GeroScience (2024) 46:2583–2604
https://doi.org/10.1007/s11357-023-01029-4

ORIGINAL ARTICLE

Analysis of epigenetic clocks links yoga, sleep, education, reduced meat intake, coffee, and a SOCS2 gene variant to slower epigenetic aging

Rezvan Noroozi · Joanna Rudnicka · Aleksandra Pisarek · Bożena Wysocka · Aleksander Masny · Michał Boroń · Kamila Migacz-Gruszka · Paulina Pruszkowska-Przybylska · Magdalena Kobus · Dagmara Lisman · Grażyna Zielińska · Aleksandra Iljin · Joanna A. Wiktorska · Małgorzata Michalczyk · Piotr Kaczka · Michał Krzysztofik · Aneta Sitek · Andrzej Ossowski · Magdalena Spólnicka · Wojciech Branicki · Ewelina Pośpiech

Received: 13 September 2023 / Accepted: 23 November 2023 / Published online: 16 December 2023

Abstract DNA methylation (DNAm) clocks hold promise for measuring biological age, useful for guiding clinical interventions and forensic identification. This study compared the commonly used DNAm clocks, using DNA methylation and SNP data generated from nearly 1000 human blood or buccal swab samples. We evaluated different preprocessing methods for age estimation, investigated the association of epigenetic age acceleration (EAA) with various lifestyle and sociodemographic factors, and undertook a series of novel genome-wide association analyses for different EAA measures to find

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11357-023-01029-4.

R. Noroozi · J. Rudnicka
Doctoral School of Exact and Natural Sciences, Jagiellonian University, Krakow, Poland

R. Noroozi
Johns Hopkins University School of Medicine, Baltimore, MD, USA

J. Rudnicka · E. Pośpiech
Malopolska Centre of Biotechnology, Jagiellonian University, Krakow, Poland

A. Pisarek · W. Branicki
Institute of Zoology and Biomedical Research of the Jagiellonian University, Krakow, Poland

B. Wysocka · A. Masny · M. Boroń
Central Forensic Laboratory of the Police, Warsaw, Poland

K. Migacz-Gruszka
Jagiellonian University Medical University, Krakow, Poland

P. Pruszkowska-Przybylska · A. Sitek
Department of Anthropology, University of Lodz, Lodz, Poland

M. Kobus
Institute of Biological Sciences, Faculty of Biology and Environmental Sciences, Cardinal Stefan Wyszynski University in Warsaw, Warsaw, Poland

D. Lisman · G. Zielińska · A. Ossowski · E. Pośpiech (✉)
Department of Forensic Genetics, Pomeranian Medical University in Szczecin, Szczecin, Poland
e-mail: ewelina.pospiech@pum.edu.pl

A. Iljin
Department of Plastic, Reconstructive and Aesthetic Surgery, Medical University of Lodz, Lodz, Poland

J. A. Wiktorska
MelissaMed M. Gauer, General Partnership, Lodz, Poland

M. Michalczyk · P. Kaczka · M. Krzysztofik
Department of Sport Nutrition, The Jerzy Kukuczka Academy of Physical Education in Katowice, Katowice, Poland

M. Spólnicka
Center for Forensic Science University of Warsaw, Warsaw, Poland

W. Branicki
Institute of Forensic Research, Krakow, Poland

associated genetic variants. Our results highlighted the Skin&Blood clock with ssNoob normalization as the most accurate predictor of chronological age. We provided novel evidence for an association between the practice of yoga and a reduction in the pace of aging (DunedinPACE). Increased sleep and physical activity were associated with lower mortality risk score (MRS) in our dataset. University degree, vegetable consumption, and coffee intake were associated with reduced levels of epigenetic aging, whereas smoking, higher BMI, meat consumption, and manual occupation correlated well with faster epigenetic aging, with FitAge, GrimAge, and DunedinPACE clocks showing the most robust associations. In addition, we found a novel association signal for *SOCS2* rs73218878 ($p = 2.87 \times 10^{-8}$) and accelerated GrimAge. Our study emphasizes the importance of an optimized DNAm analysis workflow for accurate estimation of epigenetic age, which may influence downstream analyses. The results support the influence of genetic background on EAA. The associated *SOCS2* is a member of the suppressor of cytokine signaling family known for its role in human longevity. The reported association between various risk factors and EAA has practical implications for the development of health programs to improve quality of life and reduce premature mortality associated with age-related diseases.

Keywords DNA methylation age · Epigenetic age acceleration · Epigenetic clock · Yoga · Coffee · SOCS2

Introduction

DNA methylation (DNAm) is a well-studied epigenetic modification with an age-related pattern of changes that can serve as a surrogate measure of biological aging in various tissues. Therefore, several DNAm age estimators have been developed to quantify epigenetic aging as a powerful tool for monitoring the effectiveness of geroprotective interventions [1, 2] or to precisely predict calendar age in forensic research [3]. The commonly used DNAm-based age estimators, known as DNAm clocks, consider various aspects of the aging process and vary in the precision of predicting age or the strength of association with various diseases. The first generation of DNAm clocks was trained on chronological age on mixed-age samples, ranging from children to older adults, using CpG sets with strongly time-dependent alterations in methylation patterns. These DNAm clocks such as Hannum [4], Horvath2013 [5], and Skin&Blood clock [6] estimate DNAm age in units of years, and the difference between their calculated DNAm age and chronological age can show that an individual is biologically younger or older than expected. Corresponding measures of age-adjusted epigenetic age acceleration (EAA) have been shown to be associated with different age-related health conditions such as obesity [7], all-cause mortality [8], physical and cognitive fitness [9], as well as various lifestyle-related risk factors [10], and diseases such as Down syndrome [11] or Alzheimer's disease [12].

Subsequently, the second generation of DNAm-based prediction models shifted to predicting health span and life span rather than chronological age. These models, which aim to capture time to death, include the PhenoAge clock [13], which is trained on information obtained from clinical biomarkers of physiological status and additional information from chronological age, and the GrimAge clock [14], a mortality risk estimator trained on DNAm-based surrogate measures of plasma proteins, and pack-years of smoking, plus sex and age. The corresponding age-adjusted GrimAge Acceleration (GrimAgeAccel) was reported to outperform the EAA measures obtained from first-generation clocks and PhenoAgeAccel in predicting time to cancer and time to coronary heart disease.

Another type of DNAm model is a speedometer of the pace of aging, DunedinPoAm [15] and its newer version DunedinPACE [16]. These are trained on DNAm surrogate measures of clinical biomarkers, similar to second-generation clocks, while instead of being trained on mixed-age samples, they are longitudinal estimates of biological aging in same-age individuals. The output of PoAm and PACE is reported as a measure of age-related physiological decline per year and is strongly associated with physical fitness, cognitive ability, and facial aging. Also, recently a novel DNAm base indicator of biological aging, named DNAmFitAge, has been developed, based on blood-based DNAm surrogate measures for fitness parameters plus DNAm GrimAge. Physically fitter individuals showed younger DNAmFitAge associated with decreased risk of mortality and coronary heart

disease [17]. Yet a new concept of DNAm age clock is emerging, aiming to capture changes observed in various biological processes and functions, based on a complex systems theory of aging, which suggests that the hallmarks of aging are caused by disruptions in the integration of regulatory mechanisms [18, 19].

Applying different DNAm age clocks simultaneously to a given dataset shows a weak correlation between different EAAs [20] and their disagreement in association with health conditions. This may indicate that, due to a different set of information and samples used to train the models, the biological aging measures obtained by individual calculators may differ and relate to different aspects of the aging process [16]. Therefore, a mechanistic understanding of the underlying biology of the measure captured by each clock is important to select the appropriate model for a given biological study [21]. On the other hand, while the difference between DNAm age and calendar age appears to be biologically meaningful in medical research, other criteria define the appropriate DNAm clock for forensic purposes. The first generation of DNAm clocks show practical values in suspect identification by accurately predicting chronological age [5, 6, 22–25]. In addition, biological age and individual pace of aging, derived from the second and third generation of DNAm estimators, respectively, can be informative for forensic purposes in predicting age-related physical appearance traits [26–28].

In the current study, we applied different DNAm age prediction models to the novel set of methylation array data to compare the outputs of different epigenetic clocks available in the literature. To increase the reproducibility of the methylation data, the selection of the best preprocessing method to control different known or unknown noises is critical, since the reliability of the outputs of DNAm clocks is sensitive to technical variations [29]. Therefore, we assessed the effect of different commonly used preprocessing methods on the accuracy of the age predicted by each clock to suggest the most effective and precise model for chronological age prediction. In addition, we presented statistical characteristics of epigenetic aging in the Polish population and assessed the association of different epigenetic age measures with sociodemographic data and lifestyle-related risk factors to further evaluate the relevance of each clock in aging research. Finally, we performed a series of novel genome-wide association studies on various EAA parameters to assess the effect of genetics vs. environment on the rate of epigenetic aging.

Material and methods

Study design

The number of 741 blood and 221 buccal swab samples, obtained from healthy Polish individuals (age range 20–81 years; mean ± SD: 46.5 ± 14.7, 49.7 ± 17.8 for blood and buccal swabs, respectively), were analyzed. Samples were collected from volunteers as part of the Polish epigenome project to collect methylation and SNP data representative of the Polish population. For the purposes of the analyses performed in this study, relatives and individuals under the age of 20 were excluded. Written informed consent forms were obtained from all participants and the study was approved by the Bioethics Committee of the Jagiellonian University in Krakow (decision no. 1072.6120.132.2018). DNA was extracted using an automated method and the Maxwell RSC Blood DNA Kit (Promega Corporation), and next assessed for purity using NanoDrop (Thermo Scientific, MA, USA), and for concentration using the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific). SNP data were collected using the Global Screening Array (GSA, Illumina, San Diego, CA) with 200 ng as the input DNA material. For the purpose of DNA methylation profiling, the amount of 500 ng of DNA per sample was subjected to bisulfite conversion using the EZ-96 DNA Methylation kit (Zymo Research Corp., CA, USA) following the manufacturer's instructions for Infinium assays. The DNA methylation profile, for samples randomized using the web-based application RANDOMIZE [30], was obtained using the Illumina Infinium Methylation EPIC microarray (Illumina, San Diego, CA, USA) [31].

Participants

For each participant, demographic characteristics, lifestyle data, and dietary information were collected through a comprehensive self-report questionnaire. Education, occupation, socioeconomic status, and place of residence were reported by the participants. Education level was divided into two categories and those with no university degree (primary school, high

school, or vocational school) were compared to participants with a university degree. In addition, self-reported socioeconomic status (SES) was collected. We asked participants to define their socioeconomic status as very low, low, average, and above average, and we suggested that they consider several key factors that may collectively affect their quality of life, including their income, education, occupation, professional status, and wealth. For the purpose of statistical analyses, the categorization of socioeconomic status was simplified from four to three categories by combining the categories low and very low into "low," average into "medium," and above average into "high" SES. Place of residence was categorized as village, city with less than 100,000 inhabitants (City < 100 K), and city with more than 100,000 inhabitants (City > 100 K). Job type was categorized as physical/manual, mental work partially sedentary (up to 4 h per day; sedentary mental; SM), mental work only sedentary (more than 4 h per day; long hours sedentary mental; LSM), or retired/unemployed. For statistical comparisons, the unemployed/retired group was excluded from the analyses. Harmful factors in the workplace were also recorded, including low/high temperature in the workplace, exposure to pesticides/chemicals, toxins/heavy metals/air pollution, ionizing radiation, sun, and stress.

The physical activity levels of the participants was also recorded. Self-reported physical activity, described as any type of physical activity (e.g., exercising, jogging, cycling), was collected, and those who reported being active at least once a week or every day were classified as active, and those who reported being active once a month or not at all were classified as inactive. Data were also collected on the specific types of physical activity engaged in and categorized into three types of physical activity: aerobic (e.g., soccer, volleyball, basketball, running, jogging, dancing, karate, swimming, fishing, walking, tennis, and cardio), strengthening (e.g., bodybuilding, body shaping, cross fit, weight lifting, strength training, and gym), and yoga/balance. Specific types of physical activity was compared with inactivity. Participants' smoking status was categorized into three discrete groups: current smokers, former smokers who had quit smoking for at least one year or more, and never smokers. Frequency of alcohol consumption was classified into three categories: non-drinkers, occasional drinkers (drinking once a week), and frequent drinkers (drinking at least three times a week). Information was also collected on the number of meals per day, portion of fruit and vegetables, frequency of fish and meat consumption, and cups of coffee per day. Sleep hours were recorded and individuals were divided into two groups with less than or equal to 8 h of sleep and more than 8 h of sleep. Detailed sociodemographic and lifestyle-related characteristics are reported in Table S1.

DNA methylation analysis, preprocessing, and quality control

Primary quality control of the generated DNAm data was done using GenomeStudio software (Methylation Module v1.8, Illumina Inc, 2008). The Illumina internal controls and background subtraction were applied to the samples. The internal control metrics were generated, considering Illumina-recommended cut-offs, and samples were filtered based on the Illumina guide. Further quality control and preprocessing steps were done using R version 4.2.2.

After loading raw data files into R, low-quality samples and probes were filtered based on the threshold of 0.05 detection *P*-values. Then minfi package was used to apply different preprocessing steps for background correction and probe-type and dye-bias adjustment. Five different preprocessing pipelines were used including the Illumina pipeline, implemented in the GenomeStudio Methylation Module and the minfi package, subset-quantile within array normalization (SWAN), a single sample of normal-exponential out-of-band (ssNoob) background correction method, Quantile and Functional normalization (FunNorm) [32–37].

DNA methylation age estimation

Different DNAm-based biomarkers were calculated using DNAm predictors including the DNAm Age (Horvath2013) [5], DNAm Age Hannum [4], DNAm Age Skin&Blood [6], DNAm PhenoAge [13], DNAm GrimAge [14], and DNAm FitAge [17]. Respective epigenetic age acceleration (EAA) measures were calculated, as the residuals of the DNAm age regressed on chronological age, respectively as the intrinsic epigenetic age acceleration (IEAA), extrinsic epigenetic age acceleration (EEAA), DNAm Skin&Blood age acceleration (Skin&BloodAgeAccel), DNAm

PhenoAge acceleration (PhenoAgeAccel), DNAm GrimAge acceleration (GrimAgeAccel), and DNAm FitAge acceleration (FitAgeAccel). Also, DNAm-based pace of aging (DunedinPACE and PoAm) and Mortality risk score (MRS) were measured [15, 16, 38]. Descriptions of the analyzed epigenetic estimators are shown in Table 1.

DNAm ages for each model were obtained using the Horvath online calculator, or the R packages including methylCIPHER and Methylclock [39, 40]. The accuracy of each clock for chronological age prediction was evaluated by calculating the mean absolute error (MAE). The MAEs obtained from the output of methylclock and methylCIPHER R packages were compared to examine the concordance among available methods regarding estimating DNAm age. In addition, we tested the reliability of the results of various epigenetic clocks in a set of 47 blood samples for which we had access to technical replicates per sample (age range 23–78). The estimated age was expected to be equal for each pair of technical replicates and the deviation between two measures in each pair was calculated for Horvath2013, Hannum, Skin&Blood, PhenoAge, and GrimAge clocks as well as the principal component-based version of each clock (PC-clocks). The R codes available for PC clocks were used for calculating the PC version of each model [41]. All the analysis and data visualization were done using R packages [42–45] in R version 4.2.2 [46].

Statistical analysis

The association of sociodemographic characteristics including education, type of job, socioeconomic status, area of residence and lifestyle risk factors including stress status, yoga practice, sleeping hours, physical activity, BMI, smoking status, frequency of alcohol drinking, diet type, coffee consumption, vegetable, fruit, and meat consumption with different EAA measures including IEAA, EEAA, PhenoAgeAccel, GrimAgeAccel, and FitAgeAccel as well as DunedinPACE, DunedinPoAm, and MRS was assessed. For each set of association analyses, two linear regression models were fitted. Firstly, each DNAm biomarker was analyzed as the outcome and each sociodemographic or lifestyle factor at a time was introduced as the predictor, adjusting for age and sex (Model 1). Then, fully adjusted models were fitted for each DNAm age acceleration measure as the outcome and a single sociodemographic factor of interest as the predictor adjusting for age, sex, selected lifestyle-related risk factors, including smoking status, physical activity, and BMI, as well as blood cell compositions [47] (Model 2, full model). In addition, the results of yoga's association with epigenetic aging were adjusted for physical activity, frequency of meat and vegetable consumption, hours of sleep, and cups of coffee per day. For EAA measures the coefficients were interpreted as the degree of epigenetic age acceleration associated with the sociodemographic factor. For DunedinPACE, the coefficients were interpreted as a change in the rate of aging per year associated with each sociodemographic factor. And for the MRS the coefficients were interpreted as the score of mortality risk associated with each sociodemographic factor.

Table 1 Description of the different epigenetic age estimators

Clock	No. CpGs	Measured parameter	Available method for calculation
Horvath 2013	353	Chronological age	HOC, methylCIPHER, methylClock
Horvath Skin&Blood	391	Chronological age	HOC, methylCIPHER, methylClock
Hannum	71	Chronological age	HOC, methylCIPHER, methylClock
PhenoAge	513	Biological age, health span	HOC, methylCIPHER, methylClock
GrimAge	1030	Biological age, mortality risk	HOC
FitAge	627	Biological age, fitness biomarkers	R code for DNAm FitAge
MRS	10	Mortality risk score	methylCIPHER, methylClock
DunedinPACE	173	Pace of aging	R code for DunedinPACE

HOC: Horvath Online Calculator https://DNAmAge.clockfoundation.org/; R code for DNAmFitAge: https://github.com/kristenmcgreevy/DNAmFitAge; R code for DunedinPACE: https://github.com/danbelsky/DunedinPACE

Genome-wide association analysis

Genotyping of DNA samples was conducted using the Illumina (San Diego, CA) Global Screening Array (GSA 24V3). Primary quality control was done using GenomeStudio 2.0 software and variants with a call rate of 0% were removed. Further quality control was conducted using PLINK V1.9 based on the 0.95 call rate for samples and variants [48]. Also, variants were filtered based on minor allele frequencies (MAF < 0.01) and deviation from Hardy–Weinberg equilibrium (HWE, $p < 1.0 \times 10^{-6}$). Biallelic SNPs mapped to a unique genomic location based on GRCh37 were retained for further analysis. Sample relatedness was checked, and principal component analysis (PCA) was done on the pruned SNPs to detect population stratification and genetic outliers. Imputation was done using Beagle version 5.4 [49].

The genome-wide association study (GWAS) was conducted using 719 blood samples and 477,827 quality-controlled and imputed SNPs on autosomal chromosomes to find significant associations with different epigenetic age acceleration measures including IEAA, EEAA, Skin&BloodAgeAccel, GrimAgeAccel, PhenoAgeAccel, FitAgeAccel, PACE, and PoAm. Multivariate linear regression adjusted for age and sex was used and the $p < 5 \times 10^{-8}$ (FDR-adjusted $p < 0.05$) was considered for genome-wide statistical significance.

Results

The effect of different preprocessing methods on the mean absolute error of age prediction

The accuracy of each DNAm age clock was assessed through the examination of the impact of different preprocessing approaches, with a focus on comparing the mean absolute error (MAEs). In blood samples, the smallest MAE was obtained for the DNAm age Skin&Blood (MAE = 2.47) after ssNoob normalization. Also for buccal swab samples, the Skin&Blood clock showed the highest accuracy of calendar age prediction after ssNoob normalization (MAE = 3.86), while in general the calculated MAEs for all DNAm age clocks were higher for buccal samples compared to blood samples. The results for the PC version of all DNAm clocks for buccal swab samples, showed a higher MAE compared to the original models, and the same results were obtained for blood samples, but exceptionally the PC version of the PhenoAge showed a smaller MAE than the original models for all normalization methods except quantile normalization and the PC version of Hannum age showed a smaller MAE compared to the original model after ssNoob and Funnorm normalization. The calculated MAEs for all models and all normalization methods showed consistency between methylclock and methylCIPHER R packages. Detailed results of the comparison of the effect of different preprocessing methods for different clocks are summarized in Tables S2&S3 and illustrated in Fig. S1.

Correlation among different epigenetic age acceleration measures

Assessing pairwise correlation indicated a significant correlation ($r > 0.90$, $p < 0.05$) between chronological age and all DNAm ages. The mean and range of each DNAm measure were compared among three age categories: early adulthood (20–39 years old), middle adulthood (40–59 years old), and old age (60 + years old). It was observed that all DNAm ages increased with age as expected, while the trend of EAA changes weakly corresponded to the age groups. However, the mean of EAAs in middle adulthood was marginally higher than in both early adulthood and old age groups. Also, the mean of DunedinPACE, PoAm, and MRS showed a slight increase with age categories from early adulthood to old age, and the categorical form of the mortality risk score showed a broader range in old ages. The summary statistics of each measure are reported in Table S4 and Figs. S2 & S3. The full report of pairwise correlation analysis among all DNAm Age measures and calendar age is illustrated in Table S5 and Fig. S4.

Assessing the correlation between different DNAm age acceleration measures along with the pace of aging (DunedinPACE) and the mortality risk score (MRS) showed a varied range of correlation coefficients with the lowest correlation obtained between Skin&BloodAgeAccel and MRS ($r = 0.05$) followed by DunedinPACE ($r = 0.06$) and highest correlation between GrimAgeAccel and FitAgeAccel ($r = 0.61$) followed by PACE ($r = 0.53$). Detailed results for pairwise correlations between different EAAs,

DunedinPACE, and MRS are shown in Fig. 1 and Table S6.

Assessing the association between sociodemographic characteristics, different lifestyle-related risk factors, and EAAs

The association of each sociodemographic characteristic or lifestyle-related risk factor with different EAA measures was examined using linear regression models adjusted for age and sex, Table 2. The full models, adjusted for age, sex, smoking status, BMI, and physical activity, together with DNAm-based estimations of blood cell compositions, were used for assessing the association of each sociodemographic characteristic with different EAAs (Table 3).

Male sex showed a significant association with accelerated IEAA ($p=0.02$), GrimAgeAccel ($p=7.12\times10^{-20}$), DunedinPACE ($p=0.001$), and MRS ($p=1.75\times10^{-7}$), but it was significantly associated with younger FitAge ($p=5.37\times10^{-11}$). Higher BMI showed significant association with higher EAA values obtained from the second and third generations of DNAm age clocks, PhenoAgeAccel ($p=0.04$), GrimAgeAccel ($p=0.01$), FitAgeAccel ($p=0.02$), DunedinPACE ($p=6.20\times10^{-14}$), and DunedinPoAm ($p=0.01$), but not with MRS. In addition, a significant association of smoking status showed a consistent and dose-dependent effect of cigarette smoking on EAA. Former smoking status had a smaller effect on GrimAgeAccel (std. beta = 0.1, $p=0.003$), FitAgeAccel (beta = 0.09, $p=0.01$), and DunedinPoAm (beta = 0.1, $p=0.007$) than current smoking status (GrimAgeAccel, beta = 0.41, $p=2.40\times10^{-31}$; FitAgeAccel, beta = 0.28, $p=6.40\times10^{-14}$, DunedinPoAm, beta = 0.29, $p=1.00\times10^{-15}$) (Fig. 2); whereas, only current smoking status, but not former smoking, showed a significant association with DunedinPACE ($p=3.10\times10^{-6}$) and MRS ($p=7.13\times10^{-6}$). Physical activity was associated with slower epigenetic aging in the studied population, as measured by all clocks tested. Importantly, the effect size of the association

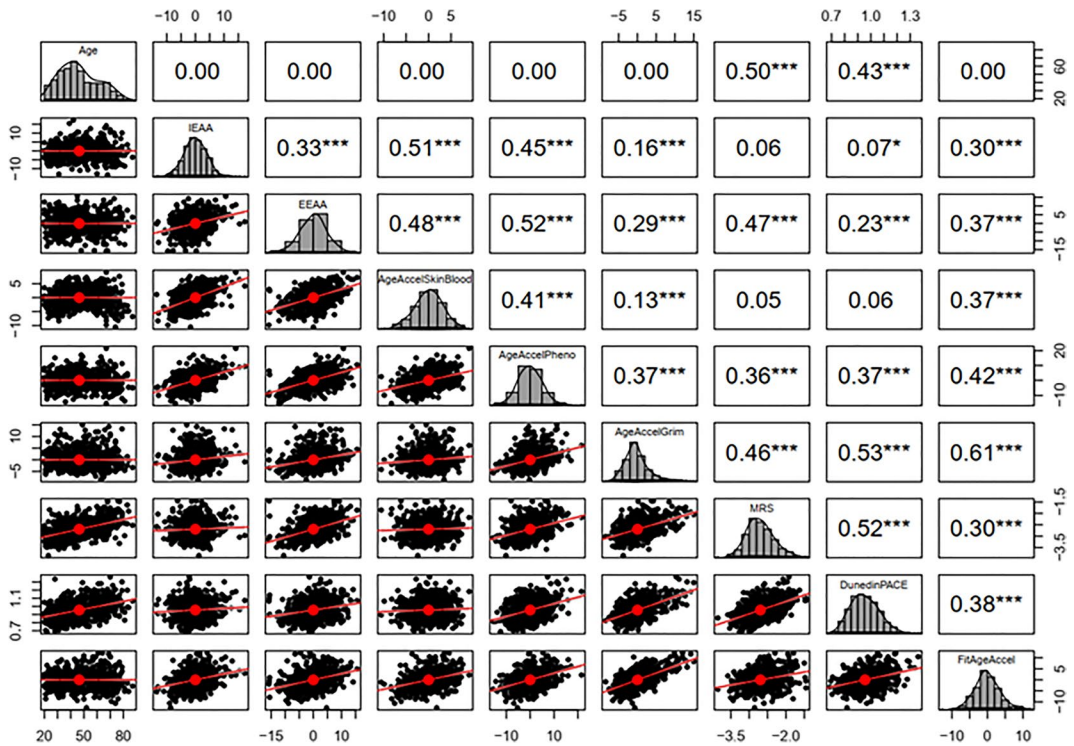

Fig. 1 Pairwise correlation among DNAm Age acceleration measures for different DNAm age clocks. The significance of correlation ($p<0.05$) is indicated with asterisks (no stars: not statistically significant, one to three stars: significant respectively at 0.1, 0.05, and 0.01 levels). Density plots indicate the distribution of each of the DNAm Age acceleration measures in the Polish population

Table 2 Associations of EAAs with lifestyle factors (one at a time) adjusted for sex and age (Model 1)

	IEAA		EEAA		PhenoAgeAccel		GrimAgeAccel		FitAgeAccel		DunedinPACE		DunedinPoAm		MRS	
	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*
Age	0.008	0.83	0.02	0.56	0.007	0.85	0.03	0.38	−0.02	0.53	**0.44**	**3.5e−36**	**0.29**	**6.7e−16**	**0.52**	**9.6e−52**
SexMale	**0.08**	**0.02**	0.22	0.79	0.07	0.05	**0.33**	**7.1e−20**	**−0.24**	**5.4e−11**	**0.11**	**0.001**	0.06	0.12	**0.17**	**1.8e−07**
Education																
University/no	0.07	0.05	0.008	0.81	−0.001	0.97	**−0.17**	**1.4e−06**	**−0.14**	**0.0001**	**−0.16**	**1.22e−06**	**−0.11**	**0.002**	−0.06	0.05
SES-low																
SES-medium	0.04	0.64	0.007	0.93	0.10	0.22	−0.07	0.40	−0.06	0.47	−0.12	0.13	−0.06	0.48	0.05	0.46
SES−high	0.05	0.54	−0.03	0.73	0.11	0.20	**−0.15**	**0.05**	−0.16	0.06	**−0.18**	**0.02**	−0.13	0.12	−0.02	0.83
Residence-village																
City < 100 K	−0.02	0.63	0.02	0.63	0.08	0.10	0.01	0.76	0.01	0.83	0.01	0.76	0.03	0.53	0.02	0.69
City > 100 K	−0.009	0.85	−0.02	0.67	0.009	0.84	0.03	0.55	0.004	0.93	−0.03	0.43	−0.0007	0.99	0.04	0.35
Manual work																
MS	0.007	0.90	−0.07	0.20	−0.01	0.85	**−0.11**	**0.03**	**−0.13**	**0.02**	**−0.13**	**0.01**	−0.08	0.11	−0.09	0.06
LMS	0.08	0.14	−0.07	0.20	0.02	0.72	**−0.12**	**0.02**	**−0.14**	**0.01**	−0.1	0.06	**−0.12**	**0.02**	−0.07	0.15
Workplace RF																
SS	0.01	0.70	−0.01	0.79	−0.05	0.16	−0.01	0.77	−0.01	0.79	−0.03	0.31	−0.04	0.29	−0.002	0.95
IR	0.03	0.36	0.03	0.38	−0.02	0.51	0.04	0.27	0.05	0.15	−0.006	0.86	0.02	0.58	0.02	0.54
EP	−0.05	0.19	−0.0002	0.99	−0.05	0.16	−0.02	0.61	0.01	0.72	−0.03	0.34	0.003	0.94	0.01	0.67
SE	−0.005	0.90	−0.02	0.65	−0.07	0.06	−0.007	0.84	0.03	0.33	−0.05	0.15	−0.05	0.20	−0.04	0.21
HT	−0.02	0.60	0.05	0.22	0.006	0.87	0.06	0.10	**0.12**	**0.002**	0.06	0.07	0.02	0.57	0.02	0.48
LT	−0.03	0.42	0.02	0.66	−0.05	0.21	0.04	0.24	**0.10**	**0.007**	−0.0007	0.98	0.02	0.54	−0.004	0.91
Lifestyle RF																
Stress	0.06	0.15	0.03	0.39	−0.008	0.84	−0.005	0.89	0.04	0.24	−0.004	0.91	−0.04	0.32	−0.02	0.58
Sleep hours	0.01	0.72	−0.003	0.94	−0.05	0.20	**−0.09**	**0.007**	−0.06	0.07	−0.04	0.26	−0.05	0.18	**−0.08**	**0.01**
Physical Activity	−0.04	0.30	−0.06	0.12	**−0.14**	**1.4e−04**	**−0.17**	**2.8e−06**	**−0.13**	**4.1e−04**	**−0.13**	**1.4e−04**	**−0.07**	**0.07**	−0.05	0.14
PA freq. (none)																
Once a month	−0.05	0.27	−0.06	0.18	−0.06	0.15	**−0.08**	**0.04**	**−0.10**	**0.01**	−0.06	0.13	−0.06	0.12	−0.06	0.1
Once a week	−0.09	0.09	−0.08	0.13	**−0.16**	**0.002**	**−0.17**	**3.8e−04**	**−0.16**	**0.001**	**−0.12**	**0.01**	**−0.11**	**0.03**	−0.05	0.22
2–3 times a week	−0.06	0.31	−0.10	0.06	**−0.24**	**2.2e−05**	**−0.27**	**5.9e−07**	**−0.21**	**1.40e−04**	**−0.22**	**2.2e−05**	**−0.13**	**0.02**	**−0.11**	**0.02**

Table 2 (continued)

	IEAA		EEAA		PhenoAgeAccel		GrimAgeAccel		FitAgeAccel		DunedinPACE		DunedinPoAm		MRS	
	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*
Every day	−0.07	0.19	−0.10	0.05	**−0.14**	**0.006**	**−0.22**	**5.3e−06**	**−0.22**	**1.34e−05**	**−0.16**	**5.2e−04**	−0.08	0.13	−0.08	0.06
Yoga	−0.01	0.70	0.003	0.93	−0.03	0.44	−0.07	0.06	−0.06	0.10	**−0.10**	**0.003**	−0.01	0.68	−0.02	0.57
Yoga v2	−0.16	0.77	0.01	0.79	−0.01	0.76	−0.03	0.44	−0.02	0.64	**−0.08**	**0.02**	−0.01	0.74	−0.003	0.92
Aerobic	0.001	0.98	−0.06	0.17	−0.05	0.27	**−0.10**	**0.01**	−0.06	0.15	**−0.08**	**0.04**	**−0.09**	**0.02**	−0.07	0.05
Strengthening	0.03	0.63	−0.11	0.06	−0.07	0.25	−0.09	0.14	−0.04	0.55	−0.08	0.12	−0.09	0.11	−0.08	0.10
BMI	0.02	0.55	−0.03	0.43	**0.08**	**0.04**	**0.09**	**0.01**	**0.09**	**0.02**	**0.26**	**6.2e−14**	**0.10**	**0.01**	−0.02	0.57
Never smoker																
Former smoker	0.04	0.32	−0.008	0.84	0.01	0.75	**0.10**	**0.003**	**0.09**	**0.01**	0.03	0.36	**0.10**	**0.007**	0.02	0.59
Current smoker	−0.04	0.36	−0.03	0.51	0.81	0.81	**0.41**	**2.4e−31**	**0.28**	**6.4e−14**	**0.16**	**3.1e−06**	**0.29**	**1.0e−15**	**0.15**	**7.1e−06**
Alcohol-non																
Occasionally	−0.006	0.90	−0.05	0.26	−0.003	0.94	0.04	0.42	−0.002	1.00	−0.04	0.40	0.03	0.53	−0.004	0.92
Frequently	0.03	0.54	−0.07	0.13	0.04	0.36	−0.005	0.92	−0.06	0.20	−0.07	0.10	0.008	0.86	−0.01	0.78
Diet-Varied																
Vegetarian	−0.04	0.32	0.03	0.45	−0.006	0.87	−0.04	0.30	**−0.07**	**0.04**	−0.01	0.73	**0.07**	**0.04**	−0.02	0.58
Protein	0.02	0.65	0.05	0.16	0.01	0.72	0.03	0.42	0.01	0.72	−0.003	0.93	0.03	0.44	0.03	0.33
Meals (2 to 3)																
3 to 4	0.02	0.60	0.03	0.44	0.06	0.15	−0.07	0.12	−0.06	0.16	−0.02	0.70	**−0.09**	**0.04**	0.0006	0.99
4 to 6 or more	0.002	0.96	0.02	0.62	0.05	0.26	−0.003	0.94	−0.02	0.58	−0.03	0.49	−0.03	0.53	0.02	0.65
Meat (never)																
1–2 times a week	0.07	0.39	0.01	0.91	0.04	0.57	0.11	0.15	**0.16**	**0.03**	0.01	0.89	−0.07	0.33	0.12	0.07
3–4 times a week	0.01	0.88	−0.01	0.89	0.01	0.88	0.12	0.14	0.15	0.06	−0.004	0.96	−0.14	0.08	0.07	0.31
Frequently	0.02	0.87	−0.05	0.58	−0.03	0.73	0.14	0.12	**0.18**	**0.04**	0.07	0.37	−0.1	0.27	0.1	0.21
Everyday	0.1	0.33	0.07	0.49	0.11	0.25	**0.19**	**0.03**	**0.28**	**0.003**	0.09	0.33	−0.12	0.19	0.11	0.19
Fish (never)																
Rarely	−0.02	0.88	0.008	0.95	−0.09	0.46	−0.08	0.52	−0.05	0.67	−0.02	0.86	−0.16	0.19	−0.13	0.22
Sometimes	0.006	0.97	−0.02	0.90	−0.10	0.43	−0.12	0.35	−0.04	0.74	−0.06	0.60	−0.19	0.13	−0.16	0.16
Frequently	0.01	0.91	−0.03	0.76	−0.14	0.19	−0.15	0.12	−0.10	0.35	−0.12	0.20	−0.20	0.05	−0.13	0.14
Portion of Veg. (0)																

Table 2 (continued)

	IEAA		EEAA		PhenoAgeAccel		GrimAgeAccel		FitAgeAccel		DunedinPACE		DunedinPoAm		MRS	
	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*
1 to 2	0.002	0.98	0.07	0.42	−0.04	0.69	**−0.21**	**0.01**	**−0.18**	**0.04**	−0.12	0.13	−0.02	0.85	0.03	0.69
3 to 4	−0.03	0.76	0.02	0.81	−0.07	0.46	**−0.23**	**0.01**	**−0.20**	**0.02**	−0.14	0.08	−0.01	0.88	0.003	0.97
≥5	0.03	0.49	0.05	0.23	−0.004	0.93	**−0.09**	**0.03**	−0.08	0.08	−0.02	0.66	−0.007	0.87	0.005	0.91
Portion of fruit (0)																
1 to 2	0.01	0.82	−0.06	0.24	−0.001	0.98	−0.03	0.64	−0.06	0.30	0.005	0.92	0.01	0.81	−0.01	0.84
3 to 4	0.002	0.97	−0.09	0.10	−0.06	0.32	−0.08	0.12	−0.09	0.10	0.006	0.91	0.02	0.66	−0.06	0.22
≥5	0.02	0.70	−0.02	0.55	−0.008	0.84	0.03	0.48	0.03	0.39	0.02	0.64	0.02	0.61	−0.03	0.40
Cups of coffee (0)																
1 to 2	0.03	0.49	**−0.14**	**0.003**	**−0.11**	**0.03**	0.02	0.65	0.03	0.58	−0.04	0.33	0.01	0.78	−0.08	0.06
3 to 4	0.05	0.33	−0.05	0.28	−0.01	0.81	0.05	0.25	0.08	0.09	−0.06	0.17	0.04	0.35	−0.02	0.66
≥5	−0.03	0.46	**−0.12**	**0.001**	−0.06	0.10	0.02	0.64	−0.04	0.25	−0.07	0.05	0.01	0.71	−0.05	0.18

IEAA intrinsic epigenetic age acceleration, *EEAA* extrinsic epigenetic age acceleration, *GrimAgeAccel* DNAm GrimAge acceleration, *PhenoAgeAccel* DNAm PhenoAge acceleration, *DunedinPACE* pace of aging. Education status: university degree versus no university degree; type of job: sedentary mental (SM), long hours sedentary mental (LSM), manual (M); workplace RFs: risk factors in the workplace including severe stress (SS), ionizing radiation (IR), environmental pollutants: tocsins/heavy metals/air pollution/dioxins/cigarette (EP), sun exposure (SE), high temperature (HT), low temperature (LT); PA: physical activity; Veg.: vegetable; SES: socioeconomic status; Std.B: the standardized regression coefficient of the respective variable from the regression model as defined below; *p*: Significant *p* values (< 0.05). Model: EAA ~ age + gender + lifestyle risk factor (one at a time); In Yoga v2, results were additionally adjusted for physical activity, frequency of meat and vegetable consumption, hours of sleep, and cups of coffee per day

Table 3 Associations of sociodemographic characteristics with EAAs adjusted for lifestyle risk factors and blood cell compositions using multivariable linear regression models

	IEAA		EEAA		PhenoAgeAccel		GrimAgeAccel		FitAgeAccel		DunedinPACE		DunedinPoAm		MRS	
	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*	Std.B	*p*
University degree	**0.08**	**0.04**	−0.01	0.66	0.005	0.89	**−0.11**	**2e−4**	**−0.09**	**0.01**	**−0.13**	**4.7e−5**	**−0.08**	**0.02**	**−0.07**	**8e−4**
SES-low																
SES-medium	0.06	0.49	−0.02	0.77	0.13	0.10	0.003	0.97	0.01	0.88	−0.03	0.64	−0.005	0.95	0.03	0.55
SES-high	0.07	0.41	−0.06	0.40	0.14	0.07	−0.06	0.43	−0.06	0.46	−0.07	0.31	−0.05	0.51	−0.03	0.53
Village																
City < 100 K	−0.02	0.68	0.02	0.66	0.09	0.05	0.05	0.22	0.02	0.68	0.02	0.57	0.04	0.31	0.01	0.62
City > 100 K	−0.01	0.84	−0.01	0.88	0.03	0.53	0.04	0.36	0.007	0.87	−0.02	0.68	−0.008	0.83	0.04	0.13
Manual work																
SM	−0.007	0.91	−0.03	0.54	0.02	0.64	−0.05	0.23	−0.08	0.14	−0.09	0.05	−0.04	0.34	−0.03	0.34
LSM	−0.06	0.28	−0.06	0.16	0.02	0.73	−0.06	0.17	−0.08	0.11	−0.06	0.17	−0.10	0.04	−0.04	0.18

IEAA intrinsic epigenetic age acceleration, *EEAA* extrinsic epigenetic age acceleration, *GrimAgeAccel* DNAm GrimAge acceleration, *PhenoAgeAccel* DNAm PhenoAge acceleration, *PACE* pace of aging, *MRS* Mortality Risk Score; education status university degree compared to no university degree; residency: area of residency in a city with a population smaller than 100,000 (City < 100 K), city with a population over 100,000 (City > 100 K); job: sedentary mental (SM), long hours sedentary mental (LSM), manual (M); *CD8T* CD8 + T lymphocyte, *CD4T* CD4 + T lymphocyte, *NK* natural killer lymphocyte, *B cell* B lymphocytes, *Mono* monocyte, *Gran* granulocytes, *Std.B* the standardized regression coefficient of the respective variable from the regression model as defined bellow, *p* significant *p* values (< 0.05). Model: EAA ~ age + gender + Sociodemographic factor + Smoking + BMI + PA + CD8T + CD4T + NK + Bcell + Mono + Gran

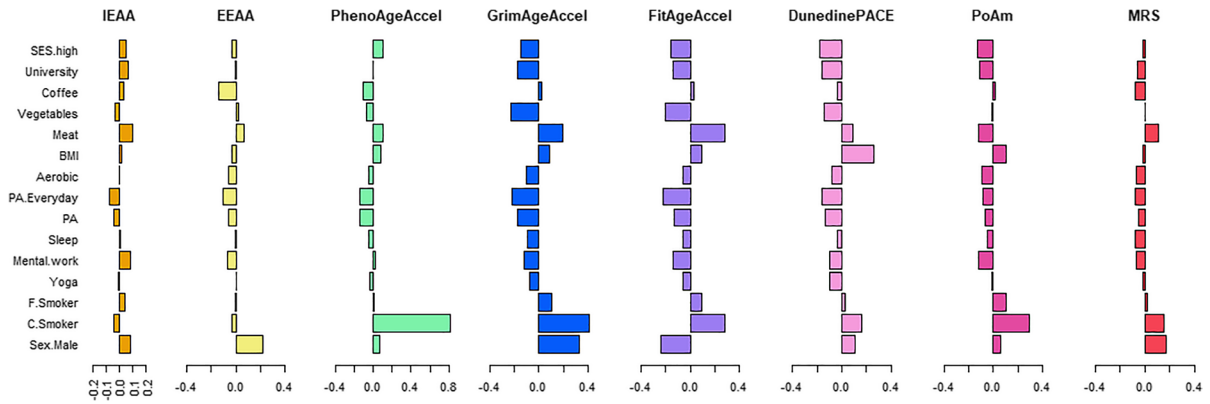

Fig. 2 Presentation of the value and direction of effect sizes obtained for age- and sex-adjusted association analysis of selected lifestyle and sociodemographic risk factors (each variable at a time) with EAAs. The X-axis shows the range of each EAA measure. The Y axis indicates the standardized regression coefficient of the respective variable from the regression model adjusted for age and sex. Variable include Education status: University degree versus no university degree; SES: High socioeconomic status compared to low SES; type of job: long hours sedentary mental (LSM) compared to manual job; (PA) physical activity versus inactivity; PA.Everyday: everyday physical activity vs. inactivity; C.Smoker: current smoker versus never smokers; F.Smoker: former smoker vs. never smokers; Coffee: 1–2 cups of coffee vs. no coffee; Meat: everyday meat consumption vs. no meat consumption; Vegetables: 3–4 portion of vegetables vs. no vegetable consumption; Sleep: more than 8 h of sleep vs. less than or equal to 8 h of sleep; Aerobic: aerobic exercises vs. no aerobic

increased with the frequency of physical activity, with daily activity showing a highly significant association for PhenoAgeAccel ($p=0.006$), GrimAgeAccel ($p=5.3\times10^{-6}$), FitAgeAccel ($p=1.34\times10^{-5}$), and PACE ($p=5.2\times10^{-4}$). Considering different types of sports, aerobic exercise was significantly associated with GrimAge and PACE. Notably, yoga practice was significantly associated with slower DunedinPACE (std. beta $=-0.1$, $p=0.003$), and the result remained significant after additional adjustment for physical activity, frequency of meat and vegetable consumption, hours of sleep, and cups of coffee per day ($p=0.02$, Table 2). Increased sleeping hours showed significant association with decreased GrimAgeAccel ($p=0.007$) and lower mortality risk score ($p=0.01$). We found a significant decrease in epigenetic aging with vegetable consumption. Individuals who consumed at least 1–2 servings of vegetables per day had slower GrimAge and FitAge ($p=0.01$ and 0.04, respectively). Frequent meat consumption was found to accelerate the process of epigenetic aging, showing a significant and positive association with GrimAge ($p=0.03$) and FitAge ($p=0.003$) when daily meat consumption was recorded. This result is consistent with a positive effect of vegetarianism on reduced FitAgeAccel ($p=0.04$). Interestingly, drinking 1–2 cups of coffee per day was associated with reduced EEAA and PhenoAgeAccel in our population ($p=0.003$ and 0.03, respectively). The comparison of direction and effect size obtained for the association of selected factors with EAAs is shown in Fig. 2.

Educational level (university degree versus no university degree) was robustly associated with reduced epigenetic aging across different clocks, including GrimAgeAccel ($p=1.38\times10^{-6}$), FitAgeAccel ($p=0.0001$), DunedinPACE ($p=1.22\times10^{-6}$), and DunedinPoAm ($p=0.002$) in the age and sex-adjusted models (Table 2). In the fully adjusted model university degree was significantly associated with decreased GrimAgeAccel ($p=2.0\times10^{-4}$), FitAgeAccel ($p=0.01$), DunedinPACE ($p=4.7\times10^{-5}$), DunedinPoAm ($p=0.02$), and MRS ($p=8.00\times10^{-4}$), but with accelerated IEAA ($p=0.04$, Table 3). Comparison of self-reported high socioeconomic status (SES) with low SES, showed a significant association with reduced DunedinPACE ($p=0.02$), while no significant effects remained in the full model (Table 3). The area of residence, comparing residency of cities with a population greater or less than 100,000 residents to those living in villages, showed no significant association with any of the DNAm measures in reduced or fully adjusted models in Model 1 or 2. Age- and sex-adjusted long hours sedentary mental work was significantly associated with reduced DNAm GrimAge

($p=0.02$), FitAge ($p=0.01$), and DunedinPoAm ($p=0.02$), but not in a full model. In addition, in model 1, being exposed to extremely high or low temperatures in the workplace was associated with higher FitAgeAccel ($p=0.002$, 0.007 respectively).

Genome-wide association study identifies novel genetic variants for EAA

The genome-wide association analysis for different epigenetic age acceleration measures was performed on 477,827 SNPs using additive genetic models implemented in PLINK [48] and the multivariate linear regression adjusted for age and sex. We found a genome-wide significant association between rs73218878, located in the *SOCS2* gene, and DNAm GrimAge acceleration ($p=2.87\times10^{-8}$). This result remained statistically significant after accounting for other lifestyle factors aside from sex and age, such as smoking (FDR = 0.01, $p=3.1\times10^{-8}$) and BMI (FDR = 0.03, $p=5.9\times10^{-8}$). The result is illustrated in the Manhattan and QQ plots in Fig. 3.

In addition to the genome-wide significant result at $p<5\times10^{-8}$, 10 additional SNPs were associated with GrimAgeAccel at a suggestive threshold ($p<1\times10^{-5}$) (Table 4). Among these SNPs, rs118072622 located at chromosome 6, and mapped to the Crystallin Beta-Gamma Domain Containing 1 (*CRYBG1*) gene had the strongest association ($p=1.7\times10^{-6}$).

Discussion

In this study, we showed extensive research presenting a wide list of environmental and

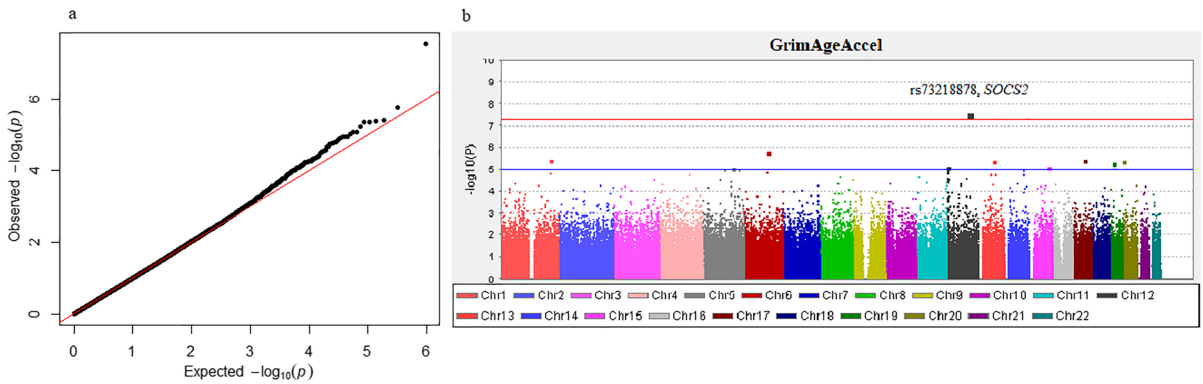

Fig. 3 **a** QQ-plot for the GWAS on GrimAgeAccel. **b** Manhattan plot showing the results of GWAS analysis for GrimAgeAccel. The GrimAgeAccel showed a genome-wide significant association ($p=2.87\times10^{-8}$) with rs73218878 after adjusting for age and sex. The red line indicates the genome-wide significant level ($p=5\times10^{-8}$) and the blue line indicates $p=1\times10^{-5}$

Table 4 Top 10 SNPs ($p<10^{-5}$) for GrimAgeAccel

Number	SNP	Chr	Position	Closest reference gene	Alleles	*p*
1	rs73218878	12	93975417	SOCS2	C > A, T	2.87e − 08
2	rs118072622	6	106912438	CRYBG1	T > C	1.74e − 06
3	rs11120686	1	216050785	USH2A	T > C	3.84e − 06
4	rs117658875	17	50219309	CA10	G > A	4.20e − 06
5	rs17520509	13	67780695	PCDH9	A > C	4.38e − 06
6	rs151037173	20	3089900	UBOX5	A > C	4.46e − 06
7	rs118104025	19	19146732	ARMC6	G > A, T	5.78e − 06
8	rs117006309	15	87952919	LOC105370956	A > C, G	8.24e − 06
9	rs10744533	12	1236286	ERC1/LOC124902857	T > A, C, G	8.30e − 06
10	rs11742455	5	133599786	CDKL3	C > T	9.45e − 06

demographic factors analyzed for associations with epigenetic aging using an exhaustive list of available epigenetic clocks, while at the same time, we analyzed the technical aspects of obtaining reliable results for estimating epigenetic age and other related parameters. We applied three different types of epigenetic predictors, including the most recently introduced epigenetic biomarker of fitness, DNAmFitAge, and the updated speedometer of the pace of aging, DunedinPACE to quantify biological aging in a cohort of > 700 individuals in our dataset. Our results, showing a strong correlation between estimated DNAm age values and the chronological age, are consistent with the literature and validate the utility of DNAm Age clocks in this dataset. Our results emphasized the effect of different methylation array preprocessing methods on the accuracy of DNAm age estimations among different clocks. In our dataset, the DNAm Skin & Blood clock outperformed other models for accurate estimation of chronological age, providing the lowest MAE after ssNoob preprocessing for both blood and buccal swab samples. In addition, our study provided evidence for an association between various sociodemographic and lifestyle-related factors and EAA measures. In particular, we showed for the first time that yoga practice is significantly associated with lower DunedinPACE and that coffee consumption may be beneficial in reducing EEAA and PhenoAgeAccel levels.

The effect of different normalization methods, different types of DNAm clocks, and available R packages on the accuracy of chronological age estimation

DNA methylation arrays are being prominently used for capturing age-related DNA methylation changes. These changes are pervasive throughout the methylome and can be captured by the available cost-effective methylation array platforms such as the updated Infinium Methylation EPIC array which interrogates > 850,000 CpGs. However, the quality of the array data can be affected by experimental and technical factors. Our EPIC array data analysis highlighted the importance of preprocessing strategies for obtaining an accurate estimation of chronological age. Comparing the MAE obtained from different DNAm clocks after applying five different preprocessing methods showed that for chronological age estimation, the Skin&Blood clock outperforms the Horvath2013 and Hannum models, and the second generation of DNAm clocks, PhenoAge and GrimAge, following the ssNoob method (MAE = 2.47 for blood samples and 3.86 for buccal swab samples). This result is consistent with previous work introducing the Skin&Blood clock as a robust age predictor of blood or buccal swab samples for forensic applications (reporting the median absolute deviation of 2.5 and 2 years between DNAm age and chronological age in blood and buccal swab samples, respectively) [6]. Besides, partly because of the evolving field of methylation array technology, different computational methods, including online web calculators and different R packages, are introduced to calculate the measure for several existing DNAm age clocks. As a result, the estimated age for a specific model may show deviation between different calculators. Testing available R packages for obtaining the DNAm Ages showed slight differences at the individual level while the calculated MAE showed consistency between methylclock and methylCIPHER R packages.

In addition, we compared the performance of different DNAm age clocks with their PC version, the principal component-based clocks (PC-clocks) which are suggested as an alternative method to control the reported age estimation noise and enhance the reliability of the results of the DNAm age clocks [41]. The estimated age for PC-clocks after applying the same optimized quality control and normalization steps indicated in general a smaller MAE for original clocks compared to their respective PC-versions, suggesting that the PC-clocks did not improve the chronological age estimation unless for PhenoAge clock with different normalization methods and Hannum age clock after ssNoob or Funnorm normalization. On the other hand, PC versions revealed an enhanced reliability of age estimation between replicates. A comparison of 47 pairs of technical replicates revealed a higher noise for DNAm age estimated by the original clocks versus PC versions. For instance, the Horvath 2013 clock showed an average of 2.74 deviations, ranging from 0.15 to 9.55 years, for 47 analyzed pairs of technical replicates while the agreement between these samples

for the PC-Horvath version was improved on average by 2 years. Our results support that PC clocks may be particularly useful for longitudinal monitoring of biological aging or for tracking the effects of rejuvenating interventions in clinical trials.

Correlation between different DNAm Age and epigenetic age acceleration measures

The results of the pairwise correlation between chronological age and estimated DNAm age were consistent with previous studies indicating a strong correlation for all clocks (r above 0.90). Assessing the correlation between respective age-adjusted epigenetic age acceleration measures (EAAs), along with the pace of aging (DunedinPACE) and the mortality risk score (MRS) exhibited varied and weaker correlations. The weakest, non-significant correlation was found between Skin&BloodAgeAccel and MRS followed by Skin&BloodAgeAccel and DunedinPACE, while the highest correlation was found between GrimAgeAccel and FitAgeAccel ($r = 0.61$). These results suggest that the varied range of correlations between different EAA measures obtained from various clocks probably reflect their different capability in capturing different aspects of biological aging. We observed the highest correlations among second and third-generation clocks and the smallest correlation between first and third-generation clocks. This result was expected since DNAm clocks with improved accuracy for forensic age estimation (Skin&Blood clock) are constructed from CpGs strongly correlated with calendar age, which may limit their capability of capturing mortality or morbidity risk. Consistent with our observation, previous studies also reported that EAA from different clocks are not strongly correlated [20, 21]. In this regard, a systematic evaluation of the correlation between three different epigenetic clocks reported that their measures were correlated with each other in the range of 0.3 to 0.5, suggesting that the same chronological age individuals varied in measures of biological aging calculated by epigenetic clocks [50].

Association between sociodemographic characteristics and EAA measures

Different sociodemographic and lifestyle-related risk factors have been studied as determinants of all-cause mortality and aging-related diseases [51]. In our study, the highest significance and the largest effect sizes for the association with epigenetic age acceleration were observed for smoking. The association was found for five of the eight clocks analyzed, and the effect was stronger in current smokers than in former smokers. This is consistent with other research indicating that smoking is the predominant cause of epigenetic aging [14, 38, 52]. The second factor that showed a strong influence on EAA measures in our population is physical activity. Daily exercise was associated with reduced PhenoAge, FitAge, GrimAge, and PACE. This is not surprising, as there are several studies in the literature showing the beneficial effects of regular exercise on slowing epigenetic aging [53–55]. Although we did not observe a statistically significant association of EAA with strength sport type in this dataset, a previous study conducted on a smaller group of our population enriched for male bodybuilders confirmed the positive impact of this type of sport on DNAmFitAge [17]. Interestingly, accelerated epigenetic aging has been observed elsewhere in elite athletes characterized by extremely intense physical training, but a complex impact of intense exercises on aging has been hypothesized, suggesting that increased DNA methylation at selected loci may potentially contribute to the observed lower risk of cardiovascular disease and cancer in elite athletes [56].

Importantly, practicing yoga or meditation as a mind–body therapy was found in our study to be significantly associated with a slower pace of aging, and the effect remained significant after adjusting for other lifestyle factors, including diet. To the best of our knowledge, this is the first study directly linking yoga with epigenetic aging and this result fits in well with the decelerating effect of intensive relaxing training on DNAm age, reported in a 60-day longitudinal study but calculated based on a different model using the methylation level of 5 methylation markers only [57]. This result may suggest that DNAm

age clocks may serve as a proper monitoring tool for tracking the effectiveness of relaxing techniques for decreasing mortality and morbidity risk. Interestingly, reduced mean methylation levels of the tumor necrosis factor (*TNF*), and serum level of immunological inflammatory markers were reported in women with psychological distress after practicing yoga for 8 weeks [58, 59]. Also, differentially methylated loci were reported through genome-wide methylation analysis of a group of adolescents with a history of adverse childhood experiences after a one-week multimodal intervention, including daily yoga [60].

Various eating behaviors were also found in our study to correlate well with epigenetic measures of aging. In particular, meat consumption on a daily basis and vegetable consumption were associated with increased and decreased aging, respectively. This is consistent with the finding that a vegetarian diet was associated with slower epigenetic aging. Coffee consumption was also associated in our study with decreased EEAA and PhenoAgeAccel. Reports on the impact of healthy dietary patterns on epigenetic aging are available in the literature and conclude that the consumption of red meat is significantly associated with accelerated DNAm aging, and this is consistent across different clocks [13, 61]. Studies have shown that coffee consumption leaves signatures in DNA methylation of human blood [62]. However, due to its known complex components, coffee is thought to have both good and bad effects on health, and it seems that the amount of coffee consumed matters [63]. To our knowledge, this is the first study linking coffee consumption to lower EEAA and PhenoAgeAccel. Next, in our dataset, longer sleep duration was also associated with lower mortality risk score and slower GrimAge. Sleep duration is a known factor associated with mortality and cardiovascular disease [64] and a recent study showed that short sleep was associated with GrimAge [65].

Our study provided evidence of accelerated DNAm GrimAge and PACE measurement in individuals with self-reported high socioeconomic level compared to those categorized as low SES. However, this finding requires further investigation as it lost its significance when adjusted for other lifestyle factors. There was also a statistically significant decrease in GrimAgeAccel, FitAgeAccel, DunedinPACE, and DunedinPoAm for participants with a university degree. Previous studies reported higher education as a protective causal factor associated with lower GrimAgeAccel [50] and PhenoAgeAccel [66]. However, in contrast to the study that showed a significant decelerating effect of higher education on the IEAA, EEAA, and PhenoAgeAccel [10, 67], in our study, higher education level showed a positive direction of association with IEAA, which is suggested to track both age-related changes in blood cell composition and intrinsic epigenetic changes. However, as the effect was not significant when adjusted for age and sex in model 1, but became marginally significant with additional adjustment for additional lifestyle factors, this result should be treated with caution.

Epigenetic clocks are multifactorial composites which makes the mechanistic understanding of their underlying biological insights more complicated. The clocks are constructed of distinct modules of functionally related CpGs with different time-dependent methylation patterns, which probably signal different biological consequences. Accordingly, different EAA measures were associated with various aging-related diseases. Therefore, in an attempt to better understand the biological meaning and the differences of the signals captured by each clock, multi-omics analysis of purified CD14 + monocytes and dorsolateral prefrontal cortex tissue indicated shared transcriptional associations for Horvath2013, Skin&Blood, Hannum, and PhenoAge clocks enriched in pathways linked to metabolism, immunity, and autophagy. Also, in vitro analysis assigned the accelerated PhenoAge to cellular senescence and mitochondrial dysfunction [68]. In another study, DNAm age clocks were deconstructed into distinct modules of CpGs based on their methylation alterations during aging and in response to reprogramming factors [69]. This showed that the contribution of each category of module to the total value calculated by each clock is different. That is, the values obtained by the first generation of mixed-age and multi-tissue clocks, such as Horvath2013 and Skin&Blood, and partially the Hannum clock, were mostly captured by age-associated, but weakly mortality-associated modules of CpGs. The opposite was true for the second and third generation clocks, such as GrimAge and DunedinPoAm. Yet, the PhenoAge clock, as a second-generation clock, developed to capture biological age and health span based on clinical biomarkers, showed a pattern more like the first generation of chronological age clocks which may introduce the PhenoAge clock as a hybrid model between two generations. It seems that

the distribution of the functional modules in a clock can determine the differential proportions of signal captured by each clock regarding specific aspects of epigenetic aging. This may explain why different EAAs are moderately correlated, and the disagreement in associations with aging-related outcomes or anti-aging interventions. Furthermore, there is still a lack of research focusing on individual cytosines and their actual role in the aging process. A better knowledge of the biological significance of changes at the level of individual cytosines may help in the future to understand the importance of the association of individual clocks with individual demographics and lifestyle type.

SOCS2 is associated with epigenetic age acceleration

Studies have shown that accelerated epigenetic age is influenced by genetic variants, suggesting a potential genetic component in the regulation of epigenetic aging. Previous GWAS analysis on EAAs in blood samples identified five genetic loci associated with IEAA, and three SNPs associated with EEAA and revealed the role of the telomerase reverse transcriptase gene (*TERT*) in epigenetic aging rate and its influence on the overall aging process [70]. A meta-analysis of GWAS studies on EAAs further explored the genetic architecture underlying the EAA and reported several genetic loci associated with IEAA and EEAA [71]. Additionally, a recent GWAS identified 137 genetic loci associated with IEAA, EEAA, PhenoAgeAccel, and GrimAgeAccel [72]. Here we performed GWAS analysis for different EAA measures and found a novel SNP associated with GrimAgeAccel at the GWAS level ($p < 5 \times 10^{-8}$). The rs73218878 located on chromosome 18 is mapped to the SOCS2 gene. This gene encodes a member of the suppressor of cytokine signaling (SOCS) family, which acts as a negative regulator of cytokine receptor signaling through the JAK/STAT pathway. It interacts with the cytoplasmic domain of insulin-like growth factor-1 receptor (IGF1R) and is involved in the regulation of IGF1R-mediated cell signaling. *SOCS2* is considered a negative regulator in the growth hormone/IGF1 signaling pathway [73, 74]. Previous findings have demonstrated that the absence of Socs2 expression leads to reduced lifespan in mice with high-growth characteristics, indicating the involvement of the *SOCS2* gene in regulating aging, potentially through its impact on plasma IGF1 levels [74]. Additionally, *SOCS2* is known to play critical roles in fat deposition, skeletal muscle development, central nervous system development, and biogenesis of mitochondria. Besides, the expression of fatty acid oxidation enzymes, such as MCAD, LCAD, and Cyt C, was found to be reduced in response to *SOCS2*. These reductions were consistent with a decrease in free fatty acids and the impairment of mitochondrial complexes I and III caused by *SOCS2*. Moreover, the inhibition of the JAK2/AMPK pathway, which is implicated in fatty acid oxidation, was shown to block mitochondrial fatty acid oxidation. Collectively, these studies suggest that *SOCS2* negatively influences mitochondrial fatty acid oxidation, and the LepR/JAK2/AMPK pathway plays a pivotal role in this process [75]. This information is interesting because of the disclosed in this study association of the rs73218878 in the *SOCS2* gene with the GrimAgeAccel, suggesting its role in longevity. Importantly, genetic variation in SOCS2 was found to be associated with exceptional human longevity in a GWAS of centenarians (individuals who reached 100 years or more) [76], but to our knowledge, the direct association of *SOCS2* with GrimAgeAccel or other epigenetic measures of aging has not been reported in the literature so far.

Limitations

In this study, self-declared sociodemographic and lifestyle-related risk factors were obtained through individual questionnaires, which are associated with subjective assessment and are known to suffer from bias and measurement inaccuracies. In addition, the lifestyle habits and demographic characteristics measured are often different in different biogeographic populations. In addition, the effect of unknown confounders may alter the measurements of the variables analyzed. These potential limitations may affect the effect size and consistency of associations of risk factors with DNAm age acceleration and other measures. Due to study sample size limitations, we could not perform GWAS replication analyses, but this should be considered in the future.

Conclusions

Here, we present a comprehensive analysis of epigenetic clocks by examining novel DNA methylation and SNP genome-wide data in a well-defined cohort. Overall, our analysis revealed that available and emerging DNAm age clocks vary in their degree of association with different lifestyle or sociodemographic data and can capture various aspects of age-related biological changes. Reliable quantification of biological aging can provide a useful measure for medical research to integrate different biomarkers of the aging process into a single personalized tool, and effectively monitor the benefits of antiaging interventions; therefore, selection of the most appropriate model, taking into account the differences in the underlying CpG sets used to develop these models, is important for interpreting their outputs. In addition, the optimized quality control and preprocessing methods of methylation array data can influence the accuracy of chronological age estimation using DNAm clocks. These considerations are also important for the application of these algorithms in forensic research.

In this study, the most robust associations with lifestyle factors were found for FitAge, GrimAge and, PACE clocks, while the comparison of effect sizes for different lifestyle and sociodemographic factors showed that smoking has the greatest impact on epigenetic aging, followed by BMI, meat and vegetable consumption, and physical activity. We provide evidence of an association between higher levels of education and a biomarker of physical fitness, FitAgeAccel, and other EAA measures. More importantly, we provide the first evidence for the beneficial effects of yoga practice on slowing the pace of aging (PACE), and we show that drinking coffee can help reduce epigenetic aging as measured by the Hannum and PhenoAge clocks. We provide new insights into the underlying genetic architecture of EAA by identifying a genome-wide significant association between the *SOCS2* variant and accelerated GrimAge. Given the known role of the *SOCS2* gene in longevity, this finding may have significant implications for epigenetic rejuvenation research. In addition, the provided summary statistics of epigenetic aging in the Polish population may be useful to compare the trend and causal risk factors of the aging process in different populations.

Acknowledgements The authors express their gratitude to the participants involved in this manuscript.

Author contribution RN conducted the bioinformatic and statistical analysis of the data, contributed to data interpretation, and drafted the manuscript. JR, AP, BW, AM, and MB performed laboratory experiments. KM.G, PP.P, MK, DL, GZ, AI, JAW, MM, PK, and MK collected samples, collected, and interpreted phenotypic data. EP conducted statistical analysis of the data, interpreted the results, and contributed to manuscript preparation. EP, AS, AO, MS, and WB contributed to the study design and coordination and the final interpretation of results. All authors read, evaluated, and approved the final version of the manuscript.

Funding Project no. DOB-BIO10/06/01/2019 is financed by the National Centre for Research and Development within the framework of call 10/2019 related to scientific research and studies for national defense and security.

Data availability The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval and consent to participate The study was approved by the Bioethics Committee of the Jagiellonian University in Krakow (decision no. 1072.6120.132.2018).

Consent for publication Written informed consent forms were obtained from all participants.

Competing of interest The authors declare no competing interests.

References

1. Duan R, Fu Q, Sun Y, Li Q. Epigenetic clock: a promising biomarker and practical tool in aging. Ageing Res Rev. 2022;81: 101743. https://doi.org/10.1016/J.ARR.2022.101743.
2. Noroozi R, et al. DNA methylation-based age clocks: from age prediction to age reversion. Ageing Res Rev. 2021;68. https://doi.org/10.1016/J.ARR.2021.101314. Available: https://pubmed.ncbi.nlm.nih.gov/33684551/. Accessed 20 Jan 2023.
3. Kayser M, Branicki W, Parson W, Phillips C. Recent advances in forensic DNA phenotyping of appearance, ancestry and age. Forensic Sci Int Genet. 2023;65: 102870. https://doi.org/10.1016/J.FSIGEN.2023.102870.
4. Hannum G, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359–67. https://doi.org/10.1016/J.MOLCEL.2012.10.016.

5. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):1–20. https://doi.org/10.1186/GB-2013-14-10-R115/COMMENTS. Available: https://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-10-r115. Accessed 08 Feb 2023.
6. Horvath S, et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. Aging (Albany NY). 2018;10(7):1758. https://doi.org/10.18632/AGING.101508. Available: /pmc/articles/PMC6075434/. Accessed 08 Feb 2023.
7. Horvath S, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci. 2014;111(43):15538–15543. https://doi.org/10.1073/PNAS.1412759111. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1412759111. Accessed 26 Apr 2023.
8. Marioni RE, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015;16(1):1–12. https://doi.org/10.1186/S13059-015-0584-6/FIGURES/4. Available: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0584-6. Accessed 26 Apr 2023.
9. Marioni RE, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. Int J Epidemiol. 2015;44(4):1388–1396. https://doi.org/10.1093/IJE/DYU277. Available: https://academic.oup.com/ije/article/44/4/1388/667600. Accessed 26 Apr 2023.
10. Quach A, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging (Albany NY). 2017;9(2):419. https://doi.org/10.18632/AGING.101168. Available: /pmc/articles/PMC5361673/. Accessed 26 Apr 2023.
11. Horvath S, et al. Accelerated epigenetic aging in Down syndrome. Aging Cell. 2015;14(3):491–495. https://doi.org/10.1111/ACEL.12325. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/acel.12325. Accessed 26 Apr 2023.
12. Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging (Albany NY). 2015;7(12):1198. https://doi.org/10.18632/AGING.100864. Available: /pmc/articles/PMC4712342/. Accessed 26 Apr 2023.
13. Levine ME, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging. 2018;10(4). Available: http://www.aging-us.com. Accessed 26 Jan 2023.
14. Lu AT, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging. 2019;11(2):303–327. https://doi.org/10.18632/AGING.101684. Available: https://www.aging-us.com/article/101684. Accessed 26 Jan 2023.
15. Belsky DW, et al. Quantification of the pace of biological aging in humans through a blood test, the DunedinPoAm DNA methylation algorithm. Elife. 2020;9:1–56. https://doi.org/10.7554/ELIFE.54870.
16. Belsky DW, et al. DunedinPACE, A DNA methylation biomarker of the Pace of Aging. Elife. 2022;11. https://doi.org/10.7554/ELIFE.73420.
17. McGreevy KM, et al. DNAmFitAge: biological age indicator incorporating physical fitness. Aging. 2023;15. https://doi.org/10.18632/AGING.204538. Available: https://www.aging-us.com/article/204538. Accessed 06 Apr 2023.
18. Cohen AA, et al. A complex systems approach to aging biology. https://doi.org/10.1038/s43587-022-00252-6. Available: https://doi.org/10.1038/s43587-022-00252-6. Accessed 26 Jan 2023.
19. Ng T, Carollo J, Tagawa A, Pan Z, Heyn P. Systems aging clock: A novel epigenetic aging clock modeled from organ & bodily function based mortality indices. Innov Aging. 2021;5(Suppl 1):1056. https://doi.org/10.1093/GERONI/IGAB046.3736. Available: /pmc/articles/PMC8682634/?report=abstract. Accessed 23 Apr 2023.
20. Pośpiech E, et al. Introduction of a multiplex amplicon sequencing assay to quantify DNA methylation in target cytosine markers underlying four selected epigenetic clocks. Clin Epigenetics. 2023;15(1). https://doi.org/10.1186/S13148-023-01545-2. Available: https://pubmed.ncbi.nlm.nih.gov/37563670/. Accessed 15 Nov 2023.
21. Belsky DW, et al. Eleven telomere, epigenetic clock, and biomarker-composite quantifications of biological aging: do they measure the same thing? Am J Epidemiol. 2018;187(6):1220. https://doi.org/10.1093/AJE/KWX346. Available: /pmc/articles/PMC6248475/. Accessed 25 Apr 2023.
22. Zbieć-Piekarska R, et al. Development of a forensically useful age prediction method based on DNA methylation analysis. Forensic Sci Int Genet. 2015;17:173–9. https://doi.org/10.1016/J.FSIGEN.2015.05.001.
23. Aliferi A, et al. Combining current knowledge on DNA methylation-based age estimation towards the development of a superior forensic DNA intelligence tool. Forensic Sci Int Genet. 2022;57. https://doi.org/10.1016/J.FSIGEN.2021.102637. Available: https://pubmed.ncbi.nlm.nih.gov/34852982/. Accessed 28 Jan 2023.
24. Freire-Aradas A, et al. A common epigenetic clock from childhood to old age. Forensic Sci Int Genet. 2022;60. https://doi.org/10.1016/J.FSIGEN.2022.102743. Available: https://pubmed.ncbi.nlm.nih.gov/35777225/. Accessed 28 Jan 2023.
25. Woźniak A, et al. Development of the VISAGE enhanced tool and statistical models for epigenetic age estimation in blood, buccal cells and bones. Aging. 2021;13(5):6459–6484. https://doi.org/10.18632/AGING.202783. Available: https://pubmed.ncbi.nlm.nih.gov/33707346/. Accessed 23 Jan 2023.
26. Freire-Aradas A, et al. Development of a methylation marker set for forensic age estimation using analysis of public methylation data and the Agena Bioscience EpiTYPER system. Forensic Sci Int Genet. 2016;24:65–74. https://doi.org/10.1016/j.fsigen.2016.06.005. Available: http://www.fsigenetics.com/article/S1872497316301065/fulltext. Accessed 25 Jan 2023.
27. Pośpiech E, et al. Exploring the possibility of predicting human head hair greying from DNA using whole-exome and targeted NGS data. BMC Genom. 2020;21(1). https://doi.org/10.1186/s12864-020-06926-y. Available: /pmc/articles/PMC7430834/?report=abstract. Accessed 07 Oct 2020.
28. Marcińska M, et al. Evaluation of DNA variants associated with androgenetic alopecia and their potential to predict male pattern baldness. PLoS One. 2015;10(5). https://doi.org/10.1371/journal.pone.0127852. Available: https://pubmed.ncbi.nlm.nih.gov/26001114/. Accessed 07 Oct 2020.

29. Ori APS, Lu AT, Horvath S, Ophoff RA. Significant variation in the performance of DNA methylation predictors across data preprocessing and normalization strategies. Genome Biol. 2022;23(1):1–21. https://doi.org/10.1186/S13059-022-02793-W/FIGURES/3. Available: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-022-02793-w. Accessed 12 May 2023.
30. Wani AH, Armstrong D, Dahrendorff J, Uddin M. RANDOMIZE: a web server for data randomization. Arch Proteom Bioinform. 2020;1(1):31. Available: /pmc/articles/PMC7861512/. Accessed 24 Apr 2023.
31. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics. 2016;8(3):389–399. https://doi.org/10.2217/EPI.15.114/SUPPL_FILE/SUPPL_TABLE10.XLSX. Available: https://www.futuremedicine.com/doi/10.2217/epi.15.114. Accessed 28 Jan 2023.
32. Xu Z, Niu L, Taylor JA. The ENmix DNA methylation analysis pipeline for Illumina BeadChip and comparisons with seven other preprocessing pipelines. Clin Epigenetics. 2021;13(1):1–8. https://doi.org/10.1186/S13148-021-01207-1/FIGURES/2. Available: https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-021-01207-1. Accessed 02 Feb 2023.
33. Aryee MJ, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of infinium DNA methylation microarrays. 2014;30(10):1363–1369. https://doi.org/10.1093/bioinformatics/btu049. Available: http://bioconductor.org/packages/release/bioc/html/minfi.html. Accessed 04 Feb 2023.
34. Fortin JP, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol. 2014;15(11):1–17. https://doi.org/10.1186/S13059-014-0503-2/FIGURES/10. Available: https://genomebiology.biomedcentral.com/articles/10.1186/s13059-014-0503-2. Accessed 04 Feb 2023.
35. Triche TJ, Weisenberger DJ, van den Berg D, Laird PW, Siegmund KD. Low-level processing of Illumina Infinium DNA Methylation BeadArrays. https://doi.org/10.1093/nar/gkt090. Available: http://ideas.repec.org/p/dgr/. Accessed 04 Feb 2023.
36. Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol. 2012;13(6):1–12. https://doi.org/10.1186/GB-2012-13-6-R44/FIGURES/7. Available: https://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-6-r44. Accessed 04 Feb 2023.
37. Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14(1):1–10. https://doi.org/10.1186/1471-2164-14-293/TABLES/2. Available: https://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-293. Accessed 04 Feb 2023.
38. Zhang Y, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun. 2017;8(1):1–11. https://doi.org/10.1038/ncomms14617. Available: https://www.nature.com/articles/ncomms14617. Accessed 22 Nov 2022.
39. Thrush KL, Higgins-Chen AT, Liu Z, Levine ME. R methylCIPHER: a methylation clock investigational package for hypothesis-driven evaluation & Research. bioRxiv. 2022;2022.07.13.499978. https://doi.org/10.1101/2022.07.13.499978. Available: https://www.biorxiv.org/content/10.1101/2022.07.13.499978v1. Accessed 30 Jan 2023.
40. Pelegi-Siso D, De Prado P, Ronkainen J, Bustamante M, Gonzalez JR. methylclock: a Bioconductor package to estimate DNA methylation age. Bioinformatics. 2021;37(12):1759–1760. https://doi.org/10.1093/BIOINFORMATICS/BTAA825. Available: https://academic.oup.com/bioinformatics/article/37/12/1759/5909987. Accessed 30 Jan 2023.
41. Higgins-Chen AT, et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. Nature Aging. 2022;2(7):644–661. https://doi.org/10.1038/s43587-022-00248-2. Available: https://www.nature.com/articles/s43587-022-00248-2. Accessed 30 Jan 2023.
42. Revelle W. Procedures for personality and psychological research, Northwestern University, Evanston, Illinois, USA. R package published through CRAN, vol. 1.6.12, 2016.
43. Fox J, Weisberg S. An {R} Companion to applied regression, Third Edition. Thousand Oaks CA: Sage. no. September 2012, 2019.
44. Wilkinson L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. Biometrics. 2011;67(2). https://doi.org/10.1111/j.1541-0420.2011.01616.x.
45. Hamner B, Frasco M. Metrics: Evaluation metrics for machine learning. R package version 0.1. 2018.
46. R: The R Project for Statistical Computing. Available: https://www.r-project.org/. Accessed 30 Jan 2023.
47. Houseman EA, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13(1):1–16. https://doi.org/10.1186/1471-2105-13-86/TABLES/6. Available: https://link.springer.com/articles/10.1186/1471-2105-13-86. Accessed 03 May 2023.
48. Purcell S, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75. https://doi.org/10.1086/519795.
49. Browning BL, Zhou Y, Browning SR. A one-penny imputed genome from next-generation reference panels. Am J Hum Genet. 2018;103(3):338–348. https://doi.org/10.1016/j.ajhg.2018.07.015. Available: http://www.cell.com/article/S0002929718302428/fulltext. Accessed 09 May 2023.
50. Zhao W, et al. Education and lifestyle factors are associated with dna methylation clocks in Older African Americans. Int J Environ Res Public Health. 2019;16(17). https://doi.org/10.3390/IJERPH16173141. Available: /pmc/articles/PMC6747433/. Accessed 06 Apr 2023.
51. Marioni RE, et al. Tracking the epigenetic clock across the human life course: a meta-analysis of longitudinal cohort data. J Gerontol - Ser A Biol Sci Med Sci. 2019;74(1):57–61. https://doi.org/10.1093/GERONA/GLY060.
52. Klopack ET, Carroll JE, Cole SW, Seeman TE, Crimmins EM. Lifetime exposure to smoking, epigenetic aging, and

morbidity and mortality in older adults. Clin Epigenetics. 2022;14(1). https://doi.org/10.1186/S13148-022-01286-8. Available: /pmc/articles/PMC9148451/. Accessed 15 Nov 2023.
53. Kresovich JK, et al. Associations of body composition and physical activity level with multiple measures of epigenetic age acceleration. Am J Epidemiol. 2021;190(6):984–993. https://doi.org/10.1093/AJE/KWAA251. Available: https://academic.oup.com/aje/article/190/6/984/5986666. Accessed 28 Apr 2023.
54. Fox FAU, Liu D, Breteler MMB, Aziz NA. Physical activity is associated with slower epigenetic ageing-Findings from the Rhineland study. Aging Cell. 2023;22(6). https://doi.org/10.1111/ACEL.13828. Available: https://pubmed.ncbi.nlm.nih.gov/37036021/. Accessed 15 Nov 2023.
55. Jokai M, et al. DNA methylation clock DNAmFitAge shows regular exercise is associated with slower aging and systemic adaptation. Geroscience. 2023;45(5). https://doi.org/10.1007/S11357-023-00826-1. Available: https://pubmed.ncbi.nlm.nih.gov/37209203/. Accessed 15 Nov 2023.
56. Spólnicka M, et al. Modified aging of elite athletes revealed by analysis of epigenetic age markers. Aging. 2018;10(2):241–252. https://doi.org/10.18632/AGING.101385. Available: https://pubmed.ncbi.nlm.nih.gov/29466246/. Accessed 15 Nov 2023.
57. Pavanello S, Campisi M, Tona F, Dal Lin C, Iliceto S. Exploring epigenetic age in response to intensive relaxing training: a pilot study to slow down biological age. Int J Environ Res Public Health 2019;16(17):3074. https://doi.org/10.3390/IJERPH16173074. Available: https://www.mdpi.com/1660-4601/16/17/3074/htm. Accessed 28 Apr 2023.
58. Harkess KN, Ryan J, Delfabbro PH, Cohen-Woods S. Preliminary indications of the effect of a brief yoga intervention on markers of inflammation and DNA methylation in chronically stressed women. Transl Psychiatry 2016;6(11):e965–e965. https://doi.org/10.1038/tp.2016.234. Available: https://www.nature.com/articles/tp2016234. Accessed 17 May 2023.
59. Kripalani S, Pradhan B, Gilrain KL. The potential positive epigenetic effects of various mind-body therapies (MBTs): a narrative review. J Complement Integr Med. 2022;19(4):827–832. https://doi.org/10.1515/JCIM-2021-0039/MACHINEREADABLECITATION/RIS. Available: https://www.degruyter.com/document/doi/10.1515/jcim-2021-0039/html. Accessed 17 May 2023.
60. Kaliman P, et al. Epigenetic impact of a 1-week intensive multimodal group program for adolescents with multiple adverse childhood experiences. Sci Rep. 2022;12(1):1–16. https://doi.org/10.1038/s41598-022-21246-9. Available: https://www.nature.com/articles/s41598-022-21246-9. Accessed 17 May 2023.
61. Quach A, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging. 2017;9(2):419–446. https://doi.org/10.18632/AGING.101168. Available: https://pubmed.ncbi.nlm.nih.gov/28198702/. Accessed 15 Nov 2023.
62. Chuang YH, Quach A, Absher D, Assimes T, Horvath S, Ritz B. Coffee consumption is associated with DNA methylation levels of human blood. Eur J Hum Genet. 2017;25(5):608. https://doi.org/10.1038/EJHG.2016.175. Available: /pmc/articles/PMC5437893/. Accessed 15 Nov 2023.
63. Ding Q, Xu Y-M, Lau ATY. The epigenetic effects of coffee. Molecules. 2023;28:1770. https://doi.org/10.3390/molecules28041770 Available: https://doi.org/10.3390/molecules28041770. Accessed 15 Nov 2023
64. Wang C, et al. Association of estimated sleep duration and naps with mortality and cardiovascular events: a study of 116 632 people from 21 countries. Eur Heart J. 2019;40(20):1620–1629. https://doi.org/10.1093/EURHEARTJ/EHY695. Available: https://pubmed.ncbi.nlm.nih.gov/30517670/. Accessed 15 Nov 2023.
65. Kusters CDJ, Klopack ET, Crimmins EM, Seeman TE, Cole S, Carroll JE. Short sleep and insomnia are associated with accelerated epigenetic age. Psychosom Med. 2023. https://doi.org/10.1097/PSY.0000000000001243. Available: https://journals.lww.com/psychosomaticmedicine/fulltext/9900/short_sleep_and_insomnia_are_associated_with.155.aspx. Accessed 15 Nov 2023.
66. Kong L, et al. Genetic evidence for causal effects of socioeconomic, lifestyle, and cardiometabolic factors on epigenetic-age acceleration. J Gerontol: Ser A. 2023. https://doi.org/10.1093/GERONA/GLAD078. Available: https://academic.oup.com/biomedgerontology/advance-article/doi/10.1093/gerona/glad078/7069451. Accessed 15 May 2023.
67. Fiorito G, et al. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. Aging. 2019;11(7):2045–2070. https://doi.org/10.18632/AGING.101900. Available: https://www.aging-us.com/article/101900. Accessed 28 Apr 2023.
68. Liu Z, et al. Underlying features of epigenetic aging clocks in vivo and in vitro. Aging Cell. 2020;19(10):e13229. https://doi.org/10.1111/ACEL.13229. Available: https://onlinelibrary.wiley.com/doi/full/10.1111/acel.13229. Accessed 26 Apr 2023.
69. Levine ME, Higgins-Chen A, Thrush K, Minteer C, Niimi P. Clock work: deconstructing the epigenetic clock signals in aging, disease, and reprogramming. bioRxiv. 2022;2022.02.13.480245. https://doi.org/10.1101/2022.02.13.480245. Available: https://www.biorxiv.org/content/10.1101/2022.02.13.480245v1. Accessed 26 Jan 2023.
70. Lu AT, et al. GWAS of epigenetic aging rates in blood reveals a critical role for TERT. 2018;Nat Commun. 9(1). https://doi.org/10.1038/s41467-017-02697-5
71. Gibsonid J, et al. A meta-analysis of genome-wide association studies of epigenetic age acceleration. 2019. https://doi.org/10.1371/journal.pgen.1008104. Available: https://doi.org/10.1371/journal.pgen.1008104. Accessed 09 Jan 2021.
72. McCartney DL, et al. Genome-wide association studies identify 137 genetic loci for DNA methylation biomarkers of aging. Genome Biol. 2021;22(1). https://doi.org/10.1186/s13059-021-02398-9.
73. Farquharson C, Ahmed SF. Inflammation and linear bone growth: the inhibitory role of SOCS2 on GH/IGF-1 signaling. Pediatr Nephrol. 2013;28(4). https://doi.org/10.1007/s00467-012-2271-0.

74. Horvat S, Medrano JF. Lack of Socs2 expression causes the high-growth phenotype in mice. Genomics. 2001;72(2). https://doi.org/10.1006/geno.2000.6441.
75. Zhang T, et al. SOCS2 Inhibits mitochondrial fatty acid oxidation via suppressing LepR/JAK2/AMPK signaling pathway in mouse adipocytes. Oxid Med Cell Longev. 2020;2020. https://doi.org/10.1155/2020/3742542.
76. Sebastiani P, et al. Genetic signatures of exceptional longevity in humans. PLoS One. 2012;7(1). https://doi.org/10.1371/journal.pone.0029848.