versions of the KDM BA and HD algorithms trained in data from NHANES 2007–11 [31]. PhenoAge has not yet been analyzed in CALERIE. Characteristics of the CALERIE participants included in analysis and values of the biological aging measures in the CR intervention and AL control groups are reported in Table 1.

CALERIE participants’ chronological age and V2 versions of the KDM BA and PhenoAge measures were correlated (Pearson r range=0.81–0.90, Fig. 2). CALERIE participants’ V2 HD values were not correlated with their chronological ages. At baseline, there was little difference in biological aging measures between CR and AL groups. Participants’ biological aging measures were slightly younger than their chronological ages (chronological age mean=39; KDM BA mean=35; PhenoAge mean=34). This difference may reflect the sampling frames used for CALERIE and the NHANES and volunteer bias and that CALERIE participants were selected to be in good health, whereas the NHANES sample represented the general US population. CALERIE sample baseline summary

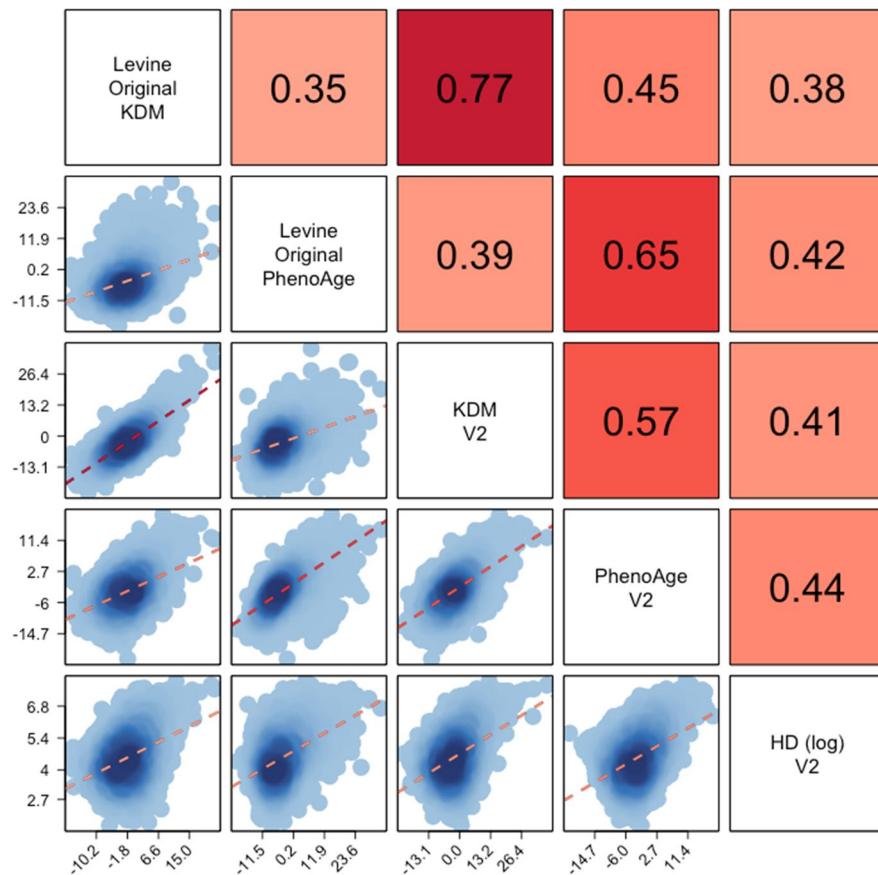


Fig. 1 Correlations of published versions of Klemera-Doubal method (KDM) biological age and PhenoAge with versions of KDM biological age, PhenoAge, and a measure computed using the homeostatic dysregulation method based on a modified set of biomarkers. The figure plots data from NHANES IV generated with the `_nhanes` commands within the BioAge package. All measures were developed from analysis of NHANES III and computed using data from NHANES IV. KDM biological age and PhenoAge values were differenced from chronological age for analysis. The “Levine Original” KDM algorithm was composed from chronological age and 10 biomarkers: albumin, alkaline phosphatase, blood urea nitrogen, creatinine, C-reactive protein, cytomegalovirus opti-

cal density, HbA1C, total cholesterol, systolic blood pressure, and forced expiratory volume in 1 s (FEV_1). The “Levine Original” PhenoAge was composed from chronological age and 9 biomarkers: albumin, alkaline phosphatase, creatinine, C-reactive protein, fasting glucose, white blood cell count, lymphocyte percentage, mean cell volume, and red cell distribution width. The “V2” versions of the KDM, PhenoAge, and HD algorithms included chronological age and 12 biomarkers: albumin, alkaline phosphatase, blood urea nitrogen, creatinine, C-reactive protein, HbA1C, total cholesterol, uric acid, white blood cell count, lymphocyte percentage, mean cell volume, and red cell distribution width

statistics for biological aging measures are reported in Table S2.2. Intercorrelations of biological aging measures are graphed in Fig. S2.2.

We tested the hypothesis that CALERIE interventions slowed biological aging using mixed-effects growth models, including participant-level random intercepts and slopes. The model took the form

$$\Delta BioAge_{it} = \gamma_1 Time_{it} + \gamma_2 CR_i + \gamma_3 Time \times CR_{it} + \chi Covariates + \mu_{0i} + \mu_{1i} Time_{it} + \epsilon_{it} \quad (4)$$

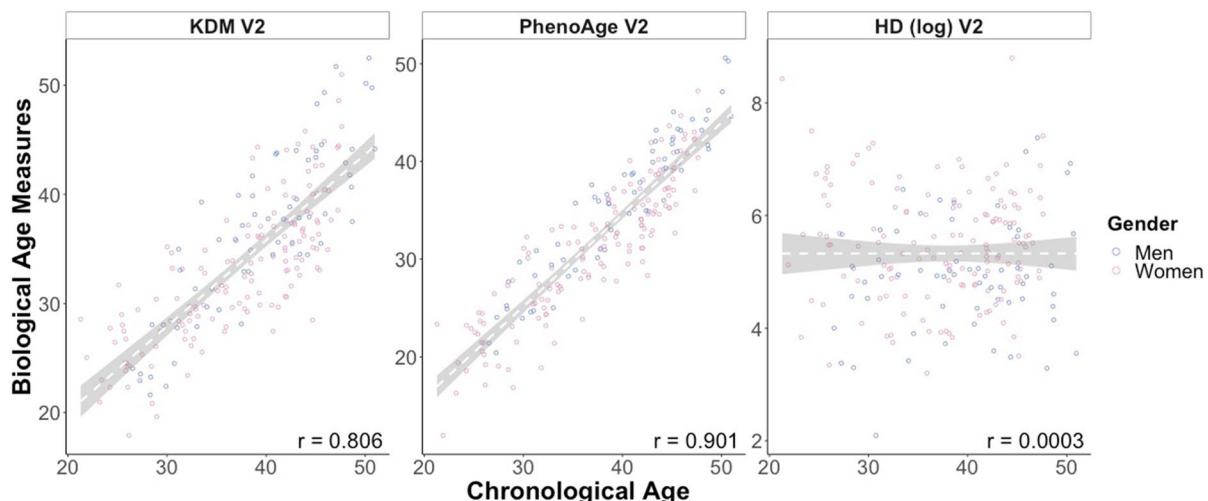
where $\Delta BioAge_{it}$ is change in the BA measure from baseline for individual “ i ” at time “ t ,” γ_1 estimates annual change in BA for ad libitum arm participants, γ_2 estimates any baseline difference BA between participants in AL and CR arms of the trial, γ_3 estimates the difference in annual change in BA

Table 1 Characteristics of CALERIE trial participants included in analysis

| | Ad libitum (n=68) | | Caloric restriction (n=139) | |
|---|----------------------|------|-----------------------------------|------|
| | %/M | SD | %/M | SD |
| Male | 26.47% | | 30.94% | |
| Age | 38.14 | 6.66 | 37.84 | 7.35 |
| V2 versions of biological age measures at pre-intervention baseline | | | | |
| KDM | 34.32 | 7.05 | 34.10 | 6.94 |
| PhenoAge | 32.32 | 7.16 | 32.80 | 7.57 |
| HD (log) | 5.27 | 0.97 | 5.32 | 1.07 |

M mean, SD standard deviation

reversed biological aging as compared to AL arm participants as measured by V2 KDM BA and PhenoAge. The average change from baseline in biological aging measures is plotted for AL and CR participants in Fig. 3. CALERIE participants randomized to the AL control group experienced an increase in V2 KDM BA of 0.69 “years” per 12-month follow-up (95% CI [0.21, 1.16]). In contrast, for participants randomized to CR, KDM BA decreased ($b = -0.2$ 95% CI [-0.55, 0.14] “years” per 12-month follow-up interval). For PhenoAge, both AL and CR groups experienced an increase over time, but the increase was slower for the CR group (AL $b = 0.83$ [0.46–1.21], CR $b = 0.17$ [-0.11, 0.44]; p value for test of difference = 0.006). HD was unchanged across follow-up for both AL and CR group participants.

**Fig. 2** Associations of Klemara-Doubal method (KDM) biological age, PhenoAge, and homeostatic dysregulation (HD) measures of biological age with chronological age among participants in the CALERIE trial at pre-intervention baseline.

The figure plots pre-intervention baseline values of the three biological aging measures against chronological age for men (blue) and women (pink) participating in the CALERIE trial ($n=207$)

between AL arm and CR arm participants, χ is a vector of covariates, and μ_{0i} and μ_{1i} are the random intercepts and slopes estimated for each individual “ i .” The coefficient γ_3 tests the hypothesis that biological aging was slowed for participants randomized to the CR arm of the trial. All models included sex and baseline age as covariates. The mixed-effects regression analysis included 611 observations of 220 individuals.

Across follow-up, CALERIE participants randomized to the trial’s CR arm experienced slower or

Regression model results are reported in Table 2.

Discussion

Quantification of biological aging is emerging as a novel approach to investigating how exposures and interventions may influence risk for chronic disease, disability, and mortality [8]. Because the aging process is ongoing from at least reproductive maturity

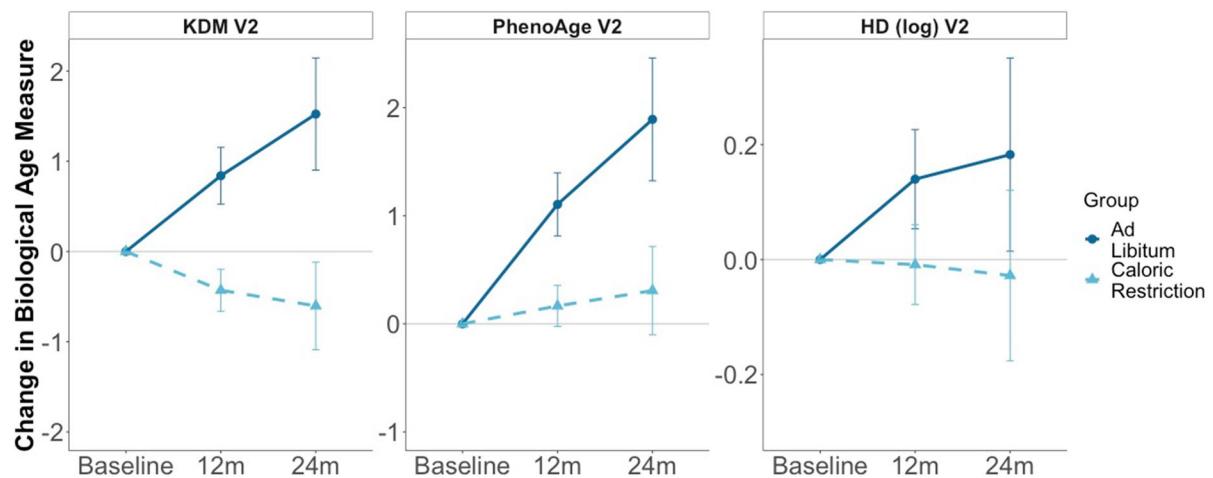


Fig. 3 Change in Klemara-Doubal method (KDM) biological age, PhenoAge, and homeostatic dysregulation (HD) from baseline to 12- and 24-month follow-ups in the ad libitum (dark blue dots) and caloric restriction (light blue triangles) groups of the CALERIE trial. The figure plots predicted values and 95% confidence intervals estimated from mixed-effects

growth models for participants in the ad libitum control group (dark blue circles, solid line) and caloric restriction intervention group (light blue triangles, dashed line). Values of KDM biological age and PhenoAge are denominated in years. Values of HD are denominated in log Mahalanobis distance units

Table 2 Estimated annual change in KDM, PhenoAge, and HD biological aging measures from baseline through 24-month follow-up in ad libitum and caloric restriction groups in the CALERIE randomized trial

| Biomarker set | | b | [95% CI] | p value | n |
|---------------|---------------------|-------|----------------|---------|-----|
| V2 | | | | | |
| KDM | Ad libitum | 0.69 | [0.21, 1.16] | 0.005 | 68 |
| | Caloric restriction | -0.2 | [-0.55, 0.14] | 0.248 | 139 |
| | Interaction | -0.89 | [-1.47, -0.31] | 0.003 | 207 |
| PhenoAge | Ad libitum | 0.83 | [0.46, 1.21] | <0.01 | 68 |
| | Caloric restriction | 0.17 | [-0.11, 0.44] | 0.232 | 139 |
| | Interaction | -0.67 | [-1.13, -0.2] | 0.006 | 207 |
| HD (log) | Ad libitum | 0.06 | [-0.08, 0.2] | 0.411 | 68 |
| | Caloric restriction | -0.01 | [-0.12, 0.09] | 0.784 | 139 |
| | Interaction | -0.07 | [-0.25, 0.1] | 0.409 | 207 |

The regression model included sex and age at baseline as covariates

CI confidence interval

and may begin even earlier [42], biological aging measures have potential to detect signs of risk decades before disease processes are established [30]. Measurements of biological aging based on algorithm-defined composites of clinical parameters are relatively understudied in this growing research area. These measures are equally or more predictive of morbidity and mortality as compared to better-studied measures based on molecular data, including telomere

length and DNA methylation clocks [20–22, 24]; they are variable in apparently healthy young adults [23, 30]; and they are sensitive to risk exposures that shorten healthy lifespan and to interventions that slow aging in animals [23, 24, 31, 43, 44]. Importantly, the data needed to calculate these algorithm-based measures are routinely collected during clinical care and health research: routine blood chemistries, complete blood counts, and assessments of lung function and

blood pressure. They therefore represent an un-tapped reservoir of information about aging processes within many existing datasets. We developed an R package to aid investigators in integrating these measurements into existing datasets that address key challenges that may have slowed their adoption.

The BioAge R package is an easy to install tool that can implement the Klemara-Dougal [27], PhenoAge [34], and homeostatic dysregulation [28] methods following the approach we have used in previous work [24, 31, 43, 45]. The package has three strengths. First, it eases implementation of published biological age algorithms in biomarker datasets. Second, it allows for parameterization of new algorithms using published methods based in existing or new datasets, although the _nhanes commands are limited to the set of biomarkers included in the NHANES database. Third, when new algorithms are composed of biomarkers included in the NHANES database, it enables head-to-head comparison with the original published versions of the biological age algorithms to evaluate comparative criterion and construct validity.

We demonstrated the BioAge package by applying it to calculate biological age values from laboratory data collected in the CALERIE randomized controlled trial, the first human randomized controlled trial of long-term calorie restriction [40]. We previously used data from NHANES 2007–2010 to develop versions of the KDM and HD algorithms to test effects of CALERIE intervention on biological aging [31]. Building on that analysis, we used the BioAge package to compute new versions of the PhenoAge, KDM, and HD algorithms based on (a) the original set of markers used in our earlier paper; (b) a set of markers matched as closely as possible to the Levine PhenoAge algorithm; and (c) a combined set of measures included in the first two sets. We used NHANES III data to train these algorithms. In comparative validation analysis using data from NHANES IV (1999–2018), the algorithm with the combined set of biomarkers performed the best in analysis predicting morbidity and mortality. We applied this algorithm to the CALERIE data to test intervention effects on biological aging. Consistent with our earlier analysis, we found that CALERIE intervention slowed biological aging as measured by the KDM and PhenoAge algorithms. However, in contrast to our previous analysis, the effect of CALERIE intervention on biological aging measured by the HD

algorithm was not statistically different from zero. For the analysis reported in this article, we developed the HD algorithm using data from NHANES III, collected during 1989–1994. In our previous article, we developed the HD algorithm using data from continuous NHANES waves 2007–2010, matching the time period during which the CALERIE trial was conducted. The difference in results highlights potential sensitivity of algorithms to the reference data used in their development.

The BioAge package offers researchers flexibility to customize biological age algorithms to chosen sets of biomarkers and reference samples. Published approaches to selecting biomarkers to compose the algorithms included in the BioAge package include screening a biomarker set for the subset meeting a correlation threshold with chronological age [29] and selection by machine learning [32]. More inclusive approaches incorporate all available biomarkers with known relationship to aging-related disease or to biological systems affected by aging [30, 33]. Empirical approaches to biomarker selection have the advantage of being objective and reproducible, although they can lead to over-fitting of algorithms to training data. Ultimately, the guiding principle of biomarker selection for physiology-based biological age algorithms should be to include markers reflecting the integrity of multiple biological systems affected by aging processes. The optimal method for identifying which systems and biomarkers should be included remains a work in progress. The BioAge R package is intended as tool that can help advance this effort as well as accelerate translation of such advances for epidemiologic and clinical research.

Conclusions

Measurements of biological aging derived from clinical parameters, including routine blood chemistries, complete blood count data, and blood pressure and lung function testing, represent a powerful approach to investigating etiology of individual differences in aging and evaluating interventions to increase healthspan. The BioAge R package facilitates implementation of training and testing of three common, validated methods to compute biological age measurements from these types of data.

Author contribution DWB and DK conceived the research and designed the software and analysis. DK wrote the software, conducted the analysis, and produced the figures. DWB and DK wrote the manuscript. Both authors had access to the data.

Funding This work was supported by the National Institute on Aging grants R01AG061378 and R01AG066887 and Russel Sage Foundation BioSS grant 1810–08987. DWB is a fellow of the Canadian Institute for Advanced Research Child Brain Development Network.

Data availability All data used in this manuscript are publicly available. Data from the US National Health and Nutrition Examinations Surveys are available from the US Centers for Disease Control and Prevention at <https://www.cdc.gov/nchs/nhanes/Default.aspx>. Data from the CALERIE trial are available from the CALERIE Research Network at <https://calerie.duke.edu/samples-data-access-and-analysis>.

Code availability All the code used in these analyses is covered by GPL-3.0 License and is available from GitHub: <https://github.com/dayoonkwon/BioAge>.

Declarations

Ethics approvals Not applicable.

Consent to participant Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

- Kirkwood TB. Understanding the odd science of aging. *Cell*. 2005;120(4):437–47. <https://doi.org/10.1016/j.cell.2005.01.027>.
- Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: linking aging to chronic disease. *Cell*. 2014;159(4):709–13. <https://doi.org/10.1016/j.cell.2014.10.039>.
- Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell*. 2013;153(6):1194–217. <https://doi.org/10.1016/j.cell.2013.05.039>.
- Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E. From discoveries in ageing research to therapeutics for healthy ageing. *Nature*. 2019;571(7764):183–92. <https://doi.org/10.1038/s41586-019-1365-2>.
- Kaeberlein M, Rabinovitch PS, Martin GM. Healthy aging: the ultimate preventative medicine. *Science*. 2015;350(6265):1191–3. <https://doi.org/10.1126/science.aad3267>.
- Barzilai N, Cuervo AM, Austad S. Aging as a biological target for prevention and therapy. *Jama-J Am Med Assoc*. 2018;320(13):1321–2. <https://doi.org/10.1001/jama.2018.9562>.
- Sierra F. Special issue: moving geroscience into uncharted waters: Guest Editorial Moving Geroscience Into Uncharted Waters. *J Gerontol a-Biol*. 2016;71(11):1385–7. <https://doi.org/10.1093/gerona/glw087>.
- Ferrucci L, Gonzalez-Freire M, Fabbri E, Simonsick E, Tanaka T, Moore Z, et al. Measuring biological aging in humans: a quest. *Aging Cell*. 2020;19(2): e13080. [https://doi.org/10.1111.acel.13080](https://doi.org/10.1111/acel.13080).
- Justice J, Miller JD, Newman JC, Hashmi SK, Halter J, Austad SN, et al. Frameworks for proof-of-concept clinical trials of interventions that target fundamental aging processes. *J Gerontol a-Biol*. 2016;71(11):1415–23. <https://doi.org/10.1093/gerona/glw126>.
- Moffitt TE, Belsky DW, Danese A, Poulton R, Caspi A. The longitudinal study of aging in human young adults: knowledge gaps and research agenda. *J Gerontol A Biol Sci Med Sci*. 2017;72(2):210–5. <https://doi.org/10.1093/gerona/glw191>.
- Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013;49(2):359–67. <https://doi.org/10.1016/j.molcel.2012.10.016>.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013;14(10):R115. <https://doi.org/10.1186/gb-2013-14-10-r115>.
- Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10(4):573–91. <https://doi.org/10.18632/aging.101414>.
- Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging-US*. 2019;11(2):303–27. <https://doi.org/10.18632/aging.101684>.
- Shireby GL, Davies JP, Francis PT, Burrage J, Walker EM, Neilson GWA, et al. Recalibrating the epigenetic clock: implications for assessing biological age in the human cortex. *Brain*. 2020;143:3763–75. <https://doi.org/10.1093/brain/awaa334>.
- Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. *Genome Biology*. 2014;15(2). <https://doi.org/10.1186/gb-2014-15-2-r24>
- Youn A, Wang S. The MiAge calculator: a DNA methylation-based mitotic age calculator of human tissue types. *Epigenetics-US*. 2018;13(2):192–206. <https://doi.org/10.1080/15592294.2017.1389361>.
- Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, Montgomery GW, et al. Improved precision of epigenetic clock estimates across tissues and its implication for

- biological ageing. *Genome Med.* 2019;11(1). <https://doi.org/10.1186/s13073-019-0667-1>.
- 19. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet.* 2018;19(6):371–84. <https://doi.org/10.1038/s41576-018-0004-3>.
 - 20. Belsky DW, Moffitt TE, Cohen AA, Corcoran DL, Levine ME, Prinz JA, et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? *Am J Epidemiol.* 2018;187(6):1220–30. <https://doi.org/10.1093/aje/kwx346>.
 - 21. Li X, Ploner A, Wang Y, Magnusson PK, Reynolds C, Finkel D, et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *Elife.* 2020;9. <https://doi.org/10.7554/elife.51507>.
 - 22. Murabito JM, Zhao Q, Larson MG, Rong J, Lin H, Benjamin EJ, et al. Measures of biologic age in a community sample predict mortality and age-related disease: the Framingham Offspring Study. *J Gerontol A Biol Sci Med Sci.* 2018;73(6):757–62. <https://doi.org/10.1093/gerona/glx144>.
 - 23. Belsky DW, Caspi A, Cohen HJ, Kraus WE, Ramrakha S, Poulton R, et al. Impact of early personal-history characteristics on the pace of aging: implications for clinical trials of therapies to slow aging and extend healthspan. *Aging Cell.* 2017;16(4):644–51. <https://doi.org/10.1111/acel.12591>.
 - 24. Hastings WJ, Shalev I, Belsky DW. Comparability of biological aging measures in the National Health and Nutrition Examination Study, 1999–2002. *Psychoneuroendocrinology.* 2019;106:171–8. <https://doi.org/10.1016/j.psyneuen.2019.03.012>.
 - 25. Levine ME, Crimmins EM. Evidence of accelerated aging among African Americans and its implications for mortality. *Soc Sci Med.* 2014;118:27–32. <https://doi.org/10.1016/j.socscimed.2014.07.022>.
 - 26. Shirazi TN, Hastings WJ, Rosinger AY, Ryan CP. Parity predicts biological age acceleration in post-menopausal, but not pre-menopausal, women. *Sci Rep-Uk.* 2020;10(1). <https://doi.org/10.1038/s41598-020-77082-2>.
 - 27. Klemara P, Doubal S. A new approach to the concept and computation of biological age. *Mech Ageing Dev.* 2006;127(3):240–8. <https://doi.org/10.1016/j.mad.2005.10.004>.
 - 28. Cohen AA, Milot E, Yong J, Seplaki CL, Fulop T, Bandeen-Roche K, et al. A novel statistical approach shows evidence for multi-system physiological dysregulation during aging. *Mech Ageing Dev.* 2013;134(3–4):110–7. <https://doi.org/10.1016/j.mad.2013.01.004>.
 - 29. Levine ME. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol a-Biol.* 2013;68(6):667–74. <https://doi.org/10.1093/gerona/gls233>.
 - 30. Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, et al. Quantification of biological aging in young adults. *Proc Natl Acad Sci USA.* 2015;112(30):E4104–10. <https://doi.org/10.1073/pnas.1506264112>.
 - 31. Belsky DW, Huffman KM, Pieper CF, Shalev I, Kraus WE. Change in the rate of biological aging in response to caloric restriction: CALERIE Biobank Analysis. *J Gerontol A Biol Sci Med Sci.* 2017;73(1):4–10. <https://doi.org/10.1093/gerona/glx096>.
 - 32. Levine ME, Crimmins EM. Is 60 the new 50? Examining changes in biological age over the past two decades. *Demography.* 2018;55(2):387–402. <https://doi.org/10.1007/s13524-017-0644-5>.
 - 33. Li Q, Wang S, Milot E, Bergeron P, Ferrucci L, Fried LP, et al. Homeostatic dysregulation proceeds in parallel in multiple physiological systems. *Aging Cell.* 2015;14(6):1103–12. <https://doi.org/10.1111/acel.12402>.
 - 34. Liu Z, Kuo PL, Horvath S, Crimmins E, Ferrucci L, Levine M. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: a cohort study. *PLoS Med.* 2018;15(12): e1002718. <https://doi.org/10.1371/journal.pmed.1002718>.
 - 35. Cohen AA. Complex systems dynamics in aging: new evidence, continuing questions. *Biogerontology.* 2016;17(1):205–20. <https://doi.org/10.1007/s10522-015-9584-x>.
 - 36. Mahalanobis PC. Mahalanobis distance. *Proceedings of the National Academy of Sciences, India.* 1936;73:1370–6.
 - 37. Parker DC, Bartlett BN, Cohen HJ, Fillenbaum G, Huebner JL, Kraus VB, et al. Association of blood chemistry quantifications of biological aging with disability and mortality in older adults. *J Gerontol A Biol Sci Med Sci.* 2019. <https://doi.org/10.1093/gerona/glz219>.
 - 38. Levine ME. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? *J Gerontol A Biol Sci Med Sci.* 2013;68(6):667–74. <https://doi.org/10.1093/gerona/gls233>.
 - 39. Ingram DD, Lochner KA, Cox CS. Mortality experience of the 1986–2000 National Health Interview Survey Linked Mortality Files participants. *Vital Health Stat 2.* 2008, (147):1–37.
 - 40. Ravussin E, Redman LM, Rochon J, Das SK, Fontana L, Kraus WE, et al. A 2-year randomized controlled trial of human caloric restriction: feasibility and effects on predictors of health span and longevity. *J Gerontol A Biol Sci Med Sci.* 2015;70(9):1097–104. <https://doi.org/10.1093/gerona/glv057>.
 - 41. Kraus WE, Bhapkar M, Huffman KM, Pieper CF, Krupa Das S, Redman LM, et al. 2 years of calorie restriction and cardiometabolic risk (CALERIE): exploratory outcomes of a multicentre, phase 2, randomised controlled trial. *Lancet Diabetes Endocrinol.* 2019;7(9):673–83. [https://doi.org/10.1016/S2213-8587\(19\)30151-2](https://doi.org/10.1016/S2213-8587(19)30151-2).
 - 42. Gladyshev VN. The ground zero of organismal life and aging. *Trends Mol Med.* 2021;27(1):11–9. <https://doi.org/10.1016/j.molmed.2020.08.012>.

43. Graf G, Crowe C, Kothari M, Kwon D, Manly J, Turney I, et al. Testing DNA-methylation and blood-chemistry measures of biological aging in models of Black-White disparities in healthspan characteristics. medRxiv. 2021:2021.03.02.21252685. <https://doi.org/10.1101/2021.03.02.21252685>.
44. Liu Z, Chen X, Gill TM, Ma C, Crimmins EM, Levine ME. Associations of genetics, behaviors, and life course circumstances with a novel aging and healthspan measure: evidence from the Health and Retirement Study. PLoS Med. 2019;16(6): e1002827. <https://doi.org/10.1371/journal.pmed.1002827>.
45. Parker DC, Bartlett BN, Cohen HJ, Fillenbaum G, Huebner JL, Kraus VB, et al. Association of blood chemistry quantifications of biological aging with disability and mortality in older adults. J Gerontol A Biol Sci Med Sci. 2020;75(9):1671–9. <https://doi.org/10.1093/gerona/glz219>.

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