EBioMedicine 63 (2021) 103151

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Research paper

DNA methylation-based biomarkers of age acceleration and all-cause death, myocardial infarction, stroke, and cancer in two cohorts: The NAS, and KORA F4

Cuicui Wang[a,*], Wenli Ni[b], Yueli Yao[b], Allan Just[c], Jonathan Heiss[c], Yaguang Wei[a], Xu Gao[d], Brent A. Coull[a,e], Anna Kosheleva[a], Andrea A. Baccarelli[d], Annette Peters[b,f], Joel D. Schwartz[a]

[a] Department of Environmental Health, Harvard T.H. Chan School of Public Health, 401 Park Drive, West of Landmark Center, Boston, MA 02215, United States
[b] Institute of Epidemiology, Helmholtz Zentrum München, Neuherberg, Germany
[c] Department of Environmental Medicine, and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[d] Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY, United States
[e] Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, United States
[f] Institute of Medical Information Science, Biometry, and Epidemiology, Ludwig Maximilians University, Munich, Germany

ARTICLE INFO

Article History:
Received 7 August 2020
Revised 13 November 2020
Accepted 16 November 2020
Available online 3 December 2020

Keywords:
DNA methylation based biomarkers of age acceleration
All-cause death
Myocardial infarction
Stroke
Cancer
Competing risk

ABSTRACT

Background: DNA methylation (DNAm) may play a role in age-related outcomes. It is not yet known which DNAm-based biomarkers of age acceleration (BoAA) has the strongest association with age-related endpoints.
Methods: We collected the blood samples from two independent cohorts: the Normative Ageing Study, and the Cooperative Health Research in the Region of Augsburg cohort. We measured epigenome-wide DNAm level, and generated five DNAm BoAA at baseline. We used Cox proportional hazards model to analyze the relationships between BoAA and all-cause death. We applied the Fine and Gray competing risk model to estimate the risk of BoAA on myocardial infarction (MI), stroke, and cancer, accounting for death of other reasons as the competing risks. We used random-effects meta-analyses to pool the individual results, with adjustment for multiple testing.
Findings: The mean chronological ages in the two cohorts were 74, and 61, respectively. Baseline GrimAgeAccel, and DNAm-related mortality risk score (DNAmRS) both had strong associations with all-cause death, MI, and stroke, independent from chronological age. For example, a one standard deviation (SD) increment in GrimAgeAccel was significantly associated with increased risk of all-cause death [hazard ratio (HR): 2.01; 95% confidence interval (CI), 1.15, 3.50], higher risk of MI (HR: 1.44; 95% CI, 1.16, 1.79), and elevated risk of stroke (HR: 1.42; 95% CI, 1.06, 1.91). There were no associations between any BoAA and cancer.
Interpretation: From the public health perspective, GrimAgeAccel is the most useful tool for identifying at-risk elderly, and evaluating the efficacy of anti-aging interventions.
Funding: National Institute of Environmental Health Sciences of U.S., Harvard Chan-NIEHS Center for Environmental Health, German Federal Ministry of Education and Research, and the State of Bavaria in Germany.

1. Introduction

The world population is rapidly growing older, leading to higher risk of age-related deaths and diseases, such as myocardial infarction (MI), and cancer. Notably, individuals of the same chronological age may exhibit different susceptibility to age-related endpoints, which is likely reflective of differences in their underlying biological aging processes [1]. Understanding the molecular mechanisms underlying aging process; identifying at-risk elderly populations; and evaluating the efficacy of interventions for anti-aging are critical in both basic research, and public health practice.

In the last decade, DNA methylation (DNAm) has been reported to be related to aging processes [2-7], and different DNAm ages have been linked with age-related outcomes [8-15]. More specifically, there are two generations of DNAm-based biological age metrics. The first generation includes the two clocks developed by Horvath [5], and Hannum [6] in 2013. Most recently, a second generation of DNAm ages incorporated additional age-related markers [12, 13] It includes DNAm PhenoAge developed by Levine et al. in 2018 [12], and

* Corresponding author.
E-mail address: cuicuiwang@hsph.harvard.edu (C. Wang).

https://doi.org/10.1016/j.ebiom.2020.103151

Research in Context

Evidence before this study

Elderly people with the same chronological age may have substantially dissimilar risk of death and diseases. DNA methylation (DNAm) has been reported to be related to the aging process. However, the potential molecular mechanisms are unclear. Several DNAm-based biomarkers of aging acceleration (BoAA) have relationships to age-related endpoints. It is not yet known which BoAA has the strongest association with specific age-related endpoints.

Added value of this study

We applied the Fine and Gray competing risk model, which considers the deaths from other outcomes as competing risks, to estimate the relationships between BoAA and MI, stroke, and cancer. The associations' estimations were more accurate compared to the conditional Cox model. In addition, this study had sizable samples with two independent cohorts.

Implications of all the available evidence

Our study examined multiple DNAm-based BoAA and diverse health endpoints. Our present study suggests that GrimAgeAccel could be the most useful tool for identifying at-risk elderly, and evaluating the efficacy of anti-aging interventions, independent from chronological age.

DNAm GrimAge generated by Lu et al. in 2019 [13], respectively. However, because of the moderate correlation between DNAm ages, and chronological age, it is not surprising to find predictive power of DNAm clocks for the prevalence, and incidence of age-related health outcomes. Recent studies have revealed that age acceleration (AA), which is the raw residual in the linear regression model with DNAm age regressed on chronological age, is a candidate biomarker for all-cause death, and age-related diseases [13, 16-22]. The well documented AA measures are epigenetic age acceleration of intrinsic, and extrinsic (i.e., IEAA, and EEAA) [17]. Most recently, Horvath et al. showed that PhenoAgeAccel, and GrimAgeAccel had strong predictive power for all-cause death, coronary heart disease, and cancer [13]. However, most related studies, to date, focused on IEAA, and EEAA [17-20], and few investigated the associations between PhenoAgeAccel / GrimAgeAccel and multiple health endpoints [13, 21].

Besides the epigenetic biomarkers mentioned above, Zhang et al. developed a continuous DNAm-related mortality risk score (i.e., DNAmRS) among 1900 older participants in ESTHER study [23]. In their study, they independently validated the strong association of DNAmRS and mortality among 1700 subjects in cooperative health research in the region of augsburg cohort (KORA F4) [23]. DNAmRS is a linear combination of 10 probes selected by least absolute shrinkage and selection operator (LASSO) regression and has been reported as a strong predictor for all-cause, and cardiovascular mortality [23, 24] Recently, our group found that DNAmRS was more predictive for all-cause, cardiovascular, and cancer mortality compared with telomere length, and PhenoAgeAccel [25]. However, when we conducted this analysis, GrimAgeAccel had not been generated yet; neither IEAA, nor EEAA were examined. Moreover, all the above studies including ours used conventional methods of survival analysis without considering the competing risk of death [26, 27] which may hinder the observation of the event of interest or modify the probability the event occurs. For example, when the onset of MI is of interest, competing risk from death of stroke or car accident before MI occurrence prevent us from observing it. Since subjects who die are no longer eligible to become MI patients, but may have had different MI risks than the uncensored subjects, this introduces bias into the effect estimates. In this case, methods such as the Fine and Gray competing risk model [28] that considers the competing risk should be more appropriate.

Herein, the goal of this present study was to assess the relationships of GrimAgeAccel / PhenoAgeAccel / IEAA / EEAA / DNAmRS and diverse health outcomes (i.e., all-cause death, MI, stroke, and cancer). Since higher AA, and DNAmRS were both associated with increased mortality [13, 19, 23] we defined all of them as biomarkers of age acceleration (BoAA) in this study. We then chose the most useful BoAA for identifying at-risk elderly, and evaluating the efficacy of anti-aging interventions, based on their hazard ratios (HRs) for a one standard deviation (SD) increment, and adjusted *P*-values. Analyses on MI, stroke, and cancer referred to the combination of both morbidity, and mortality, and were accounted for the competing risk of death [28].

2. Methods

2.1. Study population

2.1.1. Normative ageing study (NAS)

The NAS is a closed cohort study that was established in the Greater Boston in 1963 (N = 2280) [29]. The participants from NAS were all men, who were free of known chronic medical conditions at enrolment. The subjects have undergone examinations every 3 to 5 years on a continuous rolling basis. Information on clinical, and other health data was collected at these visits. We restricted to the participants with DNA samples from blood, which were collected from subjects starting in 1999 [30]. We excluded non-whites (3%), and visits that did not have complete information on covariates (see below), leaving a final sample size of 737 men. The examination at which a DNA sample was first obtained was considered as the baseline examination. Because of the rolling examinations, and since some individuals missed visits, the baseline year ranged from 1999 to 2012. The last available update was in 2016. All participants provided written informed consent. Both the Harvard T.H. Chan School of Public Health, and the Institutional Review Boards of the Department of Veterans Affairs approved this study.

2.1.2. Cooperative health research in the region of augsburg cohort (KORA F4)

KORA F4 is a population-based cohort in south Germany established between 2006 and 2008 (N = 3080) [31]. Methylation profiles were generated from 1802 participants who were randomly selected [32]. We included 1725 subjects (882 women) since 75 samples were removed after quality control, and 2 participants dropped out. The last available update was in 2016. All participants supplied written informed consent that was approved by the Ethics Committee of the Bavarian Medical Association.

2.2. Measures of events

Only endpoints that occurred after the baseline when the first DNA sample was collected were treated as events. Vital status of participants was assessed by follow-up mailings in the two studies. The events included all-cause death, MI, stroke, and cancer; the latter three referred to the combination of both morbidity, and mortality.

All-cause death. Vital status was monitored by regular checking in two cohorts. For participants who have died, their documents were reviewed to assign cause of death codes according to the 9th revision of the International Classification of Disease (ICD-9) [33-35].

Incidence and mortality of MI. In NAS, the diagnostic criteria for MI were adapted from those used in the Framingham Heart Study, and were adjudicated by a research physician [36]. In KORA F4, MI was assessed in the Augsburg MI registry [37]. Death of MI was coded as ICD-9 410.

Incidence and mortality of stroke. In NAS, non-fatal strokes were defined as neurological deficit of sudden or rapid onset that persisted for at least 24 hours [38]. In KORA F4, non-fatal strokes were assessed by participants' hospital records [39]. Death of stroke was coded as ICD-9 431−436.

Incidence and mortality of cancer. In NAS, cancer information was obtained from questionnaires, and confirmed by medical records [40]. In KORA F4, incidence of cancer was unavailable, but death of cancer was certificated by local health authorities [41], and was coded as ICD-9 140−239.

2.3. Measures of DNAm-based BoAA

DNA samples were extracted from whole blood using QIAamp DNA Blood Kit (Qiagen, CA, USA), and were bisulphite converted using the EZ-96 DNA Methylation Kit (Zymo Research, Orange, CA, USA) as described previously [42, 43]. We measured DNAm across ~ 485,000 5′-C-phosphae-G-3′ (CpG) sites at a single nucleotide resolution, with Illumina Infinium HumanMethylation450 BeadChip (Infinium HD Methylation protocol, Illumina, San Diego, CA, USA), in both NAS, and KORA F4. We utilized a two-stage age-stratified algorithm to assign the samples to plates randomly, which standardized the age distribution across plates [44].

In NAS, the blood samples were collected between 1999 and 2012. In the quality control step, we removed problematic samples that might be mislabelled, or contaminated, or had poor performance due to technological issues [45]. At the probe level, we used non-specific fluorescence via the *ewastools* package in GitHub to remove probes with a detection $P > 0.01$ [46]. We corrected the data for dye-bias [47]. Instead of normalizing the DNAm data directly [11, 13, 17] we normalized DNAm data by controlling for the normalization factors in the outcome regression. We used LASSO and an elastic-net regularized generalized linear model to extract five important experimental covariates – Non polymorphic Red, Specificity I Red, Bisulfite Conversion I Red, Bisulfite Conversion II, Extension Red – from the control metrics monitoring the execution of diverse experimental steps [48]. We then controlled for these five experimental covariates in our outcome regression (see 2.4 Statistical method). This normalization approach generally works better [49], and we have applied it previously [30].

In KORA F4, the blood was collected between 2006 and 2008. The data pre-processing, including data quality (removal of records according to functional beads, detection *P*, and SNP frequency), data improvement (background, and dye bias correction), and probe type adjustment (normalization using *BMIQ* package), has been described in detail previously [50].

Measures of AA. In the *New Methylation Age Calculator* website (https://dnamage.genetics.ucla.edu/new), we uploaded our DNAm data, and sample annotation file. After selecting "Normalize Data", and "Advanced Analysis", we submitted the data, and got outputs including four AA metrics (i.e., GrimAgeAccel, PhenoAgeAccel, IEAA, and EEAA) automatically via email. A positive AA measure indicates that the subject is older than expected based on the chronological age at the baseline, and vice versa.

Measure of DNAmRS. We calculated DNAmRS based on 10 selected CpGs by Zhang et al. [23]. The formula is:

$$\begin{aligned}\text{DNAmRS} = {} & \text{cg01612140} \times (-0.38253) + \text{cg05575921} \\ & \times (-0.92224) + \text{cg06126421} \times (-1.70129) + \text{cg08362785} \\ & \times (2.71749) + \text{cg10321156} \times (-0.02073) + \text{cg14975410} \\ & \times (-0.04156) + \text{cg19572487} \times (-0.28069) + \text{cg23665802} \\ & \times (-0.89440) + \text{cg24704287} \times (-2.98637) + \text{cg25983901} \\ & \times (-1.80325)\end{aligned}$$

To facilitate comparisons of effect sizes across BoAA metrics, we expressed our results as differences in HRs per one SD increase in BoAA.

2.4. Statistical methods

We performed a Pearson correlation analysis, and examined the data distribution among the five different BoAA (i.e., GrimAgeAccel, PhenoAccel, IEAA, EEAA, DNAmRS) in each cohort.

For all-cause death, we used Cox proportional hazards model to estimate the HRs associated with each BoAA collected at baseline. For MI, stroke, and cancer, we used the Fine and Gray competing risk model to evaluate the HRs with death from other reasons considered to be the competing risks [28]. The coefficients estimated from a traditional Cox model reflect the effect of a covariate on the cause-specific hazard, considering competing risk events as censoring without information. Fine-Gray model is a regression on the sub-distribution hazard which estimates the effect of a covariate on the cumulative incidence, taking into account the informative censoring of the competing risk [27, 51].

We controlled for covariates *a priori* based on the relevant literature [16, 19, 23]: 5-year categories of chronological age, gender (KORA F4 only), body mass index (BMI), cigarette pack years, smoking status (never, ever), alcohol consumption ($<$ 2 drinks/day, $\geq$ 2 drinks/day in NAS; g/day in KORA F4), years of education, 10-mg/dl categories of serum high-density lipoprotein, serum triglyceride, estimated cell types via the Houseman et al. method [52], and batch effect. All covariates were measured at baseline. All the categorical variables were adjusted by *strata* function in the models. For IEAA, and EEAA, we did not adjust for cell types [53]. For DNAmRS in the NAS, we additionally adjusted for the above mentioned five experimental covariates because this biomarker was calculated by 10 CpGs, which were not normalized directly [48] (We selected "Normalize Data" when we calculated four AA metrics in the online *Calculator*).

We performed random-effects meta-analyses [54] to combine the estimates from the NAS, and KORA F4, and evaluated the overall relationships between each of the five BoAA and each of the four health endpoints. In order to capture modification by smoking, we further added an interaction term between BoAA and smoking status (never, ever) in the models. We tested the proportional hazards assumption in the Cox model using the *cox.zph()* function in the *survival* package, and the sub-proportional hazards assumption in the Fine and Gray model using *diag_crr()* function in the *crskdiag* package in R. The two assumptions were fulfilled for all variables in two models accordingly.

Results from survival analyses were expressed as HRs for a one SD increase in BoAA [54]. We applied Benjamini-Hochberg false discovery rate ($FDR_{B\text{-}H}$) to adjust for multiple comparison. The threshold for statistical significance was 0.05 [55]. We used the *I*-squared (I^2) test on random-effects estimates to access heterogeneity. $I^2 < 0.50$, and *P*-value > 0.05 were considered as homogeneous [50].

All statistical analyses were performed using R Version 3.6.3 (R Core Team, Vienna, Austria) with *survival* (Cox proportional hazards model), *cmprsk* (Fine and Gray competing risk model), and *metafor* (meta-analyses) packages.

3. Results

3.1. Population description

Table 1 presents the summary statistics of participants at baseline. In the NAS, there were 737 men enrolled. The mean chronological age was 74 years (SD = 7). Mean of BMI was 28.0 kg/m^2 (SD = 4.1). The subjects were well educated, with a mean 15.0 education years (SD = 3.0). Furthermore, 229 (31.1%) of the 737 individuals were never smokers, and 602 (81.7%) had fewer than two drinks per day (Table 1). In KORA F4, there were 1725 participants (882 women) enrolled. The mean

Table 1
Demographic and characteristics of participants at baseline from the NAS (*N* = 737) and KORA F4 (*N* = 1725).

Variables	NAS (Year 1999–2012)	KORA F4 (Year 2006–2008)
Chronological age (years), mean ± SD	74 ± 7	61 ± 9
BMI (kg/m^2), mean ± SD	28.0 ± 4.1	28.1 ± 4.8
Missing, n (%)	0 (0)	6 (0.3)
Women, n (%)	0 (0)	882 (51.1)
Smoking, n (%)		
Never	229 (31.1)	721 (41.8)
Ever	508 (68.9)	1002 (58.1)
Missing, n (%)	0 (0)	2 (0.1%)
Pack-year smoked (years), mean ± SD	21.2 ± 25.3	13.1 ± 21.4
Alcohol consumption	Drinks/day, n (%) ≥2; 135 (18.3) <2; 602 (81.7)	g/day (mean ± SD) 15.5 ± 20.5
Education (years), mean ± SD	15.0 ± 3.0	11.5 ± 2.6
Missing, n (%)	0 (0)	3 (0.2)
Serum high-density lipoprotein (mg/dL)	49.0 ± 12.8	56.5 ± 14.6
Missing, n (%)	0 (0)	0 (0)
Serum triglyceride (mg/dL)	137.5 ± 87.1	133.0 ± 88.6
Missing, n (%)	0 (0)	0 (0)
Estimated cell types,%		
Monocytes	10.5	7.9
B cells	1.5	7.3
CD4+ *T* lymphocytes	11.5	16.5
CD8+ *T* lymphocytes	4.1	5.6
Natural killer cells	7.3	7.3
Missing, n (%)	0 (0)	0 (0)

Abbreviations: NAS, Normative Aging Study; KORA F4, Cooperative Health Research in the Region of Augsburg cohort; MI, myocardial infarction; SD, standard deviation; BMI, body mass index.

chronological age was 61 (SD=9). Mean of BMI was 28.1 kg/m^2 (SD = 4.8), which was almost the same with that in NAS. The percentage of never smokers was 41.8, larger than that of NAS (31.1). The years of education in KORA F4 (mean=11.5, SD=2.6) were less than that in NAS (Table 1).

Table 2 shows the summary of events of interest during follow-up period. In NAS, there were 337 all-cause death after baseline. Sixty-nine participants had non-fatal MIs or died of MI, 30 had non-fatal strokes or died of stroke, and 243 had cancer or died of cancer. The median follow-up years of all-cause death, MI, stroke, and cancer were 11.0, 14.0, 14.0, and 12.0 years, respectively (Table 2). In KORA F4, 146 subjects (90 men, and 56 women) died in total. Ninety participants had non-fatal MI or died of MI, 81 had non-fatal stroke or died of stroke, and 55 died of cancer (KORA F4 did not record the information on cancer incidence). The median follow-up years of all-cause death, MI, stroke, and cancer were 8.5, 8.8, 8.8, and 8.5 years, respectively in KORA F4 (Table 2).

Table 2
Summary of events in the Normative Aging Study (*N* = 737) and KORA F4 (*N* = 1725).

Events	NAS	KORA F4
Last available update year	2016	2014
Follow-up time (years), median ± SD		
All-cause death	11.0 ± 4.4	8.5 ± 1.4
MI	14.0 ± 3.9	8.8 ± 1.1
Stroke	14.0 ± 3.1	8.8 ± 1.4
Cancer	12.0 ± 5.5	8.5 ± 1.4
Counts of events, n (%)		
All-cause death	337 (45.7)	146 (8.5)
MI	69 (9.4)	90 (5.2)
Incidence	45 (6.1)	62 (3.6)
Mortality	24 (3.3)	28 (1.6)
Stroke	30 (4.1)	81 (4.7)
Incidence	4 (0.5)	72 (4.2)
Mortality	26 (3.5)	9 (0.5)
Cancer	243 (33.0)	55 (3.2)
Incidence	138 (18.7)	NA[a]
Mortality	105 (14.2)	55 (3.2)

[a] KORA F4 did not have the record on the incidence of cancer during the study period. Abbreviations: NAS, Normative Aging Study; KORA F4, Cooperative Health Research in the Region of Augsburg cohort; MI, myocardial infarction.

We presented the Pearson correlations, and data distributions among five BoAA in Supplementary Material (Figure S1). Generally, they were significantly correlated with each other in both NAS, and KORA F4. GrimAgeAccel, and DNAmRS had the strongest correlation.

We stratified the 737 NAS men into two groups based on each of the four AA, and estimated the difference in Kaplan-Meier survival probability between the two groups. Taking GrimAgeAccel as an example, we defined a group as biologically older if GrimAgeAccel was greater than zero, and otherwise defined it as biologically younger. The number of men in the older, and younger group were 293, and 444, respectively. We plotted a Kaplan-Meier survival curve for each group. Similarly, we divided our participants into two groups based on the other three AA (i.e., PhenoAgeAccel, IEAA, EEAA), and plotted Kaplan-Meier survival curves for all groups. Fig. 1a suggested the survival probabilities were significantly higher in GrimAgeAccel younger, PhenoAgeAccel younger, and the EEAA younger groups (Fig. 1a). We divided the subjects into two groups, and compared the Kaplan-Meier survival curves based on the same rule in KORA F4, and found that the survival probabilities were higher in younger groups classified by the same three DNAm AA metrics (i.e., GrimAgeAccel, PhenoAgeAccel, and EEAA) (Fig. 1b). We also plotted the cumulative incidence curves of all-cause death, MI, stroke, and cancer (Supplementary Material, Figure S2).

3.2. Relationships between baseline DNAm-based BoAA and event of interests

The meta-analyses indicated significant increased risk of all-cause death associated with GrimAgeAccel, PhenoAgeAccel, and DNAmRS (Fig. 2a). More specifically, each increment of a one SD in GrimAgeAccel (4.07 y), and PhenoAgeAccel (6.13 y) was associated with all-cause death with multivariate-adjusted HRs of 2.10 (95%CI: 1.15, 3.50, $FDR_{B\text{-}H}$ = 0.014), and 1.24 (95%CI: 1.05, 1.45, $FDR_{B\text{-}H}$ = 0.010), respectively. The HR of all-cause death was 1.67 (95%CI: 1.04, 2.70, $FDR_{B\text{-}H}$ = 0.034) for a one SD increment in DNAmRS (0.44 score) (Fig. 2a). The meta-analyses of the Fine and Gray competing risk models suggested significant increased risks of MI, and stroke associated with increments in GrimAgeAccel, PhenoAgeAccel, EEAA, and DNAmRS (Fig. 2b, and 2c). For instance, the risks of MI and stroke were about 1.44 (95%CI: 1.16, 1.79, $FDR_{B\text{-}H}$ = 0.001), and 1.42 (95%CI: 1.06, 1.91, $FDR_{B\text{-}H}$ = 0.019) times when GrimAgeAccel increased a one SD (4.07 y). There were no significant relationships between any of the five BoAA and cancer (Fig. 2d). From the public heath perspective, GrimAgeAccel was the best BoAA for determining at-risk populations, and assessing anti-aging interventions because the HRs carried the largest risk for all-cause death, MI, and stroke. Although study-specific associations were almost all positive (Supplementary Material, Figure S3), there were significant heterogeneity for GrimAgeAccel / DNAmRS and all-cause death across two cohorts, and generally no evidence of heterogeneity in the associations between BoAA and MI (Supplementary Material, Table S1).

We also categorized our subjects into four groups based on AA: AA ≥ 0, $-2 \leq$ AA < 0, $-5 \leq$ AA < -2, and AA < -5. We set AA > 0 (e.g., GrimAgeAccel > 0) as the reference, and estimated the reduced risk of all-cause death for groups who were younger than expected (i.e., AA < 0). The meta-analyses showed that for GrimAgeAccel, all groups with AA < 0 had reduced risks of mortality while the $FDR_{B\text{-}H}$ was

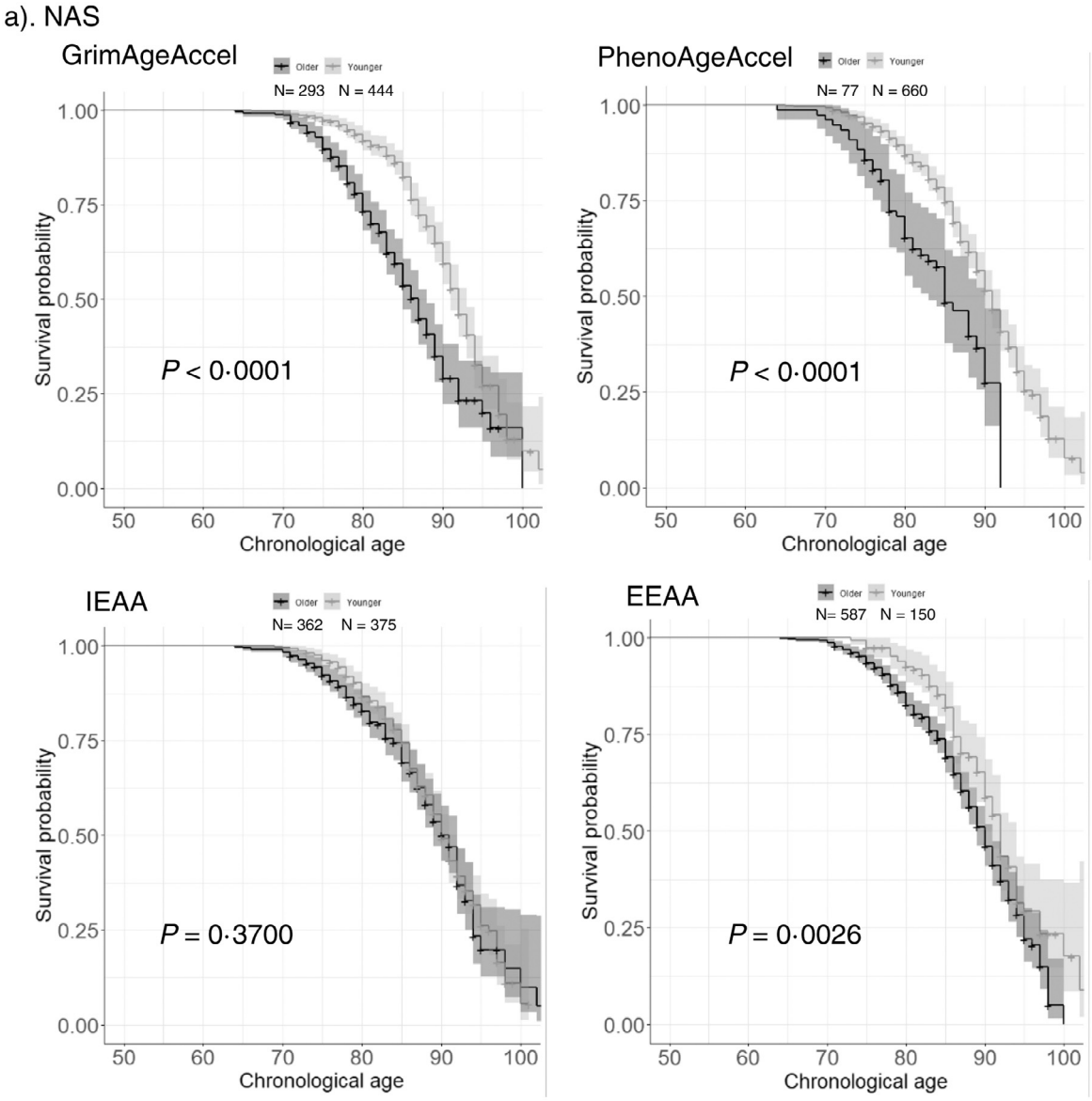

Fig. 1. Kaplan-Meier survival probability curves between two groups (AA > 0 or not).

larger than 0.05 in the group with AA < −5. The trend test showed that there were significant trends for reduced mortality as GrimAgeAccel (P = 0.00), PhenoAgeAccel (P = 0.03), and EEAA (P = 0.04) decreased (Fig. 3).

The results from models with the interaction term between BoAA and smoking status (never, ever) suggested strong, and positive associations between BoAA and all-cause death / MI / stroke among ever smokers. GrimAgeAccel carried the largest risks for endpoints (Supplementary Material, Figure S4). Among the never smokers, GrimAgeAccel, and DNAmRS remained positively associated with increased risk, however the associations were no longer significant (Supplementary Material, Figure S5).

4. Discussion

The meta-analysis of two cohorts assessed the application of five DNAm-based BoAA for differentiating risk for age-related deaths and diseases. Our findings showed that: independent of chronological age, 1) GrimAgeAccel, PhenoAgeAccel, EEAA, and DNAmRS were significantly positively associated with all-cause death, MI, and stroke; 2) from the public health perspective, GrimAgeAccel was the best biomarker because it had the largest HRs per one SD increase with events among all the BoA. Hence, while GrimAgeAccel was developed as a mortality biomarker originally [13], it is the most useful tool to identify the elderly who have high-risk of not only all-cause death, but also other age-related health outcomes.

In the U.S., heart disease ranks as top one cause of death [56]. MI occurs when blood blow reduces or pauses in a part of the heart, leading to impairment in the heart muscle. Elderly people are not only prone to die but have higher risks of having diseases [57]. Among men who are older than 75 years, over 5% have had an MI without or only little history of symptoms [58]. A stroke happens when blood cannot flow to a part of the brain, and it is the fourth leading cause of death in the U.S [59]. Ageing is the most strong risk factor for stroke – around 75% strokes occur in people aged 65 or above [60]. According to World Bank, the populations 65 year, and older were reported at 16% in the U.S., and 21% in Germany in 2018 [61], suggesting that age is a major contributor towards MI and stroke in these countries. Our stable findings of the strong associations between GrimAgeAccel / DNAmRS and all-cause death / MI / stroke

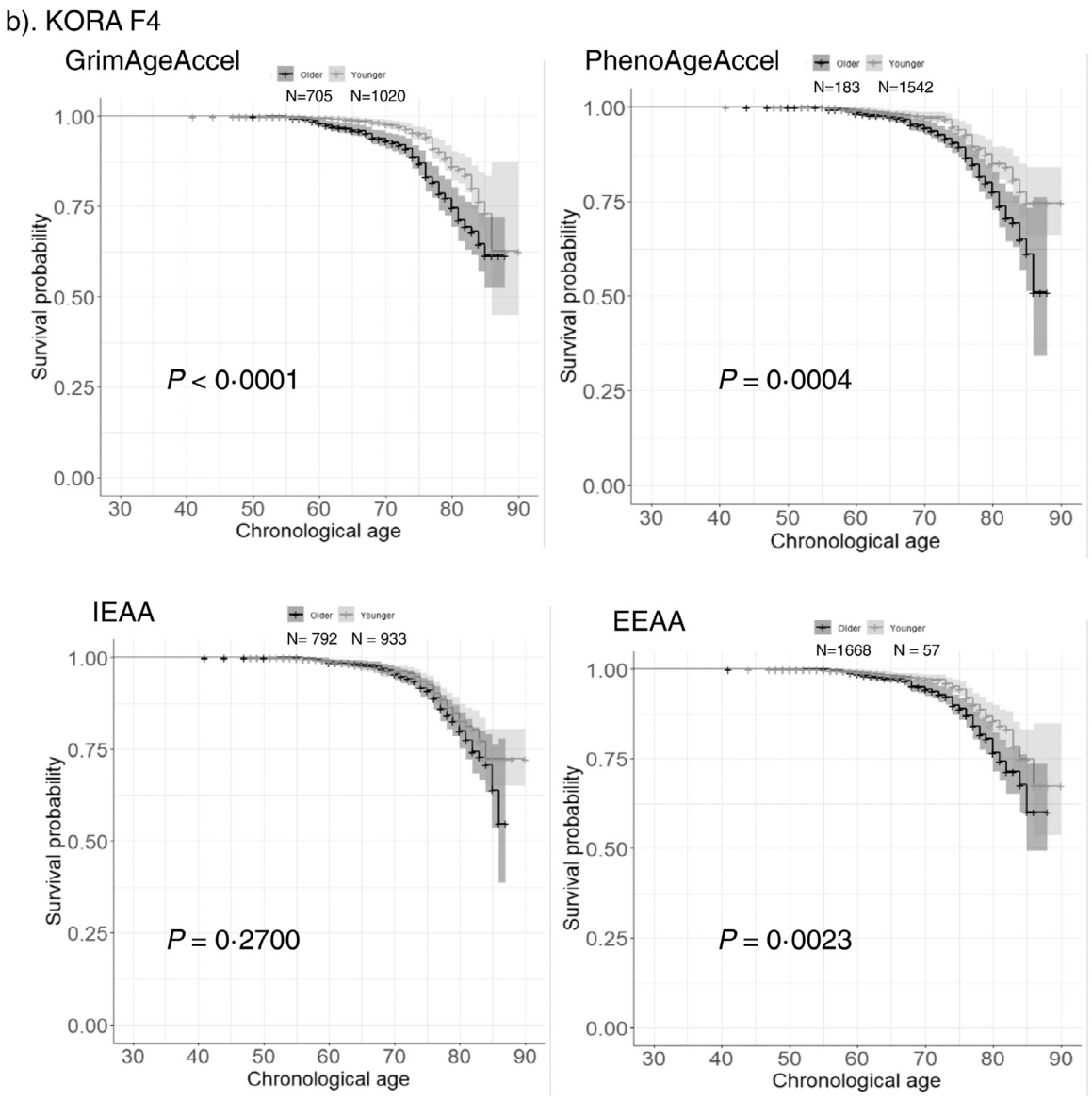

Fig. 1 Continued.

may help us understand some of the underlying biological mechanisms of the aging process. We do not know the 1030 CpG sites of GrimAgeAccel because Liu et al. have not published them yet [13]. However, the 10 probes of DNAmRS are identifiable [23]. One of the 10 CpGs – cg23665802 – is mapped to the *microRNA-19a* gene. Circulating microRNA-19a concentration could be a novel biomarker for the diagnosis of acute MI [62]. In a mouse model, Gao et al. have validated that microRNA-19a could protect the heart function from MI and secure heart activity [63]. Hence, microRNA-19a is a potential therapeutic target to prevent and treat heart disease [63]. Another CpG site (cg08362785) is mapped to the gene of *myocardin related transcription factor A*. In both murine and human hearts, *myocardin related transcription factor A* has been shown to be important for developing cardiac function, maintaining heart structure and homeostasis [64, 65] and might be expected to have a role in age-related MI.

From the public health perspective, our study provides practical implications for GrimAgeAccel in assessing the risk of age-related outcomes for the elderly. By identifying elderly populations at higher risks of death and diseases, those people can be provided with specific protection, such as medical equipment and community care. In addition, GrimAgeAccel could also be used to evaluate anti-aging interventions. As life expectancy increases in people who live up to 65 worldwide, their health span has not increased accordingly. Developing intervention (e.g., diets) to slow aging and increase health lifespan may effectively reduce age-related endpoints. Determining the effects of anti-aging intervention often require decades of follow-up. Our findings suggest that GrimAgeAccel could be taken as surrogates to facilitate assessment of anti-aging intervention efficacy.

In models with an interaction term between BoAA and smoking status (never, ever), we found consistently strong effects of GrimAgeAccel / DNAmRS on all-cause death / MI / stroke in ever smokers whereas no associations remained significant after restricting to never smokers. GrimAgeAccel and DNAmRS remained positively associated with mortality, MI, and stroke; and the lack of significance may reflect the smaller sample size of never smokers. However, it is possible that GrimAgeAccel and DNAmRS primarily pick up changes of smoking-related pathways. Because the probes behind GrimAgeAccel have not been published yet, this study is only able to investigate the 10 CpGs of DNAmRS. Among them, eight (cg01612140, cg05575921, cg06126421, cg08362785, cg14975410, cg19572487, cg23665802, cg24704287) were smoking-related probes as described by Joehanes et al., who investigated the association between cigarette

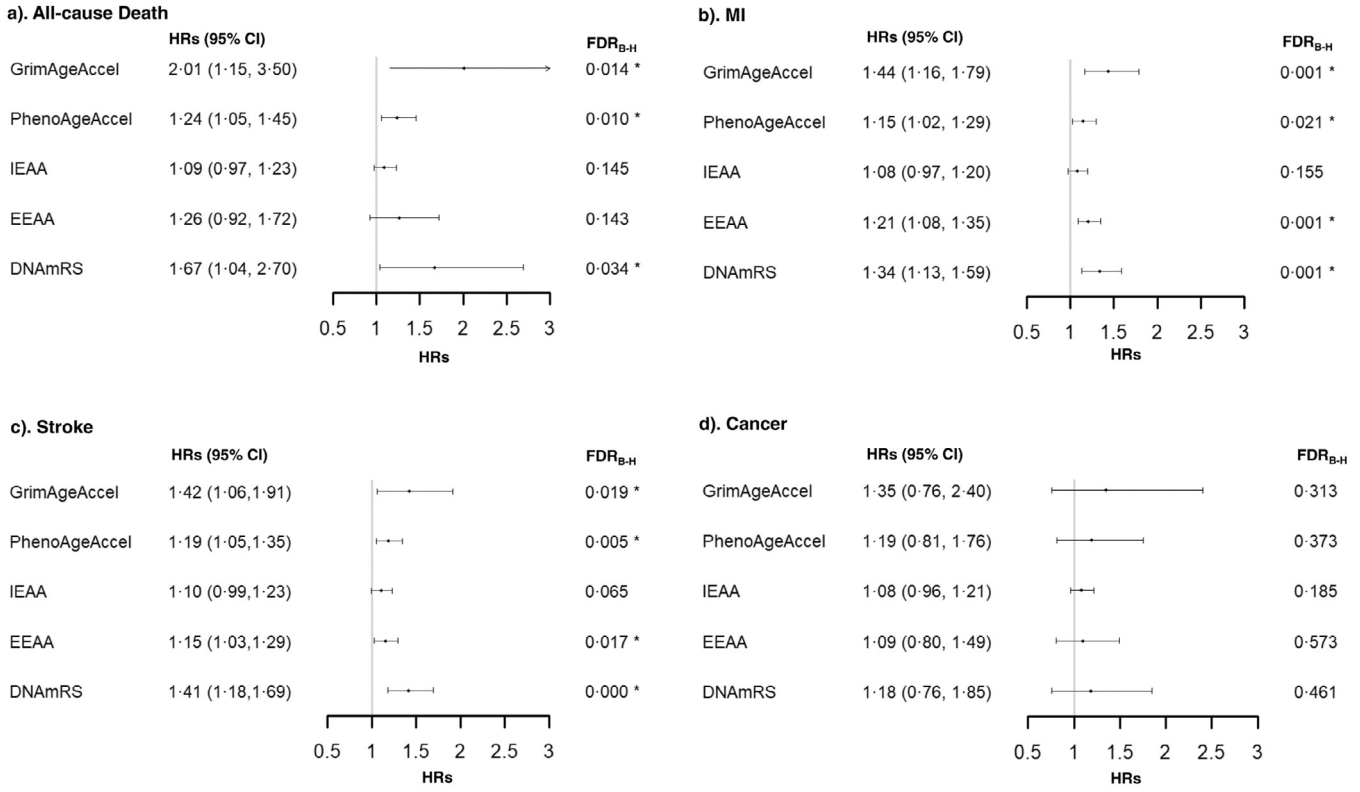

Fig. 2. Associations between baseline DNAm-based BoAA and endpoints. Abbreviations: IEAA=intrinsic epigenetic age acceleration; EEAA=extrinsic epigenetic age acceleration; HRs=Hazard Ratios; MI=myocardial infarction; CI=confidence interval; FDR=Benjamini-Hochberg false discovery rate.

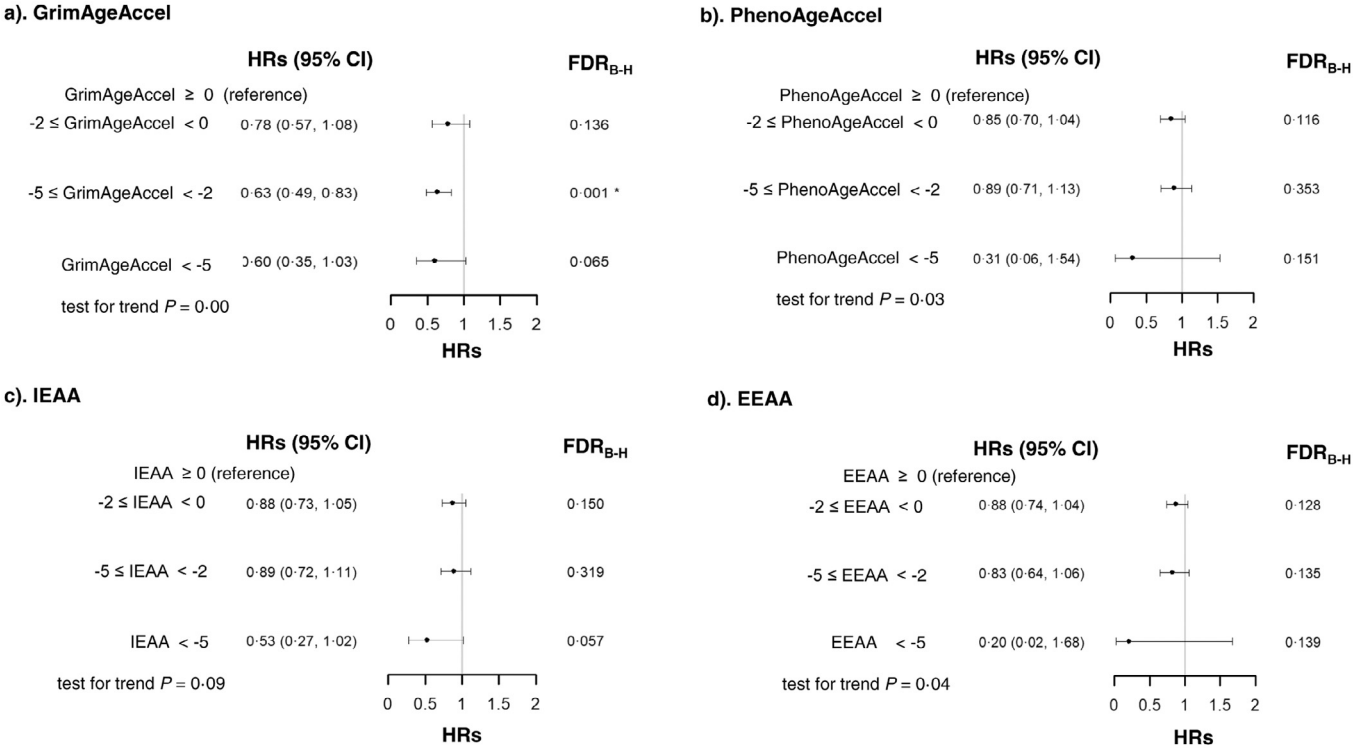

Fig. 3. Associations between baseline AA and all-cause death among four groups (AA ≥ 0, −2 ≤ AA < 0, −5 ≤ AA < −2, and AA < −5). Abbreviations: IEAA=intrinsic epigenetic age acceleration; EEAA=extrinsic epigenetic age acceleration; HRs=Hazard Ratios; CI=confidence interval; FDR=Benjamini-Hochberg false discovery rate.

smoking and DNAm from 16 cohorts with 15907 blood DNA samples [66]. However, whether DNAmRS was specific to smoking or represented the effects of smoking pathways need to be further investigated.

The traditional Kaplan-Meier approach assumes that these individuals would experience the same probability of event of interest (e.g., the onset of MI) if they were not censored by death. In fact, people who die may not have the same underlying risk of having MI. Thus, Kaplan-Meier method overestimates the cumulative incidence of the event of interest [67, 68]. For example, if we simply applied Kaplan-Meier method to estimate DNAmRS's effect on cancer, we found that a one SD increment in DNAmRS increased the risk of cancer by 1.53 times (95% CI: 1.24, 1.89) (data not shown), which was greater than the results from the Fine and Gray competing risk model (HR 1.18, 95% CI: 0.76, 1.85). In the present study, when we evaluated the relationship between BoAA and each of MI, stroke, and cancer, we considered deaths from other outcomes as competing risks, and fit the Fine and Gray competing risk models to investigate the associations, as applied in other survival analyses [69-71].

Our study does have several limitations: i) we measured these BoAA in blood samples. Findings in other tissues need to be examined in future studies; ii) we used the existing literature and a *priori* knowledge of clinical relevance to adjust for potential confounders for the health endpoints. Therefore, we may have neglected unknown confounding in our analyses; iii) the two different study locations (the U.S. and Germany) may confound our results because the health outcomes may be influenced by living environment (e.g., mortality due to ambient fine particulate matter) [72]; iv) the participants in KORA F4 were randomly selected, whereas the participants in NAS were all men and free of known chronic diseases at enrolment. The different recruiting strategies in two cohorts may limit generalizability or induce bias in our results.

The major strengths in the study are as follows: i) two cohorts with sizable samples (total N = 2462) and DNAm data in epigenome-wide; ii) assessment the application of multiple DNAm-based BoAA from currently available candidates for differentiating risks of multiple health endpoints; iii) application of the Fine and Gray competing risk model, leading to more accurate and meaningful interpretation of the association estimates; iv) compared with our previous study [25], this present NAS cohort not only had longer follower up (last available update was 2016 vs 2014), but considered both morbidity and mortality, competing risk and more BoAA.

In summary, we reported that baseline GrimAgeAccel was the strongest DNAm-based BoAA for all-cause death, MI, and stroke, adjusting for chronological age and other risk factors. GrimAgeAccel may serve as the useful tool for identifying at-risk elderly, and evaluating anti-aging intervention efficacy. Further investigation in other cohorts is needed.

Funding/Support

This work was supported by grants from the National Institute of Environmental Health Sciences (R01ES015172, 1R01ES027747, P30ES009089, R01ES021733, and R01ES025225), and HSPH—NIEHS Center for Environmental Health (ES000002) for data collection, analysis, and interpretation. The KORA F4 study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria.

Author contributions

Wang and Ni take responsibility for the integrity of the data and the accuracy of the data analysis in the study.

Concept and design: Wang, Baccarelli, Schwartz.

Acquisition, analysis, or interpretation of data: Kosheleva, Peters, Baccarelli, Schwartz, Wang, Ni, Yao, Just, Heiss.

Drafting of the manuscript: Wang, Schwartz.

Critical revision of the manuscript for important intellectual content: Coull, Peters, Baccarelli, Schwartz, Wei, Gao.

Statistical analysis: Wang, Ni, Schwartz.

Obtained funding: Baccarelli, Schwartz.

Administrative, technical, or material support: Baccarelli, Schwartz

Study supervision: Schwartz

Declaration of Competing Interest

None reported.

Acknowledgements

We thank Dr. John Hutchinson for his critical review of this manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2020.103151.

References

[1] Lowsky DJ, Olshansky SJ, Bhattacharya J, Goldman DP. Heterogeneity in healthy aging. J Gerontol A Biol Sci Med Sci 2014;69:640–9.
[2] Christensen BC, Houseman EA, Marsit CJ, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet 2009;5:e1000602.
[3] Day K, Waite LL, Thalacker-Mercer A, et al. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. Genome Biol 2013;14:R102.
[4] Xu Z, Taylor JA. Genome-wide age-related DNA methylation changes in blood and other tissues relate to histone modification, expression and cancer. Carcinogenesis 2013;35:356–64.
[5] Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14:R115.
[6] Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 2013;49:359–67.
[7] Belsky DW, Caspi A, Arseneault L, et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. Elife 2020;9:e54870.
[8] Marioni RE, Shah S, McRae AF, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol 2015;16:25.
[9] Chen BH, Marioni RE, Colicino E, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY) 2016;8:1844–65.
[10] Hillary RF, Stevenson AJ, McCartney DL, et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. Clin Epigenetics 2020;12:115.
[11] Roshandel D, Chen Z, Canty AJ, et al. DNA methylation age calculators reveal association with diabetic neuropathy in type 1 diabetes. Clin Epigenetics 2020;12:1–16.
[12] Levine ME, Lu AT, Quach A, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY) 2018;10:573–91.
[13] Lu AT, Quach A, Wilson JG, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY) 2019;11:303–27.
[14] Li X, Ploner A, Wang Y, et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. Elife 2020;9:e51507.
[15] Hillary R.F., Stevenson A.J., Cox S.R., et al. An epigenetic predictor of death captures multi-modal measures of brain health. Mol Psychiatry 2019: 1–11.
[16] Perna L, Zhang Y, Mons U, Holleczek B, Saum K, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics 2016;8:64.
[17] Horvath S, Gurven M, Levine ME, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol 2016;17:171.
[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.
[19] Zheng Y, Joyce BT, Colicino E, et al. Blood epigenetic age may predict cancer incidence and mortality. EBioMedicine 2016;5:68–73.
[20] Ambatipudi S, Horvath S, Perrier F, et al. DNA methylome analysis identifies accelerated epigenetic ageing associated with postmenopausal breast cancer susceptibility. Eur J Cancer 2017;75:299–307.

[21] Kresovich JK, Xu Z, O'Brien KM, Weinberg CR, Sandler DP, Taylor JA. Epigenetic mortality predictors and incidence of breast cancer. Aging (Albany NY) 2019; 11:11975–87.
[22] Wolf EJ, Logue MW, Stoop TB, et al. Accelerated DNA methylation age: associations with Posttraumatic Stress Disorder and Mortality. Psychosom Med 2017; 80:42–8.
[23] Zhang Y, Wilson R, Heiss J, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun 2017;8:14617.
[24] Zhang Y, Saum KU, Schottker B, Holleczek B, Brenner H. Methylomic survival predictors, frailty, and mortality. Aging (Albany NY) 2018;10:339–57.
[25] Gao X, Colicino E, Shen J, et al. Comparative validation of an epigenetic mortality risk score with three aging biomarkers for predicting mortality risks among older adult males. Int J Epidemiol 2019;48:1958–71.
[26] Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. Br J Cancer 2004;91:1229–35.
[27] Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. J Clin Epidemiol 2013;66:648–53.
[28] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.
[29] Bell B, Rose CL, Damon A. The Veterans Administration longitudinal study of healthy aging. Gerontologist 1966;6:179–84.
[30] Wang C, Just A, Heiss J, et al. Biomarkers of aging and lung function in the normative aging study. Aging (Albany NY) 2020;12:11942–66.
[31] Ruckert IM, Heier M, Rathmann W, Baumeister SE, Doring A, Meisinger C. Association between markers of fatty liver disease and impaired glucose regulation in men and women from the general population: the KORA-F4-study. PLoS ONE 2011;6:e22932.
[32] Ward-Caviness CK, Nwanaji-Enwerem JC, Wolf K, et al. Long-term exposure to air pollution is associated with biological aging. Oncotarget 2016;7:74510–25.
[33] Mroczek DK, Stawski RS, Turiano NA, et al. Emotional Reactivity and Mortality: longitudinal Findings From the VA Normative Aging Study. J Gerontol B Psychol Sci Soc Sci 2015;70:398–406.
[34] Rost S, Freuer D, Peters A, et al. New indexes of body fat distribution and sex-specific risk of total and cause-specific mortality: a prospective cohort study. BMC Public Health 2018;18:427.
[35] Koenig W, Khuseyinova N, Baumert J, Meisinger C. Prospective study of high-sensitivity C-reactive protein as a determinant of mortality: results from the MONICA/KORA Augsburg Cohort Study, 1984-1998. Clin Chem 2008;54:335–42.
[36] Todaro JF, Con A, Niaura R, Spiro AI, Ward KD, Roytberg A. Combined effect of the metabolic syndrome and hostility on the incidence of myocardial infarction (the Normative Aging Study). Am J Cardiol 2005;96:221–6.
[37] Ward-Caviness CK, Agha G, Chen BH, et al. Analysis of repeated leukocyte DNA methylation assessments reveals persistent epigenetic alterations after an incident myocardial infarction. Clin Epigenetics 2018;10:161.
[38] Baccarelli A, Wright R, Bollati V, et al. Ischemic heart disease and stroke in relation to blood DNA methylation. Epidemiology 2010;21:819–28.
[39] Zahn K, Linseisen J, Heier M, et al. Body fat distribution and risk of incident ischemic stroke in men and women aged 50 to 74 years from the general population. The KORA Augsburg cohort study. PLoS One 2018;13:e0191630.
[40] Hou L, Joyce BT, Gao T, et al. Blood telomere length attrition and cancer development in the normative aging study cohort. EBioMedicine 2015;2:591–6.
[41] Autenrieth CS, Baumert J, Baumeister S, et al. Association between domains of physical activity and all-cause, cardiovascular and cancer mortality. Eur J Epidemiol 2011;26:91–9.
[42] Wang C, Koutrakis P, Gao X, Baccarelli A, Schwartz J. Associations of annual ambient PM2. 5 components with DNAm PhenoAge acceleration in elderly men: the Normative Aging Study. Environ Pollut 2019;258:113690.
[43] Kriebel J, Herder C, Rathmann W, et al. Association between DNA Methylation in Whole Blood and Measures of Glucose Metabolism: KORA F4 Study. PLoS ONE 2016;11:e0152314.
[44] Carmona JJ, Barfield RT, Panni T, Nwanaji-Enwerem JC, Just AC. Metastable DNA methylation sites associated with longitudinal lung function decline and aging in humans: an epigenome-wide study in the NAS and KORA cohorts. Epigenetics 2018;13:1039–55.
[45] Heiss JA, Just AC. Identifying mislabeled and contaminated DNA methylation microarray data: an extended quality control toolset with examples from GEO. Clin Epigenetics 2018;10:73.
[46] Heiss JA, Just AC. Improved filtering of DNA methylation microarray data by detection p values and its impact on downstream analyses. Clin Epigenetics 2019;11:15.
[47] Xu Z, Langie SA, De Boever P, Taylor JA, Niu L. RELIC: a novel dye-bias correction method for Illumina Methylation BeadChip. BMC Genomics 2017;18:4.
[48] BeadArray controls reporter. San Diego, CA: Illumina; 2015 October1000000004009 v00.
[49] Heiss JA, Brenner H. Between-array normalization for 450 K data. Front Genet 2015;6:92.
[50] Panni T, Mehta AJ, Schwartz JD, et al. Genome-Wide Analysis of DNA Methylation and Fine Particulate Matter Air Pollution in Three Study Populations: KORA F3, KORA F4, and the Normative Aging Study. Environ Health Perspect 2016;124: 983–90.
[51] Zhang Z, Geskus RB, Kattan MW, Zhang H, Liu T. Nomogram for survival analysis in the presence of competing risks. Ann Transl Med 2017;5:403.
[52] Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 2012;13:86.
[53] 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.
[54] Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36:1–48.
[55] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57:289–300.
[56] Xu J, Murphy SL, Kochanek DD, Bastian B, Arias E. Division of Vital Statistics. Deaths: final data for 2016. U.S. Department of Health and Human Services; 2018. p. 1–75.
[57] Kennedy BK, Berger SL, Brunet A, et al. Geroscience: linking aging to chronic disease. Cell 2014;159:709–13.
[58] Valensi P, Lorgis L, Cottin Y. Prevalence, incidence, predictive factors and prognosis of silent myocardial infarction: a review of the literature. Arch Cardiovasc Dis 2011;104:178–88.
[59] https://www.nia.nih.gov/health/stroke, accessed on 10/12/2020.
[60] Yousufuddin M, Young N. Aging and ischemic stroke. Aging (Albany NY) 2019; 11:2542.
[61] https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS, accessed on 10/12/2020.
[62] Zhong J, He Y, Chen W, Shui X, Chen C, Lei W. Circulating microRNA-19a as a potential novel biomarker for diagnosis of acute myocardial infarction. Int J Mol Sci 2014;15:20355–64.
[63] Gao F, Kataoka M, Liu N, et al. Therapeutic role of miR-19a/19b in cardiac regeneration and protection from myocardial infarction. Nat Commun 2019;10:1–15.
[64] Trembley MA, Quijada P, Agullo-Pascual E, et al. Mechanosensitive gene regulation by myocardin-related transcription factors is required for cardiomyocyte integrity in load-induced ventricular hypertrophy. Circulation 2018;138:1864–78.
[65] Mokalled MH, Carroll KJ, Cenik BK, et al. Myocardin-related transcription factors are required for cardiac development and function. Dev Biol 2015;406:109–16.
[66] Joehanes R, Just AC, Marioni RE, et al. Epigenetic Signatures of Cigarette Smoking. Circ Cardiovasc Genet 2016;9:436–47.
[67] Zhang Z. Survival analysis in the presence of competing risks. Ann Transl Med 2017;5:47.
[68] Bush M, Simpson Jr. RJ, Kucharska-Newton A, Fang G, Sturmer T, Brookhart MA. Approaches to Address Premature Death of Patients When Assessing Patterns of Use of Health Care Services After an Index Event. Med Care 2018;56:619–25.
[69] Kamaraju S, Shi Y, Smith E, Nattinger AB, Laud P, Neuner J. Are aromatase inhibitors associated with higher myocardial infarction risk in breast cancer patients? A Medicare population-based study. Clin Cardiol 2018;42:93–100.
[70] Hu WS, Lin CL. Acute critical illness and cancer risk: implications from a nationwide population based study in Asia. Int J Cardiol 2018;270:319–23.
[71] de Chickera SN, Bota SE, Kuwornu JP, et al. Albuminuria, Reduced Kidney Function, and the Risk of ST - and non-ST-segment-elevation myocardial infarction. J Am Heart Assoc 2018;7:e009995.
[72] Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. N Engl J Med 2017;376:2513–22.