

DNA methylation age acceleration is associated with risk of diabetes complications

Valentin Max Vetter¹, Johanne Spieker¹, Yasmine Sommerer², Nikolaus Buchmann³, Christian Humberto Kalies¹, Vera Regitz-Zagrosek^{4,5}, Lars Bertram^{2,6} & Ilja Demuth^{1,7}✉

Abstract

Background Patients with Type 2 diabetes mellitus (T2D) are at risk for micro- and macrovascular complications. Implementable risk scores are needed to improve targeted prevention for patients that are particularly susceptible to complications. The epigenetic clock estimates an individual's biological age using DNA methylation profiles.

Methods In this study, we examined older adults of the Berlin Aging Study II that were reexamined on average 7.4 years after baseline assessment as part of the GendAge study. DNA methylation age (DNAmA) and its deviation from chronological age DNAmA acceleration (DNAmAA) were calculated with the 7-CpG clock (available at both timepoints, $n = 1,071$), Horvath's clock, Hannum's clock, PhenoAge and GrimAge (available at follow-up only, $n = 1,067$). T2D associated complications were assessed with the Diabetes Complications Severity Index (DCSI).

Results We report on a statistically significant association between oral glucose tolerance test results and Hannum and PhenoAge DNAmAA. PhenoAge was also associated with fasting glucose. In contrast, we found no cross-sectional association after covariate adjustment between DNAmAA and a diagnosis of T2D. However, longitudinal analyses showed that every additional year of 7-CpG DNAmAA at baseline increased the odds for developing one or more additional complications or worsening of an already existing complication during the follow-up period by 11% in male participants with T2D. This association persisted after covariate adjustment (OR = 1.11, $p = 0.045$, $n = 56$).

Conclusion Although our results remain to be independently validated, this study shows promising evidence of utility of the 7-CpG clock in identifying patients with diabetes who are at high risk for developing complications.

Plain language summary

Deterioration of vision, kidney function and cardiovascular function are just a few examples of diabetes-related complications. However, not all patients develop these complications, and it is desirable to detect patients that have a high risk for the complications early. In this study, we examine markers, which are based on reversible modifications of the DNA, in the context of diabetes and its complications. We found that one of these biomarkers is able to predict the development of diabetes complications over a period of about seven years in our dataset. If these results can be confirmed in other studies, our findings might help physicians to identify patients with diabetes that have an increased risk for developing complications in the future.

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, 13353 Berlin, Germany. ²Lübeck Interdisciplinary Platform for Genome Analytics (LIGA), University of Lübeck, Lübeck, Germany. ³Department of Cardiology, Charité - University Medicine Berlin Campus Benjamin Franklin, Berlin, Germany. ⁴Institute for Gender in Medicine, Center for Cardiovascular Research, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt - Universität zu Berlin and Berlin Institute of Health, Berlin, Germany. ⁵Department of Cardiology, University Hospital Zürich, University of Zürich, Zürich, Switzerland. ⁶Center for Lifespan Changes in Brain and Cognition (LCBC), Dept of Psychology, University of Oslo, Oslo, Norway. ⁷Charité - Universitätsmedizin Berlin, BCRT - Berlin Institute of Health Center for Regenerative Therapies, Berlin, Germany. ✉email: ilja.demuth@charite.de

The global number of people diagnosed with diabetes mellitus has increased fivefold to 537 million over the last 31 years^{1,2} and is projected to increase further to 784 million until 2045¹.

The disease's impact on well-being and daily living differs between individuals and is substantially driven by its complications³. Besides acute complications like diabetic ketoacidosis, hyperosmolar coma, and hypoglycemia, chronic microvascular (retinopathy, nephropathy and neuropathy) and macrovascular (coronary artery disease, peripheral artery disease and stroke) complications⁴ can result from diabetes mellitus. While some of these complications are life-threatening, others are at least disabling and can result in substantial emotional and financial burden for the affected individual and their family members^{2,5}. Furthermore, the resulting costs for the health care system are high¹ and expected to rise further⁶. A substantial part of these costs was attributed to diabetic complications^{2,7–10}.

Although complications cannot be prevented completely by the currently available therapeutic options, early interventions can delay their onset¹¹ and prevent potentially severe disease progression. Although several risk scores for diabetic complications exist (reviewed in refs. ^{12–18}), they are not used regularly in clinical practice¹⁹ nor are recommended by health organizations^{2,19,20}. Therefore, the demand for a risk score with high predictive power that is based on reliable and well-assessable data is still not met. To further improve diabetic complication prediction models, biomarkers were noted to be especially beneficial⁵.

In this study we evaluate an established biomarker of aging, DNA methylation age acceleration (DNAmAA)²¹, in a sample of 1100 participants of the Berlin Aging Study II (BASE-II) that were reassessed on average 7.4 years later as part of the GendAge study. This large longitudinal cohort has been investigated with respect to prevalence and incidence of T2D before²².

In a first step, the cross-sectional relationship between DNAmAA estimated from five epigenetic clocks (7-CpG clock, Horvath's clock, Hannum's clock, PhenoAge, GrimAge), prevalent T2D, and several T2D-associated blood parameters was examined. Previously published studies reported contradictory results. Dugue and colleagues²³, Irvin and colleagues²⁴, and Roetker and colleagues²⁵ found statistically significant associations between Horvath and Hannum clock estimates and T2D. However, McCartney and colleagues²⁶ and Horvath and colleagues²⁷ did not find any association between both variables.

In a second step, we analyzed these epigenetic clocks with respect to T2D-associated complications which were measured with the Diabetes Complications Severity Index (DCSI). To our knowledge, this is the first study that explores this relationship.

Cross-sectional associations were found for the Hannum and PhenoAge DNAmAA and longitudinal analyses suggest that the 7-CpG DNAmAA can help in the prediction of diabetic complications in men.

Methods

BASE-II and GendAge study. The BASE-II is an explorative multi-disciplinary study that examines factors promoting healthy vs. unhealthy aging. The medical part of this study included 1671 participants of the greater Berlin metropolitan area between the age of 60–85 years as assessed at baseline between 2009 and 2014. An additional group of 500 younger participants was assessed as well (age range 20 to 37 years) but is not analyzed in this study. After an average follow-up period of 7.4 years, 1083 participants of the older age group were reassessed as part of the GendAge study and are therefore available for longitudinal analyses. An additional 17 participants were only assessed at follow-up

examination. Participants who were assessed at baseline but not at follow-up were compared to participants who provided longitudinal data in a previously published manuscript. Participants without follow-up data were older and less educated, but they were no different from the follow-up participants with respect to gender and overall morbidity²².

All participants gave written informed consent. The Ethics Committee of the Charité – Universitätsmedizin Berlin approved the studies (approval numbers EA2/029/09 and EA2/144/16). The studies were conducted in accordance with the Declaration of Helsinki and were registered in the German Clinical Trials Registry as DRKS00009277 and DRKS00016157.

DNA methylation age. DNA methylation age (DNAmA) estimated from the 7-CpG clock was available for baseline and follow-up examination. This clock was trained on samples obtained from BASE-II participants during baseline examination²⁸ and is calculated from methylation data determined by methylation-sensitive single-nucleotide primer extension (MS-SNuPE)²⁹. Briefly, genomic DNA was isolated from whole blood samples and bisulfite converted. Subsequently, the areas of interest were amplified by a multiplex polymerase chain reaction. Finally, the SNuPE reaction was performed and the methylation fraction of the seven CpG sites of interest was measured by a 3730 DNA Analyzer (Applied Biosystems, HITACHI). The average change of 7-CpG DNAmA between baseline and follow up per year was 0.75 years (SD = 0.64 years, range: –4.8 to 5.4 years, $n = 965^{30}$), and correlation between DNAmA at T0 and T1 was high ($r = \text{Pearson's } r = 0.81$, $n = 965^{30}$). A detailed description of methods was previously published for baseline²⁸ and follow-up^{30,31}.

For follow-up only, four additional DNAmA measures were available that were estimated from epigenome wide DNAm profiles obtained through Illumina's Infinium MethylationEPIC array (Illumina Inc., USA). Horvath's clock³², Hannum's clock³³, PhenoAge³⁴ and GrimAge³⁵ were used to estimate DNAmA by the manual on Steve Horvath's website (<https://horvath.genetics.ucla.edu/html/dnamage/>). Briefly, outliers were identified by R's *outlyx* function (bigmelon³⁶ package), which determines them based on the interquartile range ($\text{iqr} = \text{TRUE}$) and the *pcout* function of R's *mvoutlier* package ($\text{mvP} = 0.15$). The *pcout* function uses the method described by Filzmoser, Maronna, and Werner³⁷ for fast identification of outliers in high-dimensional datasets. Outliers (identified by the *outlyx* and *pcout* function in the R-package bigmelon³⁶) and samples with a bisulfite conversion efficiency below 80% were removed. The resulting sample set was reloaded and normalized with the package's *dasen* function. Samples that had a root-mean squared deviation of 0.1 or more in beta-values before and after normalization were excluded. For epigenetic clock estimations, the raw (i.e., not normalized) DNAm values were uploaded to the website, according to the instructions in the manual. DNAmA of all five epigenetic clocks were moderately correlated with each other (Pearson's $r = 0.4$ to 0.6 , Supplementary Fig. 2A of ref. ³⁰). Detailed information on methods and longitudinal and cross-sectional descriptive statistics of the available DNAmA measures at follow-up can be found in ref. ³⁰.

DNA methylation age acceleration. To adjust for known age-associated changes in leukocyte cell composition, we calculated DNAmAA as residuals from a cell count (neutrophils, monocytes, lymphocytes, eosinophils) adjusted linear regression analysis of DNAmA on chronological age^{28,38} of all available participants. All cell counts were determined in a certified routine laboratory by flow cytometry (MVZ Labor 28 GmbH, Berlin, Germany). As

expected, correlation between DNAmAA and chronological age was low ($|\text{Pearson's } r| < 0.09^{30}$).

Diabetes mellitus type 2 (T2D) and Diabetes Complications Severity Index (DCSI). A detailed report on methodology and descriptive statistics of T2D in BASE-II and GendAge was reported before²². Diagnostic criteria of the American Diabetes Association (ADA) guidelines were used to identify participants with T2D³⁹.

The DCSI which was developed by Young and colleagues⁴⁰ incorporates information about seven categories of complications: Retinopathy, nephropathy, neuropathy, cerebrovascular disease, cardiovascular disease, peripheral vascular disease, and metabolic complications. One-on-one interviews between a study physician and the BASE-II participants were conducted and physician's letters and other medical documents were taken into account to assess previously diagnosed complications in the noted categories. A detailed list of the complications assessed in the BASE-II cohort and the corresponding items of the original manuscript of Young and colleagues is described in Supplementary Table 2 of ref. 22.

The resulting score is calculated based on the severity (as defined by Young and colleagues⁴⁰) of the respective complication (0=none, 1=some and 2=severe), except for neuropathy which can only result in a score of 0 or 1. Relevant information on the metabolic status were not available in this sample, therefore a maximum of 11 points was achievable. Dichotomized DCSI was calculated to distinguish the group of participants whose DCSI score did not increase (not increased) during follow-up from those whose DCSI score increased by at least one point (increased). The average change in DCSI per year during follow-up period was calculated as difference between DCSI at T1 and T0 divided by follow-up time. A detailed description on how the DCSI was assessed in this cohort is provided in ref. 22.

Covariates. Sex-differences in aging and in DNAmAA are well documented^{27,30,41,42}. Therefore, sex was included in all regression models and sex-stratified analyses were performed. Smoking behavior (in packyears) and alcohol consumption (baseline: g/d via food frequency questionnaire⁴³, follow-up: yes/no) were assessed in one-to-one interviews. Body mass index was calculated by height and weight measures obtained from the 763 seca measuring station (SECA, Germany). Diabetic medication was assessed by participants self-reports and/or from the medication list participants provided during both examinations.

Statistical analyses. Statistical analyses and all figures were conducted in the software package R 3.6.2⁴⁴. Logistic regression analyses were performed using the *glm* function (family=binomial) of the R's stats package. Linear regression analyses were computed with the *lm* function (stats package). ROC curves were calculated with the *proc* package's *roc* function⁴⁵. The same package was used to draw the Kaplan–Meier curves (*ggroc* function). Significance of difference between ROC curves was computed with the *roc.test*⁴⁵ function that employs the approach described by DeLong and colleagues⁴⁶.

The *ggplot2* package⁴⁷ was used to draw the figures included in this manuscript. Beeswarm plots were drawn with the *geom-beeswarm* function (*ggbeeswarm* package)⁴⁸. Boxplots were drawn with the *geom_boxplot* function. The median, hinges (corresponding to 25th and 75th percentile) and whiskers (1.5*inter-quartile range (IQR)) are shown. Outliers, defined as datapoints that lie beyond the whiskers, are shown as individual data points.

Statistical significance in differences between groups displayed in boxplots was assessed with the *stat_compare_means* function of the *ggpubr* package (computing a two-sided *t*-test)⁴⁹.

An available case analysis was performed. Therefore, participants who did not provide information on all variables needed for an analysis were excluded from it. The analyzed sample size is indicated for each analysis. Nominal statistical significance was defined at an alpha of 0.05.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Results

Cohort characteristics. As previously reported, 12.9% ($n = 209$ of a total sample of $n = 1,625$) of the participants were diagnosed with T2D at baseline²². On average, 7.4 years later this number increased to 17.1% ($n = 185$ of a total sample of $n = 1,083$) among the participants who completed the follow-up assessments²².

The mean age of BASE-II participants with diagnosed T2D at baseline and a completed follow-up examination was 68.0 years (SD = 3.7 years, $n = 126$, 41.3 % female, Table 1) at baseline and 75.5 years (SD = 4.1 years, Table 1) at follow-up. The DCSI of 54.8% ($n = 69$) of these participants increased during the on average 7.4-year follow-up period. The average increase in this group was 2.5 points for women (SD = 1.4, $n = 28$) and 1.9 points in the subgroup of men (SD = 1.0, $n = 41$). Male participants with T2D were found to have a significantly higher DCSI than women with T2D at baseline (difference = 0.71 points, $p = 0.005$, Supplementary Table 1). This difference was smaller and not statistically significant at follow-up (difference = 0.26, $p = 0.5$, Supplementary Table 1). Approximately the same number of men and women had a higher DCSI score at follow-up compared to baseline (women: 53.9%, men: 55.4%, Supplementary Table 1).

Association between five DNAmAA measures, diagnosed T2D and diagnostic blood parameters. At follow-up, we found participants with diagnosed T2D to have 0.93 years higher PhenoAge DNAmAA ($p = 0.042$, Fig. 1) and 0.85 year higher GrimAge DNAmAA ($p = 0.003$, Fig. 1) than participant that were not diagnosed with T2D. Sex-stratified analyses are shown in Supplementary Fig. 1. After adjustment for covariates in cross-sectional logistic regression analyses of the T2D diagnosis (yes/no) on DNAmAA of all five epigenetic clocks, no statistically significant association between these variables was found (*logistic regression*, Table 2, unadjusted models in Supplementary Table 2).

The only available epigenetic clock estimate at baseline examination, 7-CpG DNAmAA, was neither cross-sectionally associated with diagnosed T2D at baseline nor longitudinally associated with diagnosed T2D at follow-up. This was true for sex-stratified analyses as well (*logistic regression*, Supplementary Table 3 and Supplementary Table 4).

Fasting glucose was significantly associated with PhenoAge DNAmAA after adjustment for chronological age, sex, alcohol consumption (yes/no), smoking (packyears), antidiabetic medication (yes/no) and BMI (Table 3, $\beta = 0.26$, SE = 0.10, $p = 0.013$, *linear regression*, unadjusted models in Supplementary Table 5). Additionally, a statistically significant positive association was found between PhenoAge DNAmAA and results from the oral glucose tolerance test (oGTT, $\beta = 0.75$, SE = 0.26, $p = 0.003$, $n = 762$, *linear regression*, Table 3). Similarly, a statistically significant association of about the same effect size was found between Hannum's DNAmAA and oGTT in the highest adjusted model ($\beta = 0.78$, SE = 0.34, $p = 0.023$, $n = 762$, *linear regression*,

Table 1 Cohort characteristic of participants that were diagnosed with T2D at baseline and completed follow-up on average 7.4 years later ($n = 126$).

	<i>n</i>	mean, %	<i>sd</i>	min	max
Sex (female)	52	41.27			
Chronological Age (T0)	126	68.04	3.68	61.37	77.29
Chronological Age (T1)	126	75.46	4.09	66.78	85.94
7-CpG DNAmA (T0)	109	66.68	8.00	40.25	90.62
7-CpG DNAmAA (T0)	101	0.75	7.09	-20.75	22.33
Smoking (packyears, T0)	101	15.41	18.77	0	80
Alcohol intake (g/d, T0)	108	15.75	16.65	0.44	86.40
BMI (T0)	110	29.40	4.15	19.41	39.87
Antidiabetic medication (yes, T0)	66	52.80			
DCSI (T0)	126	1.13	1.40	0	7
DCSI (T1)	126	2.07	1.87	0	7
DCSI increase (yes)	69	54.76			
Fasting Glucose (T0)	121	123.45	31.61	59.00	254.00
oGTT (T0)*	42	193.64	61.45	94.00	308.00
HbA1c (T0)	119	6.48	0.73	5.00	8.90

DNAmA DNA methylation age, DNAmAA DNA methylation age acceleration, T0 baseline examination, T1 follow-up examination, g/d gram per day, DCSI Diabetes Complications Severity Index, oGTT oral glucose tolerance test, HbA1c glycated hemoglobin. *OGTT was only assessed in participants that were not already diagnosed with T2D before examination. Therefore, only comparatively few participants provide information in the subgroup shown here. Cohort characteristics of the whole dataset of participants at follow-up ($n = 1100$) are provided in Supplementary Table 8.

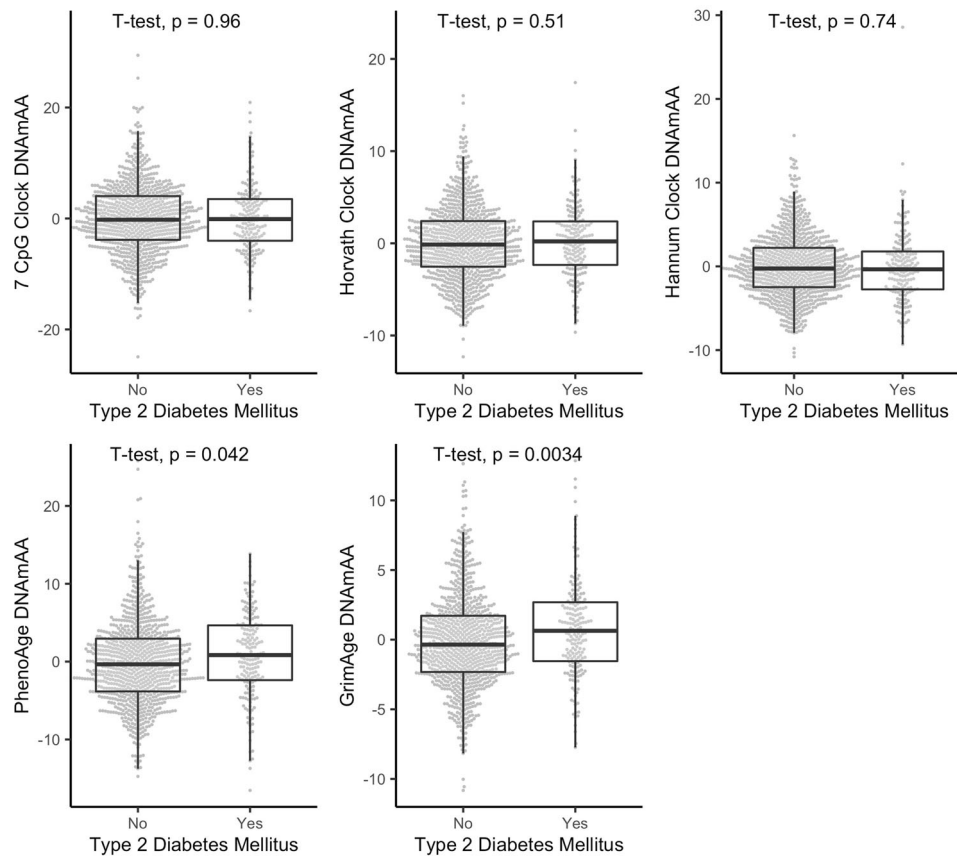


Fig. 1 Boxplots of DNAmAA estimates by five different epigenetic clocks, stratified by their diagnostic status (T2D vs. no T2D) at follow-up examination. Median, hinges (25th and 75th percentile) and Tukey-style whiskers (1.5* inter-quartile-range) are displayed. 7-CpG Clock DNAmAA: $n = 1055$; Horvath Clock DNAmAA, Hannum Clock DNAmAA, PhenoAge DNAmAA, GrimAge DNAmAA: $n = 1051$. Note: CpG=Cytosine phosphate guanine, DNAmAA=DNA methylation age acceleration.

Table 3). However, none of the examined DNAmAA measures were associated with glycosylated hemoglobin (HbA1c, $p > 0.2$, Table 3). In sex-stratified analyses, only the association between PhenoAge DNAmAA and oGTT in men remained statistically significant after covariate adjustment ($\beta = 0.96$, $SE = 0.39$, $p = 0.014$, $n = 353$, linear regression, Supplementary Table 6).

Cross-sectional association between DNAmAA of five different epigenetic clocks and DCSI. Cross-sectional analyses between 7-CpG DNAmAA and DCSI at baseline examination did not reveal statistically significant associations (Supplementary Table 7). Additional to the longitudinally available 7-CpG clock, four more epigenetic clocks were available at follow-up

Table 2 Logistic regression analyses of diagnosis of T2D on DNAmAA of five different epigenetic clocks in cross-sectional data at follow-up.

	Women and men						Women						Men							
	Estimate		SE		OR		p-value		n		Estimate		SE		OR		p-value		n	
	Estimate	SE	OR	p-value	n	Estimate	SE	OR	p-value	n	Estimate	SE	OR	p-value	n	Estimate	SE	OR	p-value	n
7-CpG DNAmAA	-0.03	0.02	0.97	0.151	977	-0.01	0.03	0.99	0.801	515	-0.05	0.03	0.95	0.112	462					
Horvath DNAmAA	0.00	0.03	1.00	0.900	974	-0.01	0.05	0.99	0.762	515	0.01	0.04	1.01	0.752	459					
Hannum DNAmAA	0.01	0.03	1.02	0.628	974	0.01	0.05	1.01	0.818	515	0.02	0.04	1.02	0.693	459					
PhenoAge DNAmAA	0.03	0.02	1.03	0.168	974	0.03	0.03	1.03	0.359	515	0.04	0.03	1.04	0.197	459					
GrimAge DNAmAA	0.06	0.04	1.06	0.148	974	0.09	0.06	1.09	0.156	515	0.02	0.06	1.03	0.659	459					

Analyses are adjusted for chronological age, sex (if applicable), alcohol consumption (yes/no), smoking (packyears), antidiabetic medication (yes/no), and BMI. All variables were assessed at follow-up. All available participants are included in these analyses.
 DNAmAA DNA methylation age acceleration, SE standard error, OR odds ratio.

Table 3 Linear regression analyses of continuous blood parameters on DNAmAA of five epigenetic clocks in cross-sectional data at follow-up.

	Fasting glucose						2h-oGTT						HbA1c											
	Estimate		SE		p-value		n		Estimate		SE		p-value		n		Estimate		SE		p-value		n	
	Estimate	SE	p-value	n	Estimate	SE	p-value	n	Estimate	SE	p-value	n	Estimate	SE	p-value	n	Estimate	SE	p-value	n	Estimate	SE	p-value	n
7-CpG DNAmAA	0.04	0.09	0.677	969	0.11	0.21	0.596	765	<0.01	<0.01	0.934	973												
Horvath DNAmAA	0.19	0.14	0.161	966	0.35	0.33	0.300	762	<0.01	<0.01	0.915	970												
Hannum DNAmAA	0.20	0.14	0.166	966	0.78	0.34	0.023*	762	<0.01	<0.01	0.314	970												
PhenoAge DNAmAA	0.26	0.10	0.013*	966	0.75	0.26	0.003**	762	<0.01	<0.01	0.201	970												
GrimAge DNAmAA	0.25	0.19	0.183	966	0.01	0.46	0.980	762	<0.01	<0.01	0.441	970												

Analyses are adjusted for chronological age, sex, alcohol consumption (yes/no), smoking (packyears), antidiabetic medication (yes/no), and BMI. All variables were assessed at follow-up. All available participants were included in this analysis.
 DNAmAA DNA methylation age acceleration, SE standard error, HbA1c glycated hemoglobin, oGTT oral glucose tolerance test. * $p < 0.05$, ** $p < 0.01$.

examination. Cohort characteristics of analyzed participants at follow-up are displayed in Supplementary Table 8. Cross-sectional linear regression analyses were performed to analyze the relationship between these biological age measures and DCSI in participants with diagnosed T2D at follow-up. No association between any of the available epigenetic clocks and DCSI was statistically significant after adjustment for covariates (*linear regression*, Table 4). Unadjusted and minimally adjusted models are displayed in Supplementary Table 9.

7-CpG DNAmAA predicts additional complications after 7.4-year follow-up time in subgroup of men. Participants who reported additional or worsened complications during the follow-up period had a 2.3 year higher 7-CpG DNAmAA at baseline ($p = 0.1$, Fig. 2). Sex-stratified subgroup analyses revealed a statistically significant difference of 4.5 years higher baseline 7-CpG DNAmAA in men whose DCSI increased after examination ($p < 0.01$, Fig. 2). However, in the female subgroup this association was inverse and not statistically significant (difference between means: 1.8 years, $p = 0.4$, Fig. 2).

To account for potential confounders, we performed a logistic regression analysis of the dichotomized change in DCSI on 7-CpG DNAmAA and covariates. In line with the findings reported above, we found a 11% increase of the odds for developing at least one additional or worsened complication captured by the DCSI during follow-up period for every additional year of DNAmAA in the subgroup of men. This association was independent from DCSI at baseline, chronological age, smoking, alcohol consumption, antidiabetic medication and BMI (OR = 1.11, $p = 0.045$, $n = 56$, *logistic regression*, Table 5, unadjusted models in Supplementary Table 10). No statistically significant association was found in the whole group (OR = 1.04, $p = 0.22$) and in the female subgroup (OR = 0.996, $p = 0.95$, Table 5). Similar results were found when analyzing the continuous change in DCSI per year during the follow-up period in a linear regression analysis (Supplementary Table 11). Receiver operating characteristic (ROC) curves of the highest adjusted model are displayed in Fig. 3. The area under the curve (AUC) of the ROC curve that resulted from the logistic regression model of all participants and the male and female subgroups was 0.6, 0.8 and 0.7. Interestingly, the inclusion of DNAmAA improved the predictive model in the subgroup of women of about the same degree as in the male subgroup (Fig. 3), although it did not contribute significantly to the logistic regression model. This might be the result of the small sample size available in this subgroup. The ROC-curves did not differ statistically ($p > 0.05$) which probably needs to be attributed to the small sample size as well.

Discussion

As a main result of the current study, we found that the 7-CpG clock derived variable DNAmAA was not associated with the diagnosis of T2D. However, baseline 7-CpG DNAmAA was associated with T2D severity on average 7.4 years later as operationalized by the DCSI in men. Although the four other analyzed epigenetic clocks were only available for cross-sectional analyses at follow-up, a similar predictive ability would be expected at least for PhenoAge DNAmAA. This clock derivative was associated with diagnosis of T2D (*t-test*), fasting glucose and results in the oGTT, a well-established test for impaired glucose tolerance (*linear regression*). These promising findings suggest that the 7-CpG clock (and possibly other clocks as well) might be able to substantially improve future diabetic complication risk scores.

The predictive ability of 7-CpG DNAmAA was analyzed by AUC of the logistic regression models of increase in DCSI (yes/

Table 4 Linear regression analyses of DCSI on DNAmAA and covariates in cross-sectional data at follow-up.

	Women and men			Women			Men		
	Estimate	SE	n	Estimate	SE	n	Estimate	SE	n
7-CpG DNAmAA	0.04	0.02	163	0.03	0.03	67	0.05	0.03	96
Horvath DNAmAA	-0.03	0.04	163	-0.04	0.06	67	-0.01	0.04	96
Hannum DNAmAA	0.00	0.03	163	0.07	0.06	67	-0.03	0.04	96
PhenoAge DNAmAA	0.03	0.03	163	0.00	0.04	67	0.02	0.03	96
GrimAge DNAmAA	-0.06	0.05	163	-0.14	0.09	67	-0.04	0.06	96

Included are participants with diagnosed T2D at follow-up. Analyses are adjusted for chronological age, sex (if applicable), alcohol consumption (yes/no), smoking (packyears), antidiabetic medication (yes/no), and BMI. All variables were assessed at follow-up. DNAmAA DNA methylation age acceleration, SE standard error.

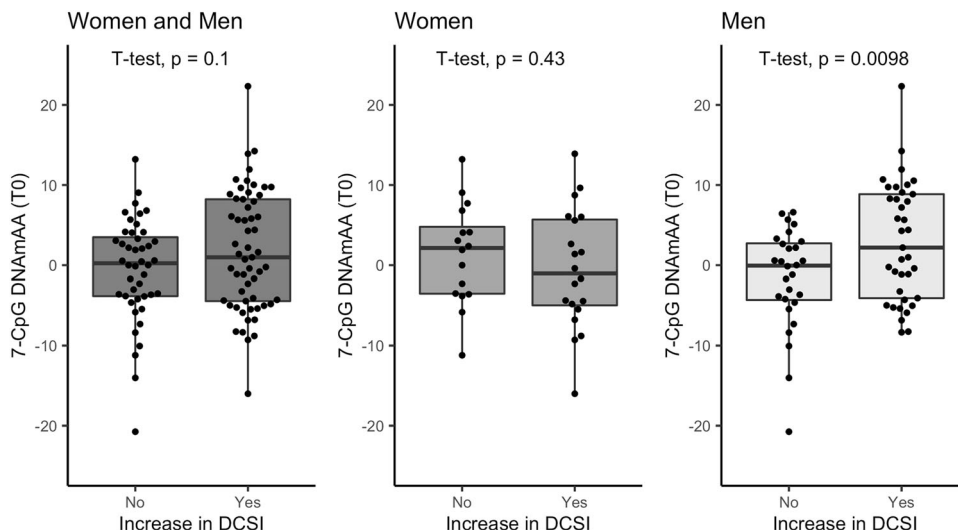


Fig. 2 Boxplot of 7-CpG DNAmAA at baseline stratified by increase in DCSI during follow-up period. Analyses were performed on all participants and in sex-stratified subgroups. Median, hinges (25th and 75th percentile) and Tukey-style whiskers (1.5* inter-quartile-range) are displayed. Women and Men: *n* = 101; Women: *n* = 36; Men: *n* = 65. Note: CpG=Cytosine phosphate guanine, DNAmAA=DNA methylation age acceleration, DCSI=Diabetes Complications Severity Index.

Table 5 Logistic regression of increase in DCSI of one or more points vs. no increase in DCSI (dichotomized) between baseline and follow-up on 7-CpG DNAmAA at baseline and covariates.

Model	Estimate	SE	OR	p-value	n
All	0.040	0.033	1.041	0.219	89
Women	-0.004	0.061	0.996	0.949	33
Men	0.101	0.051	1.106	0.045*	56

Analyses are adjusted for DCSI (T0), chronological age (T0), sex (if applicable), smoking (packyears, T0), alcohol consumption (g/d, T0), antidiabetic medication (yes/no, T0), and BMI (T0). Included are all participants with diagnosed T2D at baseline. T0 baseline, SE standard error, OR odds ratio. * *p* < 0.05.

no) during follow-up on 7-CpG DNAmAA at baseline and covariates which was in the acceptable range (AUC = 0.8, Fig. 3)⁵⁰. To our knowledge, there are no models available that aim at the prediction of the DCSI and could be used for a meaningful comparison of performance. However, numerous scores for specific micro- and macrovascular complications were reported before (reviewed in ref. 12–14,16,18). Saputro and colleagues reported on pooled C-statistics of logit-based models from cohort studies. Results for diabetic nephropathy (11 studies, 1 to 10 years of follow-up time, AUC = 0.78)¹⁸ and diabetic retinopathy (6 studies, 1 to 20 years of follow-up time, AUC = 0.82)¹⁸ were in a similar range to the model reported in this study. Similarly, Chowdhury and colleagues reported on a median AUC of 0.71 of nine models designed to predict stroke in patients with T2D (1.4 to 10.5 years of follow-up time)¹⁴ and Beulens and colleagues reported on an AUC between 0.54 and 0.81 of 20 models that aimed at the prediction of foot ulcer (1 to 5 years follow-up time)⁵¹. A detailed analysis of the previously developed risk models and the incorporated predictors is beyond the scope of this manuscript and was done before^{12–14,16,18} and the comparability between models is limited due to the high variability in the methods, cohorts and outcomes that were used to develop these risk scores. However, the results show that the 7-CpG DNAmAA informed regression model presented in this study performs in a comparable range (and sometimes better) than more complication specific risk models. This is especially intriguing as several

aspects of this still comparatively new biomarker seem very well suited for prediction modelling. First, the 7-CpG clock can be obtained through two different methodological approaches. It was developed for the MS-SNuPE method that can be conducted cost-effectively even in smaller laboratories²⁸, but additionally can be determined based on epigenome-wide array data, such as data obtained by microarrays³¹. Therefore, it can easily be applied to cohorts where epigenome-wide data is already available. Second, biomarkers allow an objective assessment of the individual complication risk. They are independent from factors that might interfere with data assessment in a clinical context, such as a language barrier, inaccurate or biased memory, unstandardized documentation, or examiner bias. An advantage of this specific biomarker is that changes in the epigenetic clock seem to result from lifelong and cumulative influences⁵². Therefore, it is expected to be robust against short-term changes and circadian differences that can complicate the use of biomarkers in a clinical context. In contrast to (poly-)genetic risk scores that are set at birth, the epigenetic clock changes throughout life. We and others have shown that these changes are potentially sensitive to interventions^{53–57} and lifestyle factors³⁸. Whether this is true for T2D specific interventions needs to be examined in sufficiently sized longitudinal studies. A meaningful relationship, however, seems plausible because high levels of glucose were shown to change DNA methylation by upregulating methylating enzymes and downregulating demethylating enzymes in in-vitro experiments of rat cells⁵⁸.

Despite these promising results in terms of complications related to T2D, no association was found between T2D diagnosis and DNAmAA of all five available clocks after adjustment for covariates (Table 2). In contrast to the first-generation clocks (7-CpG, Horvath’s and Hannum’s clock) that were trained to predict chronological age, PhenoAge and GrimAge were trained to predict phenotype-based biological age estimates. The negative findings with respect to the first-generation clocks reported here are in line with results that were reported by McCartney and colleagues²⁶ and Horvath and colleagues²⁷. In contrast, positive associations between Horvath’s or Hannum’s DNAmAA and history of diabetes were reported by Dugue and colleagues²³, Irvin and colleagues²⁴, and Roetker and colleagues²⁵. However, these cohorts differed in sex-distribution, age-range, or their

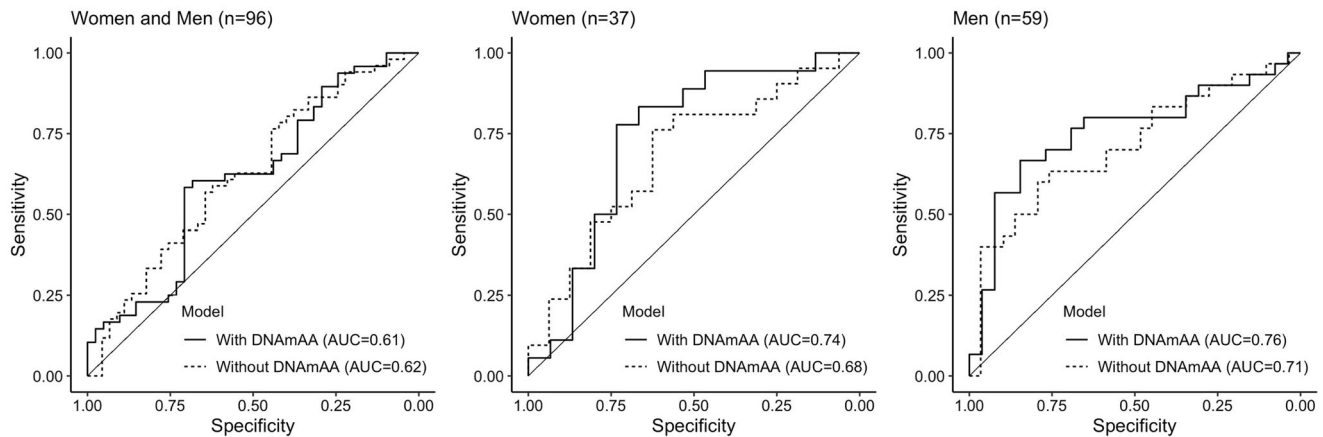


Fig. 3 Prediction of T2D complications as assessed with the DCSI. The basic model used for complication prediction includes chronological age (years), sex (if applicable), alcohol intake (g/d), packyears (years), antidiabetic medication (yes/no), and BMI. The differences between ROC curves of the model with and without 7-CpG DNAmAA were statistically not significant ($p > 0.05$). Note: DNAmAA=DNA methylation age acceleration, AUC=Area Under Curve.

statistical approach from the analyses presented here, which might at least partially explain the difference in findings. To our knowledge, GrimAge is the only second-generation clock that was examined with respect to T2D before. Kim and colleagues found an association between GrimAge DNAmAA and T2D (OR = 2.57, 95% CI: 1.61–4.11) in 318 obese participants (mean age = 40 years, 53% female). This association could not be found in the overweight and normal weight group of the study⁵⁹. An association between GrimAge DNAmAA and T2D was found in this study as well (OR = 1.07, $p = 0.002$, $n = 1051$, model 1, Supplementary Table 2), but in contrast to the findings by Kim and colleagues this association did not persist after inclusion of covariates.

To further evaluate the epigenetic clocks in the context of T2D, we performed linear regression analyses of T2D-associated blood parameters on DNAmAA. A statistically significant association between PhenoAge DNAmAA and fasting glucose was found (Table 3). This association seems plausible because serum glucose was included in the phenotypic age measure that was used to train the PhenoAge clock³⁴. The oGTT, a test used to assess how glucose is metabolized and that is used to diagnose T2D, was significantly associated with PhenoAge DNAmAA and Hannum's DNAmAA after covariate adjustment (Table 3). To our knowledge, this association has not been examined before.

There are several limitations to this study which we summarize as follows. First, the small sample size of participants with diagnosed T2D at baseline might be the reason for the lack of statistical significance in some analyses. A replication of these analyses in larger cohorts of patients with T2D is therefore needed. Second, we were not able to evaluate the predictive ability of the logistic regression model in an independent dataset. However, such independent validation and calibration analyses are crucial before translation into clinical practice is possible. Third, due to the exploratory approach of this study and in line with most studies of this field, we did not adjust our analyses for multiple testing. However, an increased rate of false-positive findings can therefore not be ruled out and our findings need to be replicated and validated in an independent cohort. Fourth, the outcome variable that represents diabetic complications, DCSI, reflects a general burden instead of a highly differentiated assessments of a specific T2D-associated complication. However, the general assessment of risk for complications is closer to clinical practice than a potentially better but complication-specific risk estimation.

This could help implementing a risk score in clinical practice in the future, where health care workers are challenged with the identification of patients with a generally high risk for complications rather than a high risk for specific complications. Nevertheless, further analyses examining the relationship between DNAmAA and individual T2D-associated complications are needed.

Strengths of this study include the wide variety of different epigenetic clock estimates assessed and the longitudinally available 7-CpG clock, which allowed a comprehensive evaluation of this still comparatively new biomarker of aging. Furthermore, the analyzed BASE-II cohort is well characterized with respect to T2D²² and provides a robust data base that allows the comparison of different T2D-associated variables and adjustment for relevant covariates.

Conclusion

None of the epigenetic clocks analyzed in this study were associated with diagnosis of T2D after adjustment for covariates. However, Hannum and PhenoAge DNAmAA showed a statistically significant association with oGTT. Furthermore, PhenoAge DNAmAA was associated with fasting glucose. Despite the lack of association between the 7-CpG clock with T2D and associated blood parameters, a one-year increase in 7-CpG DNAmAA was associated with 11% increase in the odds for development of additional T2D-related complications on average 7.4 years later in men. Although this association still lacks external validation, the results suggest that DNAmAA as a biomarker might be able to improve the identification of the group of patients with T2D that is especially prone to complications. Therefore, already existing prediction models might be substantially improved by the inclusion of 7-CpG DNAmAA which ultimately might accelerate their translation to clinical practice. However, independent, and further analyses are necessary before this biomarker can be translated into practice.

Data availability

Source data for Fig. 3 is provided as Supplementary Data. Due to concerns for participants privacy as well as data protection regulations, BASE-II raw data cannot be made publicly available. Because of their solely descriptive nature, this applies to source data for Figs. 1 and 2 as well. Interested investigators are invited to contact the scientific coordinator of BASE-II, Ludmila Müller (lmuller@mpib-berlin.mpg.de), to obtain source data for Figs. 1 and 2 or apply for raw data access. Additional information can be found on the BASE-II website: <https://www.base2.mpg.de/7549/data-documentation>.

Received: 12 July 2022; Accepted: 24 January 2023;

Published online: 10 February 2023

References

- Federation, I. D. IDF Diabetes Atlas, 10th edn. Brussels, Belgium: International Diabetes Federation. (2021).
- Organization, W.H. Global report on diabetes. 2016. (2017).
- Ndjabou, R. et al. Predictive models of diabetes complications: protocol for a scoping review. *Syst. Rev.* **9**, 1–14 (2020).
- Atlas, I. D. IDF diabetes atlas, 9th edn. Brussels and Belgium; 2019. (2021).
- Nickerson, H. D. & Dutta, S. Diabetic Complications: Current Challenges and Opportunities. *J. Cardiovas. Transl. Res.* **5**, 375–379 (2012).
- Bloom, D. E., et al. The global economic burden of noncommunicable diseases. (Program on the Global Demography of Aging, 2012).
- Seuring, T., Archangelidi, O. & Suhrcke, M. The economic costs of type 2 diabetes: a global systematic review. *Pharmacoeconomics*. **33**, 811–831 (2015).
- Church, V. Economic costs of diabetes in the US in 2002. *Diabetes Care*. **26**, 917–932 (2003).
- Kähm, K. et al. Health care costs associated with incident complications in patients with type 2 diabetes in Germany. *Diabetes Care*. **41**, 971–978 (2018).
- Zhuo, X., Zhang, P. & Hoerger, T. J. Lifetime direct medical costs of treating type 2 diabetes and diabetic complications. *American J. Prevent. Med.* **45**, 253–261 (2013).
- Organization, W.H. WHO package of essential noncommunicable (PEN) disease interventions for primary health care. (2020).
- Beulens, J. et al. Prediction models for the risk of diabetic foot in people with type 2 diabetes: a systematic review and external validation study. *Diabetologia* **62**, S459–S460 (2019).
- Chowdhury, M. Z., Yeasmin, F., Rabi, D. M., Ronksley, P. E. & Turin, T. C. Prognostic tools for cardiovascular disease in patients with type 2 diabetes: A systematic review and meta-analysis of C-statistics. *J. Diabetes Complications*. **33**, 98–111 (2019).
- Chowdhury, M. Z. I., Yeasmin, F., Rabi, D. M., Ronksley, P. E. & Turin, T. C. Predicting the risk of stroke among patients with type 2 diabetes: a systematic review and meta-analysis of C-statistics. *BMJ Open*. **9**, e025579 (2019).
- Cichosz, S. L., Johansen, M. D. & Hejlesen, O. Toward Big Data Analytics: Review of Predictive Models in Management of Diabetes and Its Complications. *J. Diabetes Sci. Technol.* **10**, 27–34 (2016).
- Haider, S., Sadiq, S. N., Moore, D., Price, M. J. & Nirantharakumar, K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. *Eye* **33**, 702–713 (2019).
- Lagani, V., Koumakis, L., Chiarugi, F., Lakasing, E. & Tsamardinos, I. A systematic review of predictive risk models for diabetes complications based on large scale clinical studies. *J. Diabetes Complications*. **27**, 407–413 (2013).
- Saputro, S. A., Pattanaprteep, O., Pattanateepaporn, A., Karmacharya, S. & Thakkinstian, A. Prognostic models of diabetic microvascular complications: a systematic review and meta-analysis. *Systematic Rev.* **10**, 288 (2021).
- Nijpels, G., Beulens, J. W., van der Heijden, A. A. & Elders, P. J. Innovations in personalised diabetes care and risk management. *Eur. J. Prev. Cardiol.* **26**, 125–132 (2019).
- Federation, I. D. Recommendations For Managing Type 2 Diabetes In Primary Care. www.idf.org/managing-type2-diabetes. (2017).
- Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* **19**, 371–384 (2018).
- Spieker, J., et al. Diabetes Type 2 in the Berlin Aging Study II: Prevalence, Incidence and Severity Over up to Ten Years of Follow-up. *PREPRINT (Version 3) available at Research Square* (2021).
- Dugué, P.-A. et al. Association of DNA methylation-based biological age with health risk factors and overall and cause-specific mortality. *American J. Epidemiol.* **187**, 529–538 (2018).
- Irvin, M. R. et al. Metabolic and inflammatory biomarkers are associated with epigenetic aging acceleration estimates in the GOLDN study. *Clinical Epigenetics*. **10**, 1–9 (2018).
- Roetker, N. S., Pankow, J. S., Bressler, J., Morrison, A. C. & Boerwinkle, E. Prospective study of epigenetic age acceleration and incidence of cardiovascular disease outcomes in the ARIC study (Atherosclerosis Risk in Communities). *Circulation: Genomic Precision Medicine*. **11**, e001937 (2018).
- McCartney, D. L. et al. Investigating the relationship between DNA methylation age acceleration and risk factors for Alzheimer’s disease. *Alzheimer’s Dement.* **10**, 429–437 (2018).
- Horvath, S. et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol.* **17**, 1–23 (2016).
- 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).
- Kaminsky, Z. A., Assadzadeh, A., Flanagan, J. & Petronis, A. Single nucleotide extension technology for quantitative site-specific evaluation of metC/C in GC-rich regions. *Nucl. Acids Res.* **33**, e95 (2005).
- Vetter, V. M., et al. Relationship between five Epigenetic Clocks, Telomere Length and Functional Capacity assessed in Older Adults: Cross-sectional and Longitudinal Analyses. *J Gerontol A Biol Sci Med Sci.* **77**, 1724–1733 (2022).
- 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).
- Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biology.* **14**, R115 (2013).
- Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular 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).
- Gorrie-Stone, T. J. et al. Bigmelon: tools for analysing large DNA methylation datasets. *Bioinformatics* **35**, 981–986 (2019).
- Filmzmoser, P., Maronna, R. & Werner, M. Outlier identification in high dimensions. *Comput. Stat. Data Analysis.* **52**, 1694–1711 (2008).
- Quach, A. et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging* **9**, 419–446 (2017).
- Association, A. D. Classification and diagnosis of diabetes: standards of medical care in diabetes—2019. *Diabetes Care.* **42**, S13–S28 (2019). 2.
- Young, B. A. et al. Diabetes complications severity index and risk of mortality, hospitalization, and healthcare utilization. *American J. Managed Care.* **14**, 15 (2008).
- Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E. & Boks, M. P. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing Res. Rev.* **69**, 101348 (2021).
- 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 Biol Sci Med Sci.* **75**, 2056–2063 (2020).
- Nöthlings, U., Hoffmann, K., Bergmann, M. M. & Boeing, H. Fitting portion sizes in a self-administered food frequency questionnaire. *J. Nutr.* **137**, 2781–2786 (2007).
- Team, R. C. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* URL <https://www.R-project.org> (2017).
- Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* **12**, 1–8 (2011).
- 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).
- Wickham, H. *Elegant graphics for data analysis (ggplot2)*. (New York, NY: Springer-Verlag, 2009).
- Clarke, E. & Sherrill-Mix, S. Ggbeeswarm: Categorical scatter (violin point) plots. *R package version 0.6. 0* (2017).
- Kassambara, A. *ggpubr: ggplot2-Based Publication Ready Plots*. R package version 0.4. 0. *Computer software*. <https://cran-r-project.org/web/packages/ggpubr/index.html> (2020).
- Mandrekar, J. N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thoracic Oncol.* **5**, 1315–1316 (2010).
- Beulens, J. W. J. et al. Prognostic models for predicting the risk of foot ulcer or amputation in people with type 2 diabetes: a systematic review and external validation study. *Diabetologia* **64**, 1550–1562 (2021).
- Zannas, A. & Chrousos, G. Epigenetic programming by stress and glucocorticoids along the human lifespan. *Molecular Psychiatry.* **22**, 640–646 (2017).
- Fahy, G. M. et al. Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell.* **18**, e13028 (2019).
- Fiorito, G. et al. DNA methylation-based biomarkers of aging were slowed down in a two-year diet and physical activity intervention trial: the DAMA study. *Aging Cell* **20**, e13439 (2021).
- Fitzgerald, K. N. et al. Potential reversal of epigenetic age using a diet and lifestyle intervention: a pilot randomized clinical trial. *Aging* **13**, 9419–9432 (2021).
- Chen, L. et al. Effects of Vitamin D3 Supplementation on Epigenetic Aging in Overweight and Obese African Americans With Suboptimal Vitamin D Status: A Randomized Clinical Trial. *J. Gerontol. Series A, Biological Sciences and Medical Sciences* **74**, 91–98 (2019).
- Vetter, V. M. et al. Vitamin D supplementation is associated with slower epigenetic aging. *GeroScience* **44**, 1847–1859 (2022).
- Ishikawa, K. et al. Long-term pancreatic beta cell exposure to high levels of glucose but not palmitate induces DNA methylation within the insulin gene promoter and represses transcriptional activity. *PLoS one.* **10**, e0115350 (2015).

59. Kim, K. et al. DNA methylation *grimAge* and incident diabetes: The coronary artery risk development in young adults (CARDIA) study. *Diabetes* **70**, 1404–1413 (2021).

Acknowledgements

This work was supported by grants of the Deutsche Forschungsgemeinschaft (grant number DE 842/7-1 to I.D.), the ERC (as part of the Lifebrain project to L.B.), and the Cure Alzheimer's Fund (as part of the CIRCUITS consortium to L.B.). This article uses data from the Berlin Aging Study II (BASE-II) and the GendAge study which were supported by the German Federal Ministry of Education and Research under grant numbers #01UW0808; #16SV5536K, #16SV5537, #16SV5538, #16SV5837, #01GL1716A and #01GL1716B. We thank all probands of the BASE-II/GendAge study for their participation in this research.

Author contributions

Conceived and designed the study: V.M.V. and I.D. Contributed study specific data: V.M.V., J.S., Y.S., N.B., C.H.K., V.R.Z., L.B., and I.D. Analyzed the data: V.M.V. and Y.S. Wrote the manuscript: V.M.V. and I.D. All authors revised and approved the manuscript.

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/s43856-023-00250-8>.

Correspondence and requests for materials should be addressed to Ilja Demuth.

Peer review information *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023