

<https://doi.org/10.1038/s43856-025-01233-7>

Comprehensive cross-sectional and longitudinal comparison of sixteen markers of biological aging from the Berlin Aging Study II

Check for updates

A list of authors and their affiliations appears at the end of the paper

Abstract

Background The disproportionate increase in lifespan compared to healthspan over the past decades results in a growing proportion of life marked by diseases, even if incidence rates are falling in some cases. However, not everyone ages at the same pace and some people remain in good health and preserve physical and cognitive function into old age. To quantify inter-individual differences in the biological aging process, numerous indicators of biological age have been developed.

Methods In this study, we analyzed 16 measures of biological aging including epigenetic clocks, proteomics clock, telomere length, and SkinAge, laboratory composite markers (BioAge, Allostatic Load), psychological aging, and Brain Age. These age markers were evaluated cross-sectionally as well as longitudinally in the context of age-associated outcomes covering frailty, mobility, cognitive function, depressive symptoms, autonomy in daily life, nutrition, morbidity, and chronic disease in participants of the Berlin Aging Study II (BASE-II).

Results Here, we analyze longitudinal data from 1083 participants (mean age of 68.3 years at baseline, 52% women) with an average follow-up period of 7.4 years. Allostatic Load Index and DunedinPACE show the strongest and most consistent cross-sectional and longitudinal associations with age-associated phenotypes. Furthermore, both biomarkers individually increase the accuracy of a logistic regression model trained to predict incident cases of Metabolic Syndrome, high cardiovascular risk (Lifes's Simple 7) as well as incident frailty (Fried's frailty index) 7.4 years after baseline examination by up to 24 percentage points.

Conclusions Our findings support the previously shown distinction between indicators of aging and provide a comprehensive overview of their individual strengths and weaknesses in the context of wide variety of age-associated phenotypes.

Plain language summary

People are living longer, but not everyone ages in the same way. Some people stay healthy well into old age, while others develop health problems earlier. To understand these differences and predict future health risks, various measures of biological aging have been developed, but comparisons of these measures in the same group of people are scarce. In this study, we compared 16 of these measures in more than 1000 older adults over seven years to see which aging markers best predict future health problems. We found that two measures, Allostatic Load and DunedinPACE, stood out as the most reliable for identifying people at risk of future problems such as frailty and cardiovascular risk. These findings could help physicians and scientists better understand aging and improve early detection of health problems in older adults.

Advances in healthcare, hygiene, lifestyle, and housing have led to an increase in the average lifespan in many countries over the past decades¹. However, this expansion in lifespan was not matched by the increase in healthspan, the time spent before the onset of chronic disease or age-associated impairments²⁻⁴. At the same time, individual aging trajectories vary, and this heterogeneity increases with advancing age. While some individuals maintain good physical and mental health into old age, others

show an early onset of chronic disease and impairment⁵. The geroscience hypothesis states that interventions targeting the biological process of aging may result in an increase of healthspan and prevent or at least postpone the onset of chronic disease⁶. Therefore, interventions that slow down or even reverse biological aging processes are of high individual but also societal interest⁷ but traditionally require long follow-up periods^{3,8}. To test interventions in a cost- and time-effective manner, the validation of markers of

✉ e-mail: valentin.vetter@charite.de; ilja.demuth@charite.de

aging that “either alone or in a composite predict biological age”³³ is needed. Although extensive efforts have been made to identify and develop markers that can quantify inter-individual differences in biological age, no consensus has been reached^{15,9–11}. This might be partly due to the differences in the conceptual meaning and the underlying framework of aging they aim to quantify and the methodological approach used. For example, aging clocks based on “-omics” data, like the proteomic clock or the first-generation epigenetic clocks, were trained to predict chronological age, and the difference between the predicted and the actual chronological age was shown to bear biological meaning¹². In contrast, more recent versions of the epigenetic clocks are trained to predict age- and mortality-dependent biological changes or the so-called Pace of Aging^{13,14}, and thus differ fundamentally in their underlying conceptualization. Other markers, such as BioAge¹⁵ and Allostatic Load¹⁶, are markers that aggregate laboratory measures associated with multiple organ systems into one composite biomarker of aging. Consistent with findings from our and other groups is that different indicators of aging only show a low to moderate association with each other^{5,9,17–25}. This may result from the fact that the indicators were derived from different aging domains using varying methods. Whilst existing indicators of aging are individually well-evaluated, comparisons of different markers within the same study population and their relation to age-associated phenotypes are scarce^{17,23}. Further, by focusing on selected health outcomes or high-risk populations, these studies limit our understanding of the broader predictive validity of aging markers in the general population.

To close this gap, we evaluate 16 markers of aging by investigating their relationship to a wide range of physical and mental age-associated health phenotypes in 1083 participants of the Berlin Aging Study II (BASE-II), based on both cross-sectional and longitudinal data. To assess their potential for clinical application, we additionally investigate the ability of each indicator to predict functional, physical and cognitive impairment after a mean follow-up period of 7.4 years. By doing so, we aim to provide a comprehensive overview of their individual strengths and weaknesses, in order to sharpen each biomarker’s individual profile, and contribute to a roadmap for their effective and targeted use in research and ultimately in clinical practice.

Our analyses reveal individual strengths and weaknesses of the investigated markers. DunedinPACE and Allostatic Load Index show the overall strongest and most consistent associations with the investigated outcomes. Both markers improve the discrimination of a basic clinical model developed to predict incident cases of Metabolic Syndrome, high cardiovascular risk (LS7), and frailty after a 7.4-year follow-up period by up to 24 percentage points.

Methods

Study population

BASE-II is an observational longitudinal study aiming at the identification of factors that predict and shape healthy aging trajectories²⁶. Participants were recruited through the Max Planck Institute for Human Development’s participant pool in Berlin, as well as through advertisements in local newspapers and public transportation networks. As a convenience sample, baseline BASE-II participants were characterized by higher education and better self-reported health status than the general population of Berlin and Germany²⁶. An above-average objective health status of BASE-II participants has been well documented, e.g., for diabetes mellitus type 2²⁷, cardiovascular health²⁸ and frailty¹⁹. Participants between the ages of 60 and 80 years (older cohort) and 20 and 35 years (younger cohort, not analyzed in this study) were eligible for recruitment. Men and women were recruited in equal numbers. The baseline examination (T0) of 1671 older participants (medical part) was conducted between 2009 and 2014²⁶. A follow-up assessment (T1) as part of the GendAge study was conducted between 2018 and 2020²⁹ on average 7.4 years after baseline (SD: 1.5 years, range: 3.9 to 10.4 years). Only participants who had provided information on outcomes of interest at baseline as well as at follow-up were included in the final sample ($n = 1083$, Supplementary Fig. 1), to allow comparison of the aging markers in the same cohort of individuals. Of the 588 participants who dropped out

between baseline and follow-up, 126 were confirmed to have died. Reasons for dropout among the remaining 462 participants were not systematically evaluated. Differences in the variables analyzed in this study between the main study sample and participants who dropped out are small, and we do not expect that loss to follow-up substantially altered our findings (Supplementary Data 1).

All participants gave written informed consent. All assessments at baseline and follow-up were conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Charité—Universitätsmedizin Berlin (approval numbers EA2/029/09, EA2/144/16, and EA2/224/21) and were registered in the German Clinical Trials Registry as DRKS00009277 (BASE-II) and DRKS00016157 (GendAge). The Ethics Committee of the Max Planck Institute for Human Development approved the procedure, and the Ethics Committee of the German Society for Psychology (DGPs) additionally approved of the MRI protocol. This manuscript was created in accordance with the STROBE guidelines³⁰.

Variables

The Horvath³¹, Hannum³², PhenoAge³³, GrimAge³⁴, and DunedinPACE¹⁴ epigenetic clock algorithms were applied to derive DNA methylation age (DNAmAge) from methylation data measured with the “Infinium MethylationEPIC”, version 1, array (Illumina, Inc., USA). Methylation data for the 7-CpG epigenetic clock¹⁸ was obtained through Single Nucleotide Primer Extension (SNuPE)^{18,35}. DNAmAge acceleration (DNAmAA) was calculated as the unstandardized residuals of a leukocyte cell-count adjusted linear regression analysis of DNAmAge on chronological age. Proteomics Age was derived from 248 proteins measured by liquid chromatography-mass spectrometry (LC-MS). Proteomics Age acceleration (ProteomicsAA) was calculated similarly to the DNAmAA as residuals from a linear regression of Proteomics Age on chronological age. Telomere length was assessed through quantitative real-time PCR (rLTL³⁶) as well as estimated by the algorithm proposed by Lu and colleagues³⁷ from methylation data (DNAmTL). The average results across three raters evaluating the number of lentigenes from photos of the participants’ skin were used to derive SkinAge¹⁵. BioAge is a composite score that aggregates 13 routine laboratory parameters that were identified for their association with mortality¹⁵. The Allostatic Load Index (ALI) was computed by awarding a point to every participant within the high-risk quartile of selected variables using the approach described by Seeman and colleagues¹⁶. It additionally incorporates information about intake of relevant medication to include successfully treated and thereby masked dysregulation in its calculation^{38,39}. Subjective Felt Age (SFA) was calculated as proportional discrepancy score from self-reported “felt age” and chronological age⁴⁰. Subjective Life Expectancy (SLE) is calculated as the difference between the age participants expect to live to and their chronological age at the time of assessment. Subjective Health Expectancy (SHE) is the difference between the age participants expect to remain healthy and their chronological age at the time of assessment. Both SLE and SHE were adjusted for chronological age⁴¹. In addition to these markers, BrainAge was available in a subgroup of $n = 255$ BASE-II participants who underwent Magnetic Resonance Imaging (MRI) in a 3-Tesla Siemens Magnetom Trio scanner. BrainAge was calculated using a model trained on participants from the UK Biobank⁴². Except for Proteomics Age, all other indicators of aging were individually investigated in BASE-II before^{15,18,19,24,28,43–49}.

Outcome measures investigated in this study include two measures of frailty. The Fried Frailty Phenotype (Fried FI) includes information on unintended weight loss, exhaustion, weakness, slow walking speed, and low physical activity⁵⁰. The SPRINT-BASEed frailty index (SP FI) used in this study is an adapted version⁵¹ of the deficit-based measure developed by Pajewsky and colleagues⁵². Finger Floor Distance (FFD) was measured in centimeters. The Tinetti Mobility Test assesses mobility and balance through a series of simple tasks that are rated by the test supervisor⁵³. Mini Mental State Examination (MMSE) is a well-established interviewer-administered instrument that was used to assess cognitive impairment⁵⁴.

Additionally, cognitive performance, specifically processing speed, was measured by the Digit Symbol Substitution Test (DSST), for which participants were asked to match symbols to numbers according to a given key⁵⁵. Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D)⁵⁶. Independence during everyday life and the ability to perform tasks without help was measured using the Activities of Daily Living questionnaire (ADL, “Barthel Index”)⁵⁷. The nutritional status was assessed using results from a short questionnaire and the measured circumference of the upper arm and the calf (Mini Nutritional Assessment, MNA)⁵⁸. Type 2 Diabetes (T2D) was diagnosed based on the criteria defined by the American Diabetes Association (ADA) guidelines⁵⁹. Diabetes-associated complications were quantified using the composite score developed by Young and colleagues^{27,60}. The Morbidity Index (MI) was calculated to assess the overall morbidity burden of the BASE-II participants by adapting the approach first described by Charlson and colleagues⁶¹. The Systematic Coronary Risk Evaluation assessment (SCORE2⁶²) and the SCORE2-OP (for participants >70 years, hereafter referred to as SCORE2⁶³) were calculated according to the recommendations in the respective publications to assess the risk for cardiovascular events. An adapted version^{28,64} of the Life’s Simple 7 (LS7⁶⁵) was calculated to quantify modifiable cardiovascular risk factors. Metabolic Syndrome (MetS) was diagnosed using the definition suggested by the American Heart Association/International Diabetes Federation/National Heart, Lung, and Blood Institute criteria 2009⁶⁶. The outcome variables described above were chosen because they represent specific health aspects that we analyze as downstream effects of the biological aging process quantified by the investigated markers of aging. However, in other contexts, some of the outcome variables themselves could also be investigated as aging markers and vice versa.

Confounding variables included in the regression models were chronological age (years), sex, alcohol consumption (g/d), and nicotine consumption (packyears), which were assessed during 1:1 interviews with trained study personnel. Body Mass Index (BMI) was calculated as kg/m² using measurements from an electronic measuring station (seca 763, SECA, Germany). Genetic ancestry was quantified using the first four components from a principal component analysis on genome-wide single-nucleotide polymorphism genotyping data⁶⁷. Details on the variables investigated in this study are described in the Supplementary Material.

Statistics and reproducibility

Statistical analyses and visualizations were conducted using R (version 4.3.2)⁶⁸. Missing values were imputed by multiple imputation using the mice package (for detailed information, please see the Supplementary Methods). In this study, $n = 1083$ participants of BASE-II were analyzed. BrainAge was available in a subgroup of 255 BASE-II participants. To facilitate comparison with all other markers, a separate multiple imputation procedure including all other variables was applied in this subsample. Descriptive statistics of the first imputed dataset are presented in Tables 1 and 2. Descriptive statistics of the original (non-imputed) dataset are shown in Supplementary Data 2. Intercorrelation plots for markers of aging and outcome variables are presented in Supplementary Figs. 2 and 3. The change in outcome variables was assessed by subtracting the outcome value at T0 from the value documented for T1. The relationship between markers of aging and outcome variables was assessed by logistic and linear regression models. In line with previous publications in the field as well as recommendations for increased quality in biomarker validation, we report results from an unadjusted, a minimally adjusted (chronological age and sex), and fully adjusted model⁶⁹. The longitudinal minimally and fully adjusted model is additionally adjusted for the outcome variables at T0. Since we were mainly interested in the raw predictive ability of the biomarkers (diagnostically as well as prognostically), we focus on the description of

Table 1 | Descriptive statistics of markers of aging at baseline in older participants of the BASE-II ($n = 1083$)

Variables	Mean (SD)	Min	Max
Chronological age	68.27 (3.49)	60.16	84.63
7-CpG DNAmAA	-0.02 (6.94)	-22.93	22.33
Horvath DNAmAA	0.08 (4.22)	-11.64	14.45
Hannum DNAmAA	-0.04 (3.43)	-10.9	19.15
PhenoAge DNAmAA	0.08 (4.65)	-16.88	14.28
GrimAge DNAmAA	0.07 (3.12)	-8.52	10.2
DunedinPACE	1.01 (0.11)	0.58	1.42
DNAmTL	6.97 (0.18)	6.26	7.44
rLTL	1.15 (0.23)	0.17	1.92
SkinAge	1.71 (0.90)	0	3
ProteomicsAA	-0.02 (1.34)	-4.18	5.11
BioAge	0.17 (5.91)	-16.37	20.21
Allostatic Load	3.98 (2.40)	0	12
Subj. age ^a	0.09 (0.08)	-0.13	0.41
Subj. life expectancy	15.89 (7.68)	0	47
Subj. health expectancy	11.47 (7.06)	0	46
BrainAge ^b	-0.06 (3.23)	-9.03	8.03

^aSubj. age = relative subjective age adjusted by chronological age.

^bBrainAge was available in a subgroup of 255 participants.

the unadjusted model in the main text. Statistical significance of the differences between the basic risk prediction models with and without the respective marker of aging was assessed using the *roc.test* function (pROC package). To adjust for multiple testing, a Bonferroni correction was applied to the main analyses. As cross-sectional and longitudinal analyses were performed, we set our level for statistical significance at $\alpha = 0.05/(2^{*}16 \text{ markers} * 16 \text{ outcomes}) = 0.0001$. To facilitate comparison between markers and outcome measures on differing scales, all values were z-transformed using R’s *scale* function prior to the regression analyses, and standardized effect measures are reported. Dichotomization of continuously scaled variables used in the sensitivity analyses and the impairment prediction models was done using validated cut-off values if available (Supplementary Data 3). The MMSE and ADL were excluded from analyses that examined the dichotomized variables due to the very limited number of participants showing impairment according to these instruments. For the main analyses, sex-stratified subgroup analyses were performed as suggested by Moqri and colleagues⁶⁹. No age-group stratified analyses are shown due to the already narrow age range in BASE-II⁶⁹. Further information on the statistical analyses can be found in the Supplementary Material.

Results

Participants

In this study, 1083 BASE-II participants who provided information at baseline and follow-up on average 7.4 years later were analyzed with respect to cross-sectional and longitudinal associations of 16 markers of aging (Supplementary Fig. 4) with a range of age-associated outcomes. Mean chronological age at baseline was 68.3 years (SD: 3.5 years), and 52% were women (Table 1, Supplementary Data 1). Frequency of impairment in the analyzed variables is shown in Supplementary Data 4. All age-associated outcome variables differed statistically significantly between examinations with the exception of FFD, Falls, MMSE, ADL, CES-D, and LS7 (Table 2). As described in BASE-II before²⁴, correlations between markers of aging derived from different data sources and domains were moderate to low with $r \leq$

Table 2 | Descriptive statistics of outcome variables at baseline (T0) and follow-up (T1) examination (n = 1083)

Variable	T0				T1				Mean Diff.	SMD	p-value
	Mean, n	SD, %	Min	Max	Mean, n	SD, %	Min	Max			
Chronological age	68.272	3.489	60.2	84.6	75.62	3.768	64.9	94.1	7.348	2.024	<2.2e-16
Fried frailty	0.364	0.61	0	3	0.763	0.878	0	4	0.399	0.528	<2.2e-16
Frailty SPRINT-BASEed	0.132	0.063	0.008	0.44	0.16	0.075	0	0.523	0.028	0.411	<2.2e-16
Finger floor distance	9.746	11.839	0	52	10.576	12.673	0	159	0.829	0.068	0.005
Tinetti score	27.629	1.691	0	28	26.572	2.668	7	28	-1.056	-0.473	<2.2e-16
Falls (past 12 months)	336	0.31			295	0.272					0.043
MMSE	28.59	1.335	22	30	28.578	1.524	18	30	-0.012	-0.008	0.827
DSST	45.001	8.301	5	74	40.801	8.674	2	73	-4.2	-0.495	<2.2e-16
CES-D	7.285	6.613	0	36	7.17	6.165	0	36	-0.115	-0.018	0.526
ADL	99.003	3.72	5	100	98.698	4.086	15	100	-0.305	-0.078	0.046
MNA	27.383	1.756	19	30	26.595	2.158	16	30	-0.788	-0.401	<2.2e-16
DCSI	0.777	1.163	0	7	1.264	1.486	0	8	0.488	0.365	<2.2e-16
Morbidity index	0.975	1.208	0	7	1.432	1.542	0	9	0.457	0.33	<2.2e-16
SCORE2	10.809	4.222	4	32	16.073	6.16	5	47	5.264	0.997	<2.2e-16
LS7	8.609	1.978	2	14	8.44	1.86	3	14	-0.17	-0.088	0.001
T2D (diagnosed)	127	0.117			185	0.171					1.874e-09
Metabolic syndrome (diagnosed)	388	0.358			491	0.453					5.018e-11

Two-sided paired *t*-test and McNemar’s test were used to assess the statistical significance of differences between timepoints. SMD Standardized Mean Difference, LS7 Life’s Simple Seven, MMSE mini mental state examination, SCORE2 systematic coronary risk evaluation 2, DCSI diabetes complications severity index, CES-D Center for Epidemiologic Studies Depression Scale, MNA mini nutritional assessment, DSST digit symbol substitution test, Tinetti Tinetti test, T2DM type 2 diabetes mellitus.

[0.31] while higher correlations were found between markers which were calculated based on the same data (e.g., epigenetic clocks, $r \leq 0.49$) or which were closely related to each other (SHE and SLE, $r = 0.73$, Supplementary Fig. 2).

Cross-sectional association between markers of aging and age-associated outcomes at baseline

Cross-sectional associations of markers of aging with continuously and categorically scaled age-associated outcomes at baseline were investigated using linear and logistic regression models. No statistically significant association (Bonferroni adjusted significance level defined at $\alpha = 0.0001$) with the outcome variables was found for the first-generation epigenetic clocks, PhenoAge DNAmAA, SkinAge, ProteomicsAA, and BrainAge. In contrast, GrimAge DNAmAA, DunedinPACE and both laboratory composite markers (BioAge, ALI) were statistically significantly associated with at least one of the two cardiovascular health scores (LS7, SCORE2, all $p > 0.000018$). In addition, ALI and DunedinPACE were associated with disease-related outcomes like MI ($p = 7.3 \times 10^{-6}$), diagnosed T2D ($p = 5.7 \times 10^{-6}$), and MetS ($p = 1.1 \times 10^{-10}$, Supplementary Data 5). After correction for multiple testing, both TL measures were statistically significantly associated with FFD and SCORE2, and all three psychological aging markers (SFA, SLE, SHE) were statistically significantly associated with SPRINT-BASEed FI. Overall, the strongest and most robust results after covariate adjustment were found for ALI and DunedinPACE. All results from the unadjusted cross-sectional linear and logistic regression analyses, as well as minimally (age, sex) and fully adjusted models (age, sex, genetic ancestry, alcohol intake, smoking, BMI) are displayed in Fig. 1, Supplementary Fig. 5, and Supplementary Data 5 and 6.

As a sensitivity analysis, we report on the results of logistic regression models of dichotomized continuously scaled variables on the markers of aging (Supplementary Data 6) and sex-stratified subgroup analyses (Supplementary Data 5 and 6). All analyses were repeated in the subsample of participants for which BrainAge was available (Supplementary Data 7 and 8) to allow direct comparison of the results of BrainAge with the other markers in the same individuals.

Longitudinal associations between markers of aging at baseline and age-associated outcomes after 7.4 years of follow-up

To explore the markers’ longitudinal association with the investigated outcomes, we performed regression analyses of outcomes measured at follow-up on markers assessed at baseline. Similarly to the cross-sectional analyses, no statistically significant associations with age-associated outcomes were found for measures derived from first-generation epigenetic clocks, PhenoAge, SkinAge, ProteomicsAA, and BrainAge. As already observed in the cross-sectional analyses, after correction for multiple testing, the strongest and most consistent associations were found for DunedinPACE and the ALI. Interestingly, all three psychological aging markers were associated with Fried’s frailty phenotype in addition to the association with the SPRINT-BASEed frailty index that was already observed in the cross-sectional analyses at baseline (Fig. 2, Supplementary Data 9). The longitudinal results of the dichotomized as well as of the binary outcome variables can be found in Supplementary Data 10 and Supplementary Fig. 6. The results from the analyses of the subgroup of participants who provided BrainAge data can be found in Supplementary Data 11 and 12. Additionally, we investigated the change in the continuously scaled outcome variables by regressing the change in the outcome variables on the markers of aging at baseline, which can be found in Supplementary Data 13. Similar to the other longitudinal linear regression models, ALI showed the most frequent and consistent statistically significant associations across outcomes and models.

Prediction of incident cases in outcome variables over a 7.4-year follow-up period

To promote translation into clinical practice and to evaluate the markers in a possible use-case scenario, we evaluated the markers’ ability to predict the incidence of impairment in the analyzed assessments, as well as incident cases of the investigated diseases and frailty over the average follow-up time of 7.4 years. Each marker-outcome-combination was evaluated in two scenarios. First, we simulated the case in which no other clinical information is available and assessed the markers raw predictive value. In a second step, we investigated the additional value that each marker adds to a basic prediction model.

Fig. 1 | Standardized regression coefficients of unadjusted linear regression analyses of outcome variables on markers of aging in 1083 participants from the BASE-II at baseline. BrainAge was available in a subgroup of $n = 255$ participants. *Bonferroni-corrected statistical significance ($p < 0.0001$). LS7 Life's Simple Seven, SCORE2 systematic coronary risk evaluation 2, MI Charlson's morbidity index, DCSI diabetes complications severity index, MNA mini nutritional assessment, DSST digit symbol substitution test, Tinetti Tinetti test, FFD finger floor distance, SP FI SPRINT-BASEed frailty phenotype, Fried FI Fried's frailty phenotype, MetS metabolic syndrome, T2DM type 2 diabetes mellitus.

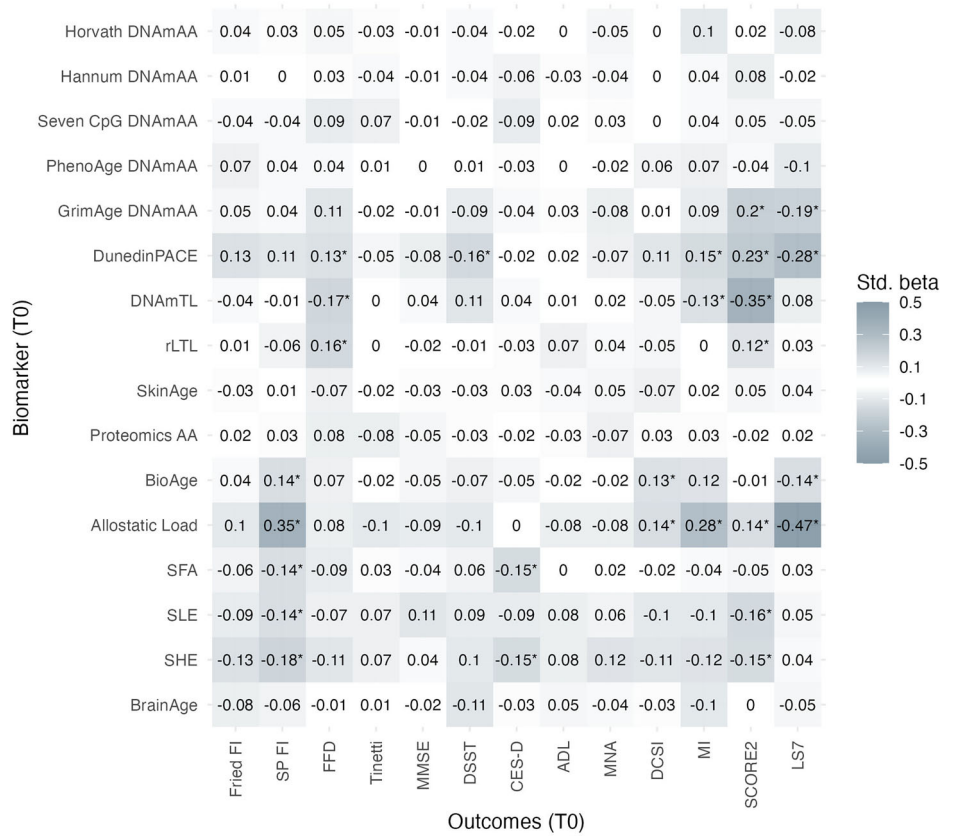
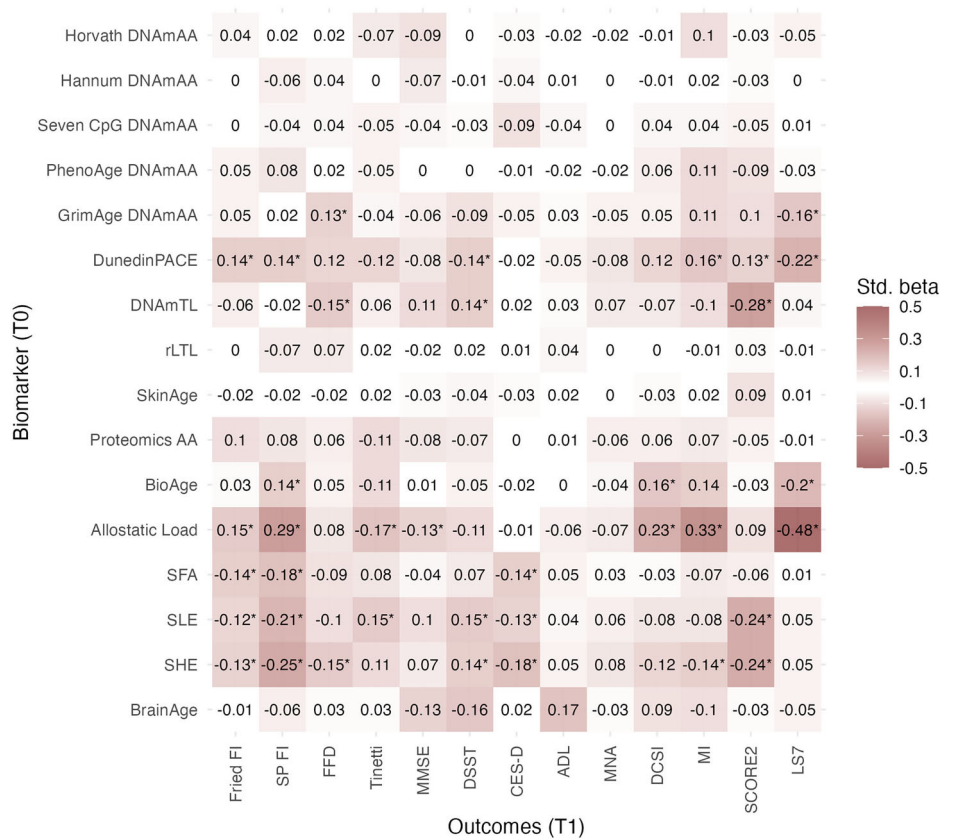


Fig. 2 | Standardized regression coefficients of unadjusted longitudinal linear regression analyses of outcome variables at T1 on markers of aging at T0 in 1083 participants from the BASE-II. BrainAge was available in a subgroup of $n = 255$ participants. *Bonferroni-corrected statistical significance ($p < 0.0001$). LS7 Life's Simple Seven, SCORE2 systematic coronary risk evaluation 2, MI Charlson's morbidity index, DCSI diabetes complications severity index, MNA mini nutritional assessment, DSST digit symbol substitution test, Tinetti Tinetti test, FFD finger floor distance, SP FI SPRINT-BASEed frailty phenotype, Fried FI Fried's frailty phenotype, MetS metabolic syndrome, T2DM type 2 diabetes mellitus.



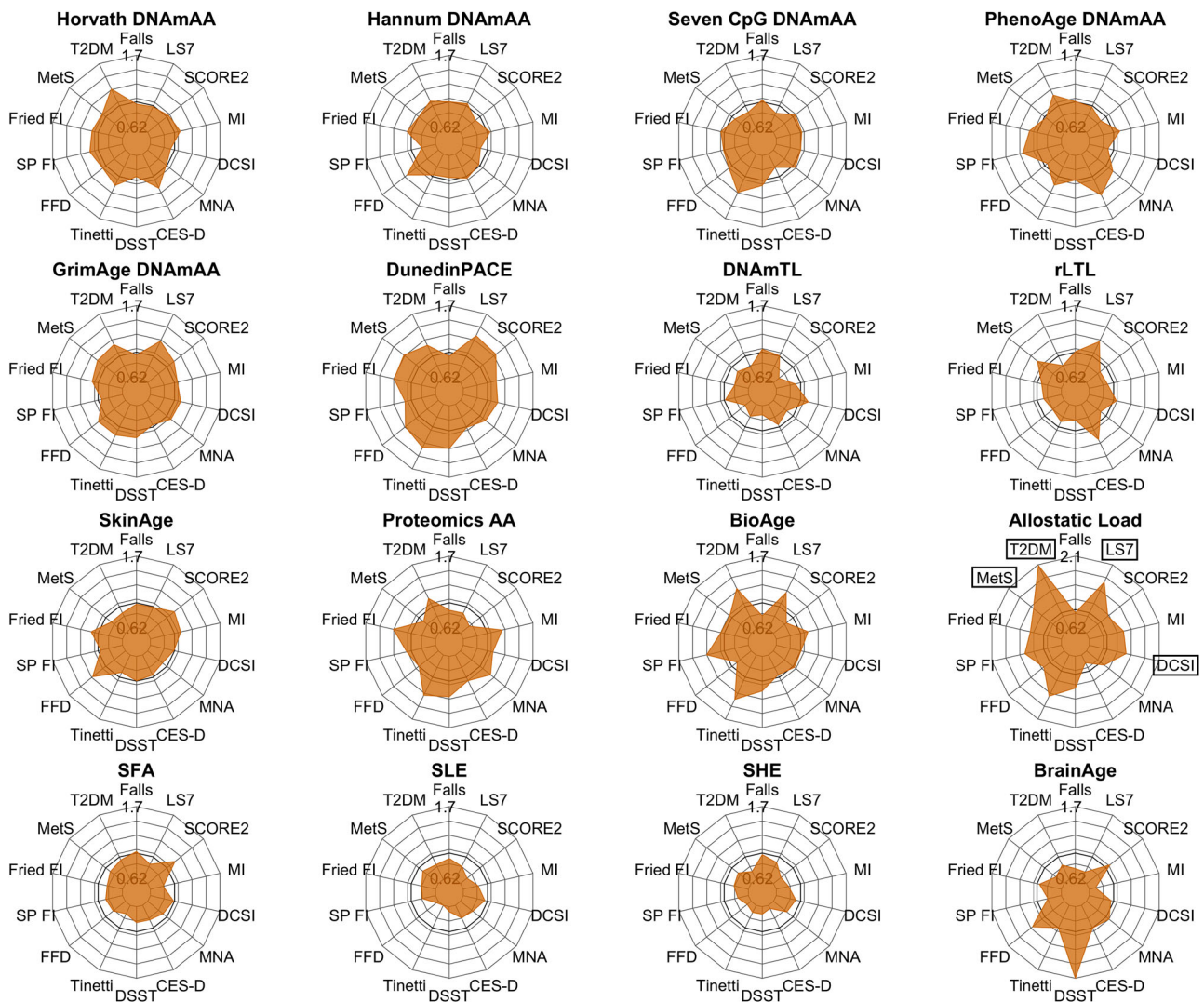


Fig. 3 | Radar plots showing the results of unadjusted logistic regression analyses of incident cases of analyzed outcome variables at T1 on markers at T0. Participants with prevalent cases at T0 were excluded from this analysis. Due to the multiple imputation procedure, the sample sizes between each imputed dataset differ slightly. The range of sample sizes is indicated in Supplementary Data 18. Due to very strong effect sizes observed for Allostatic Load, the scale of this plot was adjusted to increase readability of the other plots. Therefore, the individual axis limits are displayed for each plot individually. Variable names of associations that were

statistically significant after Bonferroni correction for multiple testing ($p < 0.0001$) are displayed within a black box. LS7 Life’s Simple Seven, SCORE2 systematic coronary risk evaluation 2, MI Charlson’s morbidity index, DCSI diabetes complications severity index, MNA mini nutritional assessment, DSST digit symbol substitution test, Tinetti Tinetti test, FFD finger floor distance, SP FI SPRINT-BASEd frailty phenotype, Fried FI Fried’s frailty phenotype, MetS metabolic syndrome, T2DM type 2 diabetes mellitus.

Of all analyzed biomarkers, Allostatic Load was the only one to be statistically significantly (after multiple testing correction) associated with incident impairment at follow-up (Fig. 3). Specifically, higher Allostatic Load was associated with increased odds for a diagnosis of T2D and MetS and its addition to the basic prediction model improved the AUC of the model by 24 and 18 percentage points (DeLong’s test, $p < 0.00001$, Fig. 4). Additionally, an increase in one SD of Allostatic Load at baseline increased the odds for the incidence of at least one diabetes-associated complication at T1 (DCSI) 1.4-times ($p < 0.0001$) and the inclusion of Allostatic Load to a basic prediction model increased the AUC by 7 percentage points ($p < 0.00001$). Odds for an incident classification as “average” or “inadequate” based on the LS7 criteria were 1.7-fold for every 1-SD increase in Allostatic Load ($p = 0.00002$, Fig. 4) at baseline and it increased the accuracy of the basic prediction model by 20 percentage points ($p < 0.00001$, Fig. 4). Interestingly, while Allostatic Load was not statistically significantly associated with the outcome in prediction models for Fried’s frailty index and MI, its inclusion still improved the respective prediction models statistically

significantly by 4 and 10 percentage points ($p < 0.00003$), respectively. To investigate the robustness of this association, we conducted a series of sensitivity analyses with modified versions of the ALI. If the analyzed outcome or variables used in its definition was also part of the ALI, as is the case for MetS, T2D, LS7, SCORE2, the respective variables were individually excluded from the calculation of the modified ALI. As expected, effect sizes of the association were stronger for the original ALI in comparison to the modified ALI scores (Supplementary Data 14–16). However, in almost all cases, the associations remained statistically significant, which increases confidence that ALI is a robust marker for biological age. Similarly, DunedinPACE did not show individual statistically significant associations in the models, but its addition to the basic prediction model increased the AUC of models predicting diagnosis of MetS and impairment in Fried’s frailty index, and LS7 by 5, 4, and 8 percentage points ($p < 0.00006$, Fig. 4). In case of LS7, GrimAge DNAmAA improved the prediction model by 5 percentage points ($p < 0.0001$, Fig. 4). As a sensitivity analysis, we recalculated all prediction models with GrimAge2 instead of GrimAge and

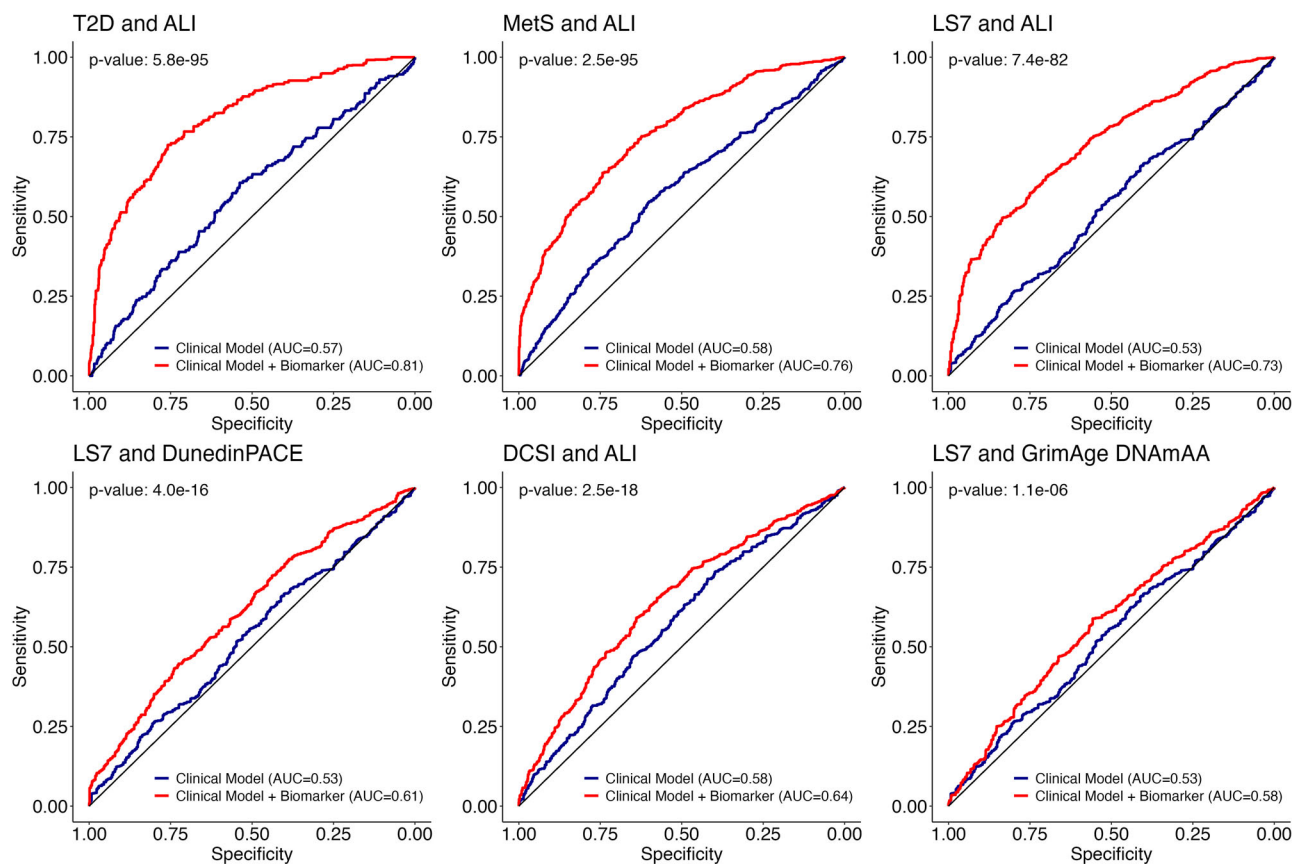


Fig. 4 | ROC curves illustrating the sensitivity and specificity of selected logistic regression models predicting impairment in age-associated phenotypes on average 7.4 years after marker assessment. The predictive performance of a basic clinical model including age and sex (blue) is shown compared to a prediction model that is extended by the respective markers of aging (red). The marker-outcome combinations presented here were selected to highlight the most compelling results

from the three most promising markers (ALI, DunedinPACE, and GrimAge). A complete list of all AUC values can be found in Supplementary Data 18. P-values for the difference between prediction models were calculated using the approach described by DeLong⁷² as part of the roc.test function (pROC package). After Bonferroni correction, the level of statistical significance was defined at $p < 0.0001$.

found no or small differences (≤ 1.6 percentage points in the AUC) between the two versions (Supplementary Data 17). A very high AUC (AUC = 0.935) was observed for the clinical prediction model with respect to identifying incident high cardiovascular risk defined by the SCORE2 algorithm. This prediction was significantly improved by 1 percentage point after adding TL measured using qPCR ($p < 0.00003$). The complete results for all marker-outcome combinations are shown in Fig. 3 and Supplementary Data 18.

Discussion

In this study, we investigated 16 markers of aging in 1083 older participants from the Berlin Aging Study II (BASE-II). All markers were cross-sectionally and longitudinally analyzed in the context of a wide range of age-associated outcome variables representing different aspects of aging, including frailty, mobility, cognitive function, depressive symptoms, autonomy, nutrition, overall health, and chronic disease. Additionally, we investigated the markers' ability to predict incident impairment in the age-associated outcomes over the average follow-up period of 7.4 years.

Our analyses showed that GrimAge DNAmAA and DunedinPACE performed best with respect to their association with cardiovascular health and cognitive capacity. Allostatic Load seems to represent frailty and overall morbidity-associated variables (MI, T2D and MetS). Longitudinally, the subjective psychological markers were associated with depressive symptoms (CES-D) and cognitive functioning (DSST, except SFA). Additionally, they were associated with the SPRINT-BASEed frailty index (cross-sectionally and longitudinally) and Fried's frailty index (longitudinally). The strongest effect sizes in cross-sectional and longitudinal analyses were found for ALI

with respect to diagnosed T2D, MetS, LS7, MI and SPRINT-BASEed frailty index. This presumably results from the way it is calculated, as many of the variables included in the ALI variable are closely related to health and are, in some instances, also part of the diagnostic criteria for T2D, MetS, LS7, and SCORE2. However, sensitivity analyses with modified ALI versions that excluded all variables that were also part of the outcomes in question (MetS, T2D, LS7, SCORE2) remained in almost all cases statistically significant (Supplementary Data 14–16). This increases confidence that ALI is indeed a robust marker for biological age.

Subsequently, we investigated a potential use-case of the markers in a clinical context by predicting incident cases of impairment at follow-up by markers assessed on average 7.4 years earlier at baseline. ALI, GrimAge DNAmAA and DunedinPACE improved our basic prediction model in its prediction of incidence of impairment by up to 24 percentage points (Fig. 4). Although this increase in accuracy of the prediction model is expected to be less substantial when the markers are added to more comprehensive and specific prediction models, these results nevertheless illustrate that the markers hold biological information that can substantially improve the prediction of impairment years before it becomes clinically apparent. To estimate how the predictive value of the markers of aging would change in comparison to more comprehensive clinical prediction models, we calculated a sensitivity analysis that investigated the value added by the markers of aging to an extended clinical model (age, sex, BMI, smoking, and alcohol) (Supplementary Data 19). Although this did not lead to substantial changes in most marker's interpretation, the added value of ALI with respect to the AUC of the prediction models for MetS, T2D, and LS7 was reduced by up to

17 percentage points compared to our original analysis. This suggests that future studies need to compare individual markers of aging with more comprehensive and specific clinical prediction models to fully understand how the markers of aging can contribute to incident case prediction in a clinical setting. Generally, we found more statistically significant results and stronger effect sizes in the longitudinal analyses compared to the cross-sectional results. This was observable especially for DunedinPACE, Allostatic Load and the psychological aging markers. Markers of aging are expected to depict the underlying aging processes, which are anticipated to happen prior to clinical manifestations. Therefore, it is expected that an accelerated aging process that did not (yet) exceed the individual mechanisms of resilience and therefore did not manifest through clinical phenotypes would be trackable through markers. The results presented here suggest that the respective markers appear to recognize an acceleration in the underlying aging process that was (at least at baseline) not advanced enough or was still compensated by the physiological and cognitive coping mechanisms to not result in clinically observable phenotypes. However, at follow-up examination, the acceleration in the biological aging process, which was already picked up at baseline by the respective markers, potentially resulted in clinical manifestations.

Another main finding of our study is the in part large differences in effect sizes of associations with the examined outcome variables, which suggests that the analyzed markers capture distinct aspects of the biological aging processes^{9,17,18,25}. This interpretation of our findings is particularly plausible due to the different concepts of biological age that underlie the examined markers. For example, the first-generation epigenetic clocks aim at the prediction of chronological age and the residuals of a regression of chronological on epigenetic age are used to derive the analyzed marker (DNAmAA)^{18,31,32}. Second- and third-generation epigenetic clocks, in contrast, are trained on more complex measures of biological aging calculated from several individual variables^{14,33,34}, and DunedinPACE captures the rate of aging comparable to a speedometer^{14,70}. Other markers, such as the psychological aging markers, rely on the subjective self-assessment of the participants. Composite markers, like Allostatic Load and BioAge, on the other hand, incorporate information from numerous systems to quantify aging, while other measures, such as telomere length and BrainAge, aim at the quantification of age-related biological changes. These differences result in unique strengths and weaknesses of each marker, which in turn define how they could potentially be used in the scientific and clinical context.

As sex-specific differences in the aging process are well known and to be in line with recommendations for the validation of markers of aging⁶⁹, sex-stratified analyses were presented in the Supplementary Material of this manuscript. In some cases, statistically significant findings that were observable for analyses including the whole study sample did not reach statistical significance in the sex-stratified subgroup analyses. This most likely results from the smaller sample sizes of the respective subgroups. On the other hand, in some cases, it can be seen that the association observed in the whole sample is driven mostly by one of the sex-stratified subgroups. For example, the cross-sectional association between ALI and DCSI (logistic regression, Supplementary Data 6) observed for the whole study sample (OR = 1.4, $p = 0.00001$) seems to be mostly driven by men (OR = 1.5, $p = 0.00005$) compared to the subgroup of women (OR = 1.3, $p = 0.02$) in which the association does not reach the Bonferroni adjusted significance level (Supplementary Data 6). Due to the large number of markers and outcomes, as with the main analyses, we refrained from a detailed discussion of all findings from these subgroup analyses.

Our results are indicative that these markers are indeed capable of depicting underlying aging processes and do so with a higher sensitivity than clinical aging measures. Thereby, they could potentially be used to identify participants who are especially prone to future impairment due to an acceleration in one of the biological aging processes long before they show this decline clinically. Furthermore, we want to note that in this study, the markers of aging are analyzed separately with respect to their ability to predict incident cases in the outcome variables. This neglects the possibility of complementary and synergistic effects when combining markers from

different domains. In addition, while the results indicate the promising potential of these markers as screening parameters, the study design does not allow to conclude on any clinical recommendations. Further studies that examine the added value of including these markers in screening processes, ideally in a randomized controlled design, with a meaningful endpoint, are needed. While screening programs are often evaluated with mortality as endpoint, in the aging context, it might be important to analyze other variables that are specific to the outcome of interest and are associated with quality of life and overall health. For example, a reduction of diabetes-associated complications might provide substantial benefit for individuals even if the overall lifespan could not be extended. Therefore, future studies with a focus on adverse outcome prevention are needed to further explore the possibility of utilizing these markers of aging as possible screening markers in a clinical context.

In a previous study by Kuo and colleagues, change in epigenetic age over time was associated with mortality⁷¹. As a sensitivity analysis, we calculated prediction models investigating the change in the two most promising markers of aging, ALI and DunedinPACE, between T0 and T1 (Supplementary Data 20). Interestingly, DunedinPACE at T0 showed better results compared to the longitudinal change in DunedinPACE, which only remained nominally statistically significant in the unadjusted logistic regression model predicting MI. On the other hand, the longitudinal difference in ALI showed stronger effect sizes when predicting falls, T2D, and MetS, but generally less strong or overall, not statistically significant associations for all other outcome variables. One limitation of our study in this context is that the second timepoint for assessing the longitudinal difference was also used to assess incident cases. Future studies with a third timepoint for incident cases assessment might be needed to evaluate the potential of the aging rate to predict outcomes other than mortality.

We want to point out several limitations to this study. First, the participants of this study are above-average health²⁶. Consequently, a comparatively low prevalence of impairment in the functional and cognitive assessments was observed, which could lead to an underestimation of the true effect sizes, and the generalizability of our results to the underlying source population might be limited. Second, the covariates used in the fully adjusted regression models were chosen because of their frequent use in other studies in the field, as well as their known or suspected association with the independent and dependent variables in our regression models. An individual covariate selection for each marker-outcome-combination was outside the scope of this investigation and would also have compromised comparability of effect sizes between markers within this study. Future studies with a stronger focus on causal associations and the mechanistic relationship between markers and outcomes are needed to deepen our understanding. Third, as with any study that reports on a large number of individual statistical tests, multiple testing is an issue. Here, we adjusted the p -values using the Bonferroni correction. While this correction is robust, it also comes with the risk of missing true effects due to its conservative approach. Therefore, it is possible that true associations did not reach statistical significance in our study. Finally, the comparatively small age range in BASE-II might be a reason for the weak correlation between some markers and chronological age, and limits our conclusions to this age group. Future longitudinal studies with a wider age range and repeated marker measurements are needed to further improve our understanding of longitudinal associations between the markers investigated here and clinical aging phenotypes.

Strengths of these analyses include the comparative analysis of a large number of markers derived from a wide variety of aging domains, including epigenetics, proteomics, telomeres, composite markers, psychological markers, SkinAge, and BrainAge. The availability of these variables, in addition to numerous age-associated outcomes in a large longitudinal sample, allowed a comprehensive comparison of these markers cross-sectionally as well as over time. We provide information on the distinct abilities of these markers, which can inform future studies and sharpen their profiles. As the measurement of numerous markers is costly, these analyses allow a targeted selection of markers that can be used in

future studies based on their strength of association with the specific aging domain of interest.

Conclusion

Our comparative analyses of 16 markers of aging in the context of a wide range of age-associated outcome variables highlight the distinction between the investigated domains of aging. Interestingly, markers were more frequently and also more strongly associated with outcomes on average 7.4 years after their assessment compared to cross-sectional analyses, suggesting their sensitivity to biological aging processes that became clinically apparent only much later during the follow-up examination. This finding underscores the potential of these markers for early risk stratification. Namely, ALI and DunedinPACE substantially contributed to the prediction of incident impairment at follow-up examination.

Data availability

Source data for Figs. 1–4 are provided as Supplementary Data 5, 9, 18, and 21. Due to concerns for participants' privacy as well as data protection regulations, BASE-II raw data cannot be made publicly available. Interested investigators are invited to contact the scientific coordinator of BASE-II, Ludmila Müller (lmueller@mpib-berlin.mpg.de), to obtain raw data access for project-specific analyses under an individually negotiated data use agreement. Data requests will be addressed in a timely manner, typically within four weeks of receipt. Additional information can be found on the BASE-II website: <https://www.base2.mpg.de/7549/data-documentation>.

Received: 20 May 2025; Accepted: 30 October 2025;

Published online: 27 March 2026

References

- Oeppen, J. & Vaupel, J. W. *Broken Limits to Life Expectancy*, 1029–1031 (American Association for the Advancement of Science, 2002).
- Crimmins, E. M. Lifespan and healthspan: past, present, and promise. *Gerontologist* **55**, 901–911 (2015).
- Moqri, M. et al. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell* **186**, 3758–3775 (2023).
- Garmany, A., Yamada, S. & Terzic, A. Longevity leap: mind the healthspan gap. *NPJ Regen. Med.* **6**, 57 (2021).
- Jylhava, J., Pedersen, N. L. & Hagg, S. Biological age predictors. *EBioMedicine* **21**, 29–36 (2017).
- Kennedy, B. K. et al. Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).
- Goldman, D. P. et al. Substantial health and economic returns from delayed aging may warrant a new focus for medical research. *Health Aff.* **32**, 1698–1705 (2013).
- Kirkland, J. L. Translating the science of aging into therapeutic interventions. *Cold Spring Harb. Perspect. Med.* **6**, a025908 (2016).
- Hägg, S., Belsky, D. W. & Cohen, A. A. Developments in molecular epidemiology of aging. *Emerg. Top. Life Sci.* **3**, 411–421 (2019).
- Lara, J. et al. A proposed panel of biomarkers of healthy ageing. *BMC Med.* **13**, 1–8 (2015).
- Cummings, S. R. & Kritchevsky, S. B. Endpoints for geroscience clinical trials: health outcomes, biomarkers, and biologic age. *Geroscience* **44**, 2925–2931 (2022).
- Macdonald-Dunlop, E. et al. A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk. *Aging* **14**, 623–659 (2022).
- Raj, K. & Horvath, S. Current perspectives on the cellular and molecular features of epigenetic ageing. *Exp. Biol. Med.* **245**, 1532–1542 (2020).
- Belsky, D. W. et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *eLife* **11**, e73420 (2022).
- Drewelies, J. et al. Using blood test parameters to define biological age among older adults: association with morbidity and mortality independent of chronological age validated in two separate birth cohorts. *Geroscience* **44**, 2685–2699 (2022).
- Seeman, T. E., Singer, B. H., Rowe, J. W., Horwitz, R. I. & McEwen, B. S. Price of adaptation—allostatic load and its health consequences: MacArthur studies of successful aging. *Arch. Intern. Med.* **157**, 2259–2268 (1997).
- Belsky, D. W. et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? *Am. J. Epidemiol.* **187**, 1220–1230 (2018).
- Vetter, V. M. et al. Epigenetic clock and relative telomere length represent largely different aspects of aging in the Berlin Aging Study II (BASE-II). *J. Gerontol. A Biol. Sci. Med. Sci.* **74**, 27–32 (2019).
- Vetter, V. M. et al. Relationship between 5 epigenetic clocks, telomere length, and functional capacity assessed in older adults: cross-sectional and longitudinal analyses. *J. Gerontol. A* **77**, 1724–1733 (2022).
- McCrary, C. et al. Epigenetic clocks and allostatic load reveal potential sex-specific drivers of biological aging. *J. Gerontol. A* **75**, 495–503 (2020).
- Kim, S., Myers, L., Wyckoff, J., Cherry, K. E. & Jazwinski, S. M. The frailty index outperforms DNA methylation age and its derivatives as an indicator of biological age. *Geroscience* **39**, 83–92 (2017).
- Zhang, Y., Saum, K.-U., Schöttker, B., Holleczek, B. & Brenner, H. Methyloomic survival predictors, frailty, and mortality. *Aging* **10**, 339 (2018).
- Li, X. et al. Longitudinal trajectories, correlations and mortality associations of nine biological ages across 20-years follow-up. *elife* **9**, e51507 (2020).
- Drewelies, J. et al. There are multiple clocks that time us: cross-sectional and longitudinal associations among 14 alternative indicators of age and aging. *J. Gerontol. A Biol. Sci. Med. Sci.* **80**, glae244 (2025).
- Kuiper, L. M. et al. Epigenetic and metabolomic biomarkers for biological age: a comparative analysis of mortality and frailty risk. *J. Gerontol. A* **78**, 1753–1762 (2023).
- Bertram, L. et al. Cohort profile: the Berlin Aging Study II (BASE-II). *Int. J. Epidemiol.* **43**, 703–712 (2014).
- Spieker, J. et al. Diabetes type 2 in the Berlin Aging Study II: cross-sectional and longitudinal data on prevalence, incidence and severity over on average seven years of follow-up. *Diab. Med.* **40**, e15104 (2023).
- Lemke, E. et al. Cardiovascular health is associated with the epigenetic clock in the Berlin Aging Study II (BASE-II). *Mech. Ageing Dev.* **201**, 111616 (2022).
- Demuth, I. et al. Cohort profile: follow-up of a Berlin Aging Study II (BASE-II) subsample as part of the GendAge study. *BMJ Open* **11**, e045576 (2021).
- von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* **370**, 1453–1457 (2007).
- Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
- Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
- Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573 (2018).
- Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303 (2019).
- Vetter, V. M., Kalies, C. H., Sommerer, Y., Bertram, L. & Demuth, I. Seven-CpG DNA methylation age determined by single nucleotide primer extension and Illumina's Infinium MethylationEPIC Array provide highly comparable results. *Front. Genet.* **12**, 759357 (2022).
- Meyer, A., Salewsky, B., Buchmann, N., Steinhagen-Thiessen, E. & Demuth, I. Relative leukocyte telomere length, hematological parameters and anemia—data from the Berlin Aging Study II (BASE-II). *Gerontology* **62**, 330–336 (2016).

37. Lu, A. T. et al. DNA methylation-based estimator of telomere length. *Aging* **11**, 5895 (2019).
38. Seeman, M., Merkin, S. S., Karlamangla, A., Koretz, B. & Seeman, T. Social status and biological dysregulation: the “status syndrome” and allostatic load. *Soc. Sci. Med.* **118**, 143–151 (2014).
39. McCrory, C. et al. How does socio-economic position (SEP) get biologically embedded? A comparison of allostatic load and the epigenetic clock(s). *Psychoneuroendocrinology* **104**, 64–73 (2019).
40. Rubin, D. C. & Berntsen, D. People over forty feel 20% younger than their age: subjective age across the lifespan. *Psychon. Bull. Rev.* **13**, 776–780 (2006).
41. Düzel, S. et al. The Subjective Health Horizon Questionnaire (SHH-Q): assessing future time perspectives for facets of an active lifestyle. *Gerontology* **62**, 345–353 (2016).
42. Jawinski, P. et al. Linking brain age gap to mental and physical health in the Berlin Aging Study II. *Front. Aging Neurosci.* **14**, 791222 (2022).
43. Vetter, V. M. et al. Low blood levels of selenium, selenoprotein P and GPx3 are associated with accelerated biological aging: results from the Berlin Aging Study II (BASE-II). *Clin Epigenetics.* **17**, 62 (2025).
44. Vetter, V. M. et al. Epigenetic aging and perceived psychological stress in old age. *Transl. Psychiatry* **12**, 410 (2022).
45. Vetter, V. M. et al. Vitamin D supplementation is associated with slower epigenetic aging. *Geroscience* **44**, 1847–1859 (2022).
46. Vetter, V. M. et al. DNA methylation age acceleration is associated with risk of diabetes complications. *Commun. Med.* **3**, 21 (2023).
47. Vetter, V. M., Spira, D., Banszerus, V. L. & Demuth, I. Epigenetic clock and leukocyte telomere length are associated with vitamin D status, but not with functional assessments and frailty in the Berlin Aging Study II. *J. Gerontol. A* **75**, 2056–2063 (2020).
48. Schmid, H. et al. Cross-sectional and Longitudinal Relationship Between Sex Hormones and 6 Epigenetic Clocks in Older Adults: Results of the Berlin Aging Study II. *J. Gerontol. A Biol. Sci. Med. Sci.* **80**, glaf106 (2025).
49. Demuth, I. et al. DunedinPACE predicts incident metabolic syndrome: cross-sectional and longitudinal data from the Berlin Aging Study II (BASE-II). *J. Gerontol. A Biol. Sci. Med. Sci.* **80**, glaf157 (2025).
50. Fried, L. P. et al. Frailty in older adults: evidence for a phenotype. *J. Gerontol. A Biol. Sci. Med. Sci.* **56**, M146–M157 (2001).
51. Vetter, V. M. et al. Change in body weight of older adults before and during the COVID-19 pandemic: longitudinal results from the Berlin Aging Study II. *J. Nutr. Health Aging* **28**, 100206 (2024).
52. Pajewski, N. M. et al. Characterizing Frailty Status in the Systolic Blood Pressure Intervention Trial. *J. Gerontol. A Biol. Sci. Med. Sci.* **71**, 649–655 (2016).
53. Tinetti, M. E., Williams, T. F. & Mayewski, R. Fall risk index for elderly patients based on number of chronic disabilities. *Am. J. Med.* **80**, 429–434 (1986).
54. Folstein, M. F., Folstein, S. E. & McHugh, P. R. “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* **12**, 189–198 (1975).
55. Wechsler, D. *Wechsler Adult Intelligence Scale—Revised* (Psychological Corporation, 1981).
56. Radloff, L. S. The CES-D scale: a self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1**, 385–401 (1977).
57. Mahoney, F. I. & Barthel, D. W. Functional evaluation: the Barthel Index. *Md. State Med. J.* **14**, 61–65 (1965).
58. Vellas, B. et al. The Mini Nutritional Assessment (MNA) and its use in grading the nutritional state of elderly patients. *Nutrition* **15**, 116–122 (1999).
59. American Diabetes Association. Classification and diagnosis of diabetes: standards of medical care in diabetes—2019. *Diab. Care* **42**, S13–S28 (2019).
60. Young, B. A. et al. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. *Am. J. Manag Care* **14**, 15 (2008).
61. Charlson, M. E., Pompei, P., Ales, K. L. & MacKenzie, C. R. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J. Chronic Dis.* **40**, 373–383 (1987).
62. SCORE2 Working Group and ESC Cardiovascular Risk Collaboration. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur. Heart J.* **42**, 2439–2454 (2021).
63. SCORE2-OP Working Group and ESC Cardiovascular Risk Collaboration. SCORE2-OP risk prediction algorithms: estimating incident cardiovascular event risk in older persons in four geographical risk regions. *Eur. Heart J.* **42**, 2455–2467 (2021).
64. König, M. et al. Historical trends in modifiable indicators of cardiovascular health and self-rated health among older adults: cohort differences over 20 years between the Berlin Aging Study (BASE) and the Berlin Aging Study II (BASE-II). *PLoS ONE* **13**, e0191699 (2018).
65. Lloyd-Jones, D. M. et al. Defining and setting national goals for cardiovascular health promotion and disease reduction: the American Heart Association’s strategic Impact Goal through 2020 and beyond. *Circulation* **121**, 586–613 (2010).
66. Alberti, K. G. et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* **120**, 1640–1645 (2009).
67. Hong, S. et al. TMEM106B and CPOX are genetic determinants of cerebrospinal fluid Alzheimer’s disease biomarker levels. *Alzheimers Dement.* **17**, 1628–1640 (2021).
68. R Core Team. *R: A Language and Environment for Statistical Computing*. <https://www.R-project.org> (R Foundation for Statistical Computing, 2022).
69. Moqri, M. et al. Validation of biomarkers of aging. *Nat. Med.* **30**, 360–372 (2024).
70. Hillary, R. F. et al. Epigenetic measures of ageing predict the prevalence and incidence of leading causes of death and disease burden. *Clin. Epigenet.* **12**, 115 (2020).
71. Kuo, P.-L. et al. Longitudinal changes in epigenetic clocks predict survival in the InCHIANTI cohort. Preprint at *medRxiv* <https://doi.org/10.1101/2024.09.13.24313620> (2024).
72. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**, 837–845 (1988).

Acknowledgements

This work was supported by grants of the Deutsche Forschungsgemeinschaft (project number 460683900 to I.D. and L.B.), the ERC (as part of the “Lifebrain” project to L.B. and C.M.L.), and the Cure Alzheimer’s Fund (as part of the “CIR-CUITS” consortium to L.B.), and a grant from the EU Joint Programme—Neurodegenerative Disease Research (JPND2021-650-289, coordinator: C.M.L.). This work was supported by the Ministry of Education and Research (BMBF), as part of the National Research Initiative ‘Mass Spectrometry in Systems Medicine’ (MSCoreSys), under grant agreement number 01EP2201 (to MR) and 16LW0239K (to MM). This article uses data from the Berlin Aging Study II (BASE-II). BASE-II was supported by the German Federal Ministry of Education and Research under grant numbers #01UW0808; #16SV5536K, #16SV5537, #16SV5538, #16SV5837, #01GL1716A, and #01GL1716B. C.M.L. was supported by the Heisenberg program of the DFG (DFG; LI 2654/4-1).

Author contributions

Conceptualization: V.M.V., I.D.; Data curation: V.M.V., J.H., J.D., S.D., L.D., P.J., E.K., S.M.; Formal analysis: V.M.V.; Funding acquisition: U.L., D.G., L.B., I.D., C.M.L.; Investigation: V.M.V., I.D.; Methodology: V.M.V., I.D.; Project administration: V.M.V., I.D.; Resources: S.K., M.M., M.R., U.L.,

C.M.L., D.G., L.B., I.D.; Supervision: I.D.; Visualization: V.M.V.; Writing—original draft: V.M.V., I.D.; and Writing—review & editing: all authors.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43856-025-01233-7>.

Correspondence and requests for materials should be addressed to Valentin Max Vetter or Ilja Demuth.

Peer review information *Communications Medicine* thanks Pei-Lun Kuo and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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

© The Author(s) 2025

Valentin Max Vetter¹ ✉, **Johanna Drewelies**^{2,3}, **Jan Homann**⁴, **Sandra Düzel**⁵, **Laura Deecke**^{4,6}, **Philippe Jawinski**³, **Simone Kühn**^{7,8}, **Elisa Kubala**¹, **Sebastian Markett**³, **Michael Müllereder**⁹, **Markus Ralser**^{10,11,12,13}, **Ulman Lindenberger**^{8,14}, **Christina M. Lill**^{4,15}, **Denis Gerstorff**³, **Lars Bertram**⁶ & **Ilja Demuth**^{1,16} ✉

¹Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Department of Endocrinology and Metabolic Diseases (including Division of Lipid Metabolism), Biology of Aging Working Group, Augustenburger Platz 1, Berlin, Germany. ²Center for Environmental Neuroscience, Max Planck Institute for Human Development, Berlin, Germany. ³Department of Psychology, Humboldt University Berlin, Berlin, Germany. ⁴Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany. ⁵Friede Springer Cardiovascular Prevention Center, Charité – Universitätsmedizin Berlin (CBF), Berlin, Germany. ⁶Lübeck Interdisciplinary Platform for Genome Analytics (LIGA), University of Lübeck, Lübeck, Germany. ⁷Center for Environmental Neuroscience, Max-Planck-Institut für Bildungsforschung, Berlin, Germany. ⁸Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, and, London, UK. ⁹Core Facility High Throughput Mass Spectrometry, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. ¹⁰The Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, Oxford, UK. ¹¹Berlin Institute of Health (BIH) at Charité – Universitätsmedizin Berlin, Berlin, Germany. ¹²Department of Biochemistry, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany. ¹³Max Planck Institute for Molecular Genetics, Berlin, Germany. ¹⁴Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany. ¹⁵Ageing Epidemiology Research Unit (AGE), School of Public Health, Imperial College London, London, UK. ¹⁶Charité - Universitätsmedizin Berlin, Berlin Institute of Health Center for Regenerative Therapies (BCRT), Berlin, Germany.

✉ e-mail: valentin.vetter@charite.de; ilja.demuth@charite.de