

EpiAge: a next-generation sequencing-based *ELOVL2* epigenetic clock for biological age assessment in saliva and blood across health and disease

David Cheishvili^{1,2,3,*}, Sonia Do Carmo^{4,*}, Filippo Caraci^{6,7}, Margherita Grasso⁷,
A Claudio Cuello^{4,5}, Moshe Szyf^{1,2}

¹EpiMedTech Global, Singapore 409051, Singapore

²HKG Epitherapeutics Ltd., Hong Kong SAR, China

³Gerald Bronfman Department of Oncology, McGill University, Montreal H4A 3T2, Canada

⁴Department of Pharmacology & Therapeutics, McGill University, Montreal H3G 1Y6, Canada

⁵Visiting Professor, Department of Pharmacology, Oxford University, Oxford OX13QT, UK

⁶Department of Drug and Health Sciences, University of Catania, Catania 95125, Italy

⁷Neuropharmacology and Translational Neurosciences Research Unit, Oasi Research Institute-IRCCS, Troina 94018, Italy

*Equal contribution

Correspondence to: Moshe Szyf; email: moshe.szyf@epimedtech.com, <https://orcid.org/0000-0003-3881-8214>

Keywords: epigenetic clock, *elovl2*, next-generation sequencing, EpiAge, Alzheimer's disease

Received: May 8, 2024

Accepted: January 6, 2025

Published: January 22, 2025

Correction: This article has been corrected. Please see Aging 2025: <https://doi.org/10.18632/aging.206263>

Copyright: © 2025 Cheishvili et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

This study introduces EpiAgePublic, a new method to estimate biological age using only three specific sites on the gene *ELOVL2*, known for its connection to aging. Unlike traditional methods that require complex and extensive data, our model uses a simpler approach that is well-suited for next-generation sequencing technology, which is a more advanced method of analyzing DNA methylation. This new model overcomes some of the common challenges found in older methods, such as errors due to sample quality and processing variations.

We tested EpiAgePublic with a large and varied group of over 4,600 people to ensure its accuracy. It performed on par with, and sometimes better than, more complicated models that use much more data for age estimation. We examined its effectiveness in understanding how factors like HIV infection and stress affect aging, confirming its usefulness in real-world clinical settings.

Our results prove that our simple yet effective model, EpiAgePublic, can capture the subtle signs of aging with high accuracy. We also used this model in a study involving patients with Alzheimer's Disease, demonstrating the practical benefits of next-generation sequencing in making precise age-related assessments.

This study lays the groundwork for future research on aging mechanisms and assessing how different interventions might impact the aging process using this clock.

INTRODUCTION

The broad interest in *ELOVL2* (Elongation of Very Long Chain Fatty Acids Protein 2) in the field of aging research is highlighted by more than a hundred scholarly articles

discussing its link with epigenetic aging. This extensive body of research confirms its critical role and widespread recognition within the scientific community. Consistently heralded as a potent epigenetic biomarker, *ELOVL2* has proven indispensable in elucidating the biological

processes associated with aging [1, 2]. *ELOVL2* has been integrated into epigenetic clocks to accurately determine age-related differences across various disease groups, highlighting its utility in age prediction [3]. Research indicates that *ELOVL2* affects the aging process through its role in regulating lipid metabolism, with its epigenetic alterations closely linked to age prediction capabilities (Li et al., 2022). Additionally, the CpG sites within *ELOVL2*'s regulatory regions play a crucial role in age prediction models, underscoring the gene's central importance in epigenetic clocks [4]. Since it is well known that there is a close correlation in methylation between juxtaposed CpG positions [5–8], we reasoned that including additional CpGs in the region could increase the accuracy of the analysis. Therefore, we designed the primers to cover not only the three CpGs discussed (cg16867657, cg21572722, and cg24724428) but also to capture ten additional CpGs within the same region.

Further research has shown that DNA methylation patterns of *ELOVL2* are associated with age-related macular degeneration, establishing it as a reliable biomarker for aging in ocular tissues (Saptarshi et al., 2021). In addition, *ELOVL2* methylation has been utilized in forensic science for age estimation, which underscores its critical value in age prediction applications (Brenna & Kothapalli, 2021).

In summary, *ELOVL2* is a fundamental gene in the development of epigenetic clocks and age prediction models. Its methylation patterns serve as a robust indicator of biological age, demonstrating significant utility across various research and clinical contexts.

Nevertheless, traditional epigenetic clocks primarily utilize Illumina Infinium BeadChips, which analyze hundreds of thousands of CpG sites. However, this approach is prone to technical variances from sample preparation, probe hybridization, chemistry, and batch effects, which often compromise data reliability [9]. Despite efforts to mitigate these issues through normalization and batch correction, significant challenges persist, particularly affecting the clocks' utility in both basic and translational research, including short-term longitudinal studies like clinical trials.

Next-generation sequencing (NGS) significantly enhances the precision and scope of epigenetic studies by addressing the limitations of traditional array technologies, such as technical noise and reliance on predefined sequences. With its high-throughput capability, base-resolution accuracy, and broader genomic coverage, NGS not only allows for detailed examination of methylation patterns across diverse genomic contexts but also facilitates the discovery of

novel methylation sites. This comprehensive approach allows for more accurate identification of differentially methylated regions, providing a robust alternative that surpasses traditional methods in both detail and scalability [10]. The accuracy of DNA methylation determination of bisulfite-converted DNA by next-generation sequencing is dependent on the number of reads to achieve statistical power because of the heterogeneity of methylation profiles even in the same tissue. Therefore, reducing the number of regions that are required to be sequenced for a given biomarker would increase the depth of sequencing and, thus, the power for accurate determination of its DNA methylation, especially when samples are multiplexed, to increase the cost-effectiveness of the biomarker. We, therefore, examined in this study whether we need the hundreds or thousands of regions that are used in all current epigenetic clocks to develop an accurate and cost-effective epigenetic clock.

Most current clocks use blood as a biological sample. Saliva has become a valuable resource for epigenetic age estimation due to its non-invasive collection, ease of handling, and rich DNA content from both epithelial and white blood cells. This mixed cellular composition captures systemic biological signals, reflecting both oral health and broader immune responses, making it suitable for large-scale studies and clinical settings where less invasive methods enhance participant compliance. Additionally, saliva mirrors the methylome of blood and other tissues, which supports its use in epigenetic studies exploring the effects of environmental exposures, lifestyle, and disease states on aging [11]. Research shows saliva's relevance in studying various health conditions, such as its association with Parkinson's disease and certain carcinomas through differential methylation patterns [12, 13]. Additionally, saliva has been utilized to evaluate epigenetic age acceleration, with adjustments made for cell type proportions in the samples [14]. It has also been instrumental in developmental studies linking birth weight with DNA methylation [15] and evaluating childhood BMI and social disparities through epigenetic markers [16].

This utility highlights saliva's potential to advance our understanding of biological aging and age-related diseases. In this study, we assess the performance of various biological clocks in saliva versus blood.

This study introduces EpiAgePublic, a new epigenetic aging model that leverages just three strategically selected CpG sites within the *ELOVL2* gene, a key marker strongly associated with aging. Inspired by minimal-marker models like the “Epigenetic age-predictor for mice based on three CpG sites [17], EpiAgePublic is optimized for next-generation sequencing technologies. Unlike

traditional array-based clocks that rely on hundreds to thousands of CpG sites, this model demonstrates the effectiveness of using fewer markers to accurately predict age, thus addressing common limitations of older methods. This method not only mitigates issues related to DNA quality and batch effects but also enhances the model's specificity and clinical applicability.

Validated against a comprehensive dataset comprising 4,625 individuals, EpiAgePublic has shown predictive accuracies on par with, or even exceeding, those of well-established epigenetic clocks such as DNAmAge (Horvath's Clock) and DNAmPhenoAge. The model's performance was rigorously evaluated in various clinical contexts, including its correlation with HIV infection and stress levels, underscoring its utility alongside more traditional, complex clocks.

Our results affirm that EpiAgePublic, despite its simplicity in focusing on only three CpG sites, effectively captures the biological intricacies of aging. This demonstrates its viability as a powerful tool for aging research and clinical application. We also determined that the clock performs well on saliva samples. We next developed a targeted amplification-next-generation sequencing assay that captures 13 CpG sites around the *ELOVL2* gene and used a proprietary model on DNA methylation data from the *ELOVL2* region. The model was further employed in a clinical study on Alzheimer's Disease, where next-generation sequencing was utilized to assess epigenetic age. This application not only highlighted EpiAgePublic's precision but also illustrated the significant advantages of next-generation sequencing in enhancing the accuracy of epigenetic age evaluations. By using saliva as the biological sample, we increase the feasibility of wide usage of this clock.

MATERIALS AND METHODS

Estimation of DNA methylation age

The Horvath's DNAmAge, Hannum's DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkin-BloodClock, and DNA GrimAge (both versions 1 and 2) methylation ages were determined using the online DNA Methylation Age Calculator, provided by the Clock Foundation: <https://dnamage.clockfoundation.org/>.

EpiAge for next-generation sequencing data from individuals in the Alzheimer's Disease study was obtained through a commercial service provided by HKG epiTherapeutics. This service utilizes a proprietary model based on DNA methylation data within the same *ELOVL2* region discussed in this paper,

specifically designed for next-generation sequencing applications. The analysis was performed in triplicate.

Development and application of the EpiAgePublic model to investigate biological aging dynamics

In this study, we developed the EpiAgePublic model to explore the impact of various clinical outcomes on biological aging. Linear regression was employed to assign weights to three targeted CpG sites: cg16867657, cg21572722, and cg24724428. We selected these three CpG sites from a set of 13 CpGs used in our previous study [18] due to their presence on standard DNA methylation microarray platforms. This enabled us to validate the model across a wide range of publicly available datasets derived from these platforms. The model calculates beta values using the formula:

$$(\beta_{cg16867657} \times 122.70 + \beta_{cg21572722} \times 24.45 + \beta_{cg24724428} \times (-30.44)) - 42.91.$$

Model training

The model was trained using a comprehensive DNA methylation dataset aggregated from four public databases: GSE55763, GSE157131, GSE40279, and GSE30870, as detailed in Table 1. This training dataset, derived from Illumina 450K and Epic array platforms, encompasses a demographic range of individuals aged 0 to 103 years, including Caucasian-European, Hispanic Mexican, and African American ethnic groups. The sex distribution was balanced, comprising 2506 males, 2079 females, and 40 individuals with unspecified gender.

The selection of this diverse dataset was critical for developing an inclusive and representative EpiAgePublic model. The model achieved an R-squared value of 0.7512 in the training cohort, indicating its robust ability to accurately reflect the relationship between DNA methylation patterns and biological age.

Model validation

After development, the EpiAgePublic model was validated on independent datasets, which included several additional public cohorts listed in Table 1. These validation datasets were not part of the training cohort and were used to test the generalizability and robustness of the model across different populations and clinical contexts. The validation results showed that the model maintained high accuracy and robustness, supporting its effectiveness as a reliable indicator of biological aging across diverse populations.

Table 1. Comprehensive dataset summary for epigenetic age model development and clinical studies.

GSE ID	Sample tissue type	Total samples	Age range	Cohort description	Methylation platform	Dataset application	Purpose
GSE55763	Blood	2711	23.7 - 75	General Population	450K	epiAgePublic model	Training
GSE157131	Blood	1218	26.41 - 94.74	General Population	450K/Epic	epiAgePublic model	Training
GSE40279	Blood	656	31 - 68	General Population	450K	epiAgePublic model	Training
GSE30870	Blood	40	30 - 48	General Population	450K	epiAgePublic model	Training
GSE78874	Saliva	258	36-88	General Population	450K	epiAgePublic validation	Validation
GSE150643	Saliva	240	9.19-15.85	General Population	450K	epiAgePublic validation	Validation
GSE92767	Saliva	54	18-73	General Population	450K	epiAgePublic validation	Validation
GSE99029	Saliva	57	21-91	General Population	450K	epiAgePublic validation	Validation
GSE67751	Blood	92	24 - 68	Healthy and HIV Patients	450K	HIV	Validation
GSE117859	Blood	608	25 - 75	HIV Patients	450K	HIV	Validation
GSE53840	Blood	120	31 - 68	HIV Patients	450K	HIV	Validation
GSE185391	Blood	86	30 - 48	HIV Patients	450K	HIV treatment	Validation
GSE167202	Blood	525	17-96	Healthy, COVID-19 and other infections	Epic	COVID-19	Validation
GSE168739	Blood	407	19-61	COVID-19 patients	Epic	COVID-19	Validation
GSE72680	Blood	422	18-77	Trauma and Psychiatric Symptoms in African Americans	450K	Stress	Validation
GSE128235	Blood	533	18-87	Depression	450K	Controls, Depression	Validation
GSE125105	Blood	847	17-87	Depression	450K	Controls, Depression	Validation
GSE144858	Blood	300	52-90	Alzheimer's	450K	Controls, MCI, Alzheimer's	Validation

Epigenetic age acceleration (EAA)

Epigenetic Age Acceleration (EAA) measures biological aging dynamics. It is computed as the difference between DNA methylation age and chronological age. A positive EAA value indicates that an individual's biological age surpasses their chronological age, suggesting accelerated aging. This metric is used to compare aging rates across different clinical conditions by analyzing the differences in EAA between a specific condition and a control group.

Data source

The DNA methylation data used in this study were derived from publicly available datasets and newly collected original data. The Down syndrome and Alzheimer's Disease cohorts represent the original data collected specifically for this study. Details regarding the sample collection, processing, and DNA extraction for these cohorts are provided below.

Demographic characteristics of the study sample cohorts

Down syndrome study

The study protocol was approved by the Institutional Review Board (IRB) of the Oasi Research Institute—IRCCS, Troina, Italy (Protocol Number: 2016/1.0/122/CE-IRCCS-OASI). Participants or their legal representatives gave written informed consent, as approved by the Ethics Committee of the Oasi Research Institute-IRCCS. This study included 22 control participants (9 males and 13 females) and 22 individuals

with Down syndrome (9 males and 13 females). The mean age for the control group was 40.95 years (SD = 9.276), and for the Down syndrome group, it was 40.65 years (SD = 8.969). An unpaired parametric t-test showed no significant age difference between the two groups (p-value = 0.9146).

Alzheimer's disease study

This research was approved by the IRB of Oasi Research Institute—IRCCS, Troina, Italy, under Protocol Number: 2019/03/18/CE-IRCCS-OASI/18. Participants or their legal representatives gave written informed consent, as approved by the Ethics Committee of the Oasi Research Institute-IRCCS. The study involved 55 participants, with a demographic composition of 10 males and 17 females in the control group and 7 males and 21 females in the Alzheimer's disease (AD) group. The mean age was 72.25 years (SD = 9.26) for the controls and 73.01 years (SD = 8.91) for the AD group. Due to technical difficulties, methylation analysis was performed on 54 participants, excluding one from the original cohort. An unpaired parametric t-test revealed that the age difference between the two groups was not statistically significant (p-value = 0.0585), although this value approaches the conventional threshold for significance.

Plasma sample processing and white blood cell isolation

Plasma samples were collected according to standard procedures. Briefly, fasting venous blood was collected in

lavender EDTA-K2 BD vacutainer tubes and centrifuged at 1900 rpm for 10 minutes for the separation of the plasma component from the cellular constituents. After a second centrifugation (3900 rpm for 10 min) to purify from biological debris, plasma samples were separated in aliquots and stored at -80°C until use. The blood was used for peripheral blood mononuclear cells (PBMCs) isolation by using the Lympholyte®-H density gradient separation medium (Cedarlane, Burlington, NC, USA) according to the manufacturer's instructions, with slight modifications. Briefly, the blood was diluted with an equal volume of 1X PBS, mixed gently and added to one part of Lympholyte®-H. After a centrifugation step (400 rcf for 30 min), a well-defined lymphocyte layer appeared at the interface, which was removed and transferred by using a serological pipette to a new sterile centrifuge tube. After a washing step by using 5 ml of 1X Red Blood Cell Lysis Buffer (Abcam, ab204733) for optimal lysis of erythrocytes in a single-cell suspension and after two additional washing steps (1X sterile PBS, 1600 rpm for 10 min), the isolated PBMCs fraction was counted by using LUNA-II Automated Cell Counter (Logos Biosystems).

DNA extraction from isolated lymphocytes

DNA was isolated from lymphocytes using the AllPrep DNA/RNA Mini Kit (Qiagen, Cat: 80204) according to the manufacturer's protocol. To enhance DNA concentration, it was eluted twice with 50 μl of Buffer EB following a 1-minute incubation at room temperature. The amount and purity of the extracted DNA were assessed using spectrophotometry and by measuring the A260/280 ratio, which ranged between 1.9 and 2.1, indicating high-quality DNA suitable for subsequent analyses.

Targeted DNA methylation sequencing with illumina NGS

DNA was bisulfite-converted using the EZ-96 DNA Methylation MagPrep kit (D5041, Zymo Research), followed by two stages of polymerase chain reaction (PCR). In the first round of PCR, primers, including an anchoring sequence targeted at a specific region of the *ELOVL2* gene, were used (Bio-Rad C1000 Touch Thermal Cycler, Bio-Rad Laboratories, CA, USA). Primer sequences are available upon request. Five microliters from the first PCR reaction were then used in a second round of PCR to amplify and barcode the samples using indexing primers. The PCR products were pooled, purified twice using AMPure XP Beads (Beckman Coulter Life Sciences, CA, USA), and quantified via real-time PCR (NEBNext® Library Quant Kit for Illumina, New England Biolabs, MA, USA). The barcoded libraries were sequenced on the Illumina platform using a MiSeq

Reagent Nano Kit V2 with a 250×2 paired-end sequencing protocol (Illumina, CA, USA).

For data processing, raw paired-end reads were trimmed of sequencing adapters and low-quality sequences using Trim-galore (parameters: trim_galore -illumina -paired -fastqc; available at <https://zenodo.org/record/5127899#.Y7RxfOzMJqs>). Cleaned data were aligned to the *ELOVL2* reference genome using Bismark [19], and reads deduplicated using UMIs in the forward primers to minimize PCR amplification bias (deduplicate_bismark --paired --barcode -bam). Methylation levels at each CpG site were determined using the Bismark methylation extractor (bismark_methylation_extractor -p --bedGraph --counts --scaffolds --no_overlap), with a minimum threshold of 100 reads per gene for inclusion in the analysis.

DNA methylation data processing and epigenetic age calculation

IDAT files were processed using the champ.load function from the ChAMP R package [20, 21]. Initial beta values were computed directly from the loaded data to represent methylation levels before normalization.

Following initial data processing, we applied BMIQ (Beta Mixture Quantile dilation) normalization to adjust the beta values. This normalization technique specifically corrects for the technical variation between type I and type II probes by aligning the distribution of type II probe beta values to match those of type I probes. This step ensures that the methylation data are more statistically reliable and comparable across different probe types [22]. Epigenetic age (EpiAge) was calculated for each individual both before and after normalization using the EpiAgePublic algorithm.

Statistical analysis

For multilinear regression analysis, we utilized Python's statsmodels library to fit an Ordinary Least Squares (OLS) model. This model was employed to examine the relationship between epigenetic age and several predictors, including cell composition and sex. Our analysis encompassed data preparation, model fitting, and the extraction of key statistics, such as coefficients and p-values.

In addition to multilinear regression, our study also involved other statistical techniques to address specific research questions:

T-tests and Mann-Whitney U tests were used to compare epigenetic ages across different conditions and

cohorts, ensuring that we appropriately addressed the distribution characteristics of the data.

Receiver Operating Characteristic (ROC) analysis was conducted to assess the discriminatory power of various epigenetic clocks in distinguishing between different clinical conditions.

ANOVA was utilized to analyze the impact of clinical interventions on epigenetic age across different treatment groups. Statistical tests applied in each figure are described in the figure legend.

RESULTS

Development of the EpiAgePublic model across blood and saliva samples

The *ELOVL2* gene has been consistently linked to chronological age across various studies [2]. Analysis of 20 public datasets derived from whole blood using the Illumina BeadChip technology identified *ELOVL2* and *FHL2* as showing the highest correlation between age and DNA methylation [23].

Our research focused on three CpG sites within the *ELOVL2* gene: cg16867657, cg21572722, and cg24724428. Notably, cg16867657 has been recognized as a key age-associated marker with a robust pattern of age-related methylation changes [24]. This was supported by a study examining 421 individuals, which found cg16867657 showing a strong positive correlation with age among 137993 sites ($R = 0.957$, $P\text{-value} = 1.20\text{e-}228$) [25]. The proximity of cg21572722 and cg24724428 to cg16867657, coupled with their known associations with aging (Bell et al., 2012; Christiansen et al., 2016; Hao et al., 2021; Horvath & Raj, 2018; Li et al., 2022; Marioni et al., 2015), justified their inclusion in our EpiAgePublic model. The development of the EpiAgePublic model utilized data from four specific datasets: GSE55763, GSE157131, GSE40279, and GSE30870 (Table 1). The detailed methodology and analysis for model development are outlined in the Methods section of our study.

Comparative analysis of epiAgePublic with established epigenetic clocks

We conducted comparative analyses to evaluate the EpiAgePublic model's performance relative to established epigenetic clocks. This analysis utilized the same datasets originally employed for the development of the EpiAgePublic model: GSE55763, GSE157131, GSE40279, and GSE30870, all derived from blood samples. This included comparisons of cg16867657, cg21572722, and cg24724428, as well as the

EpiAgePublic model, against established epigenetic age models such as DNAmAge (Horvath's Clock) [26], DNAmAgeHannum [27], DNAmPhenoAge [28], DNAmAgeSkinBloodClock [29], as well as DNA methylation GrimAge versions 1 (DNAGrimAge v1) [30] and 2 (DNAGrimAge v2) [31] (Figure 1A).

The analysis revealed high correlation coefficients for all models. The EpiAgePublic model, which integrates the three CpG sites, exhibited a correlation of 0.87, surpassing individual CpGs (cg16867657, cg21572722, and cg24724428) and closely following DNAmAgeSkinBloodClock, which scored the highest with 0.93. Other clocks displayed slightly lower correlations, with DNAmAgeHannum showing the lowest at 0.65 (Figure 1A).

Assessing the impact of cell composition and sex on epigenetic age predictions

Our next objective was to determine whether epigenetic age predictions could be affected by blood cell composition and the individual's sex. To precisely assess the potential confounding effects and correct for variations related to blood cell composition (CD8T cells, CD4T cells, NK cells, B cells, monocytes (Mono), and granulocytes (Gran)), which are inferred from DNA methylation patterns indicative of these cells' presence [32], we executed multilinear regression analyses adding these confounders as covariates. This detailed examination focused individually on the cg16867657, cg21572722, and cg24724428 CpG sites within the *ELOVL2* gene, as well as collectively within the context of the EpiAgePublic model, comparing these values against DNAmAge (Horvath's Clock), DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. These analyses validated the significant and consistent predictive value of chronological age across all individual CpG sites and the EpiAge, even when these confounders are included in the regression (Supplementary Table 1).

Intriguingly, a distinct feature of EpiAgePublic emerged during our analysis: both the individual CpG site assessments and the aggregated EpiAgePublic model demonstrated no significant correlation with sex, exhibiting a $P\text{-value}$ of 0.182. This finding contrasts with other epigenetic clocks, where sex showed a significant influence with a $P\text{-value}$ of less than 0.05. Specifically, the regression analysis for EpiAgePublic indicated a sex coefficient of 0.388 (std err: 0.291), which was not statistically significant ($P > |t|$: 0.182), diverging from the trends observed in other epigenetic clocks where the $P\text{-value}$ for sex was effectively 0, signifying a highly statistically

significant correlation. This distinctive attribute of EpiAgePublic underscores its robustness by highlighting its ability to provide age predictions that are unbiased by sex differences.

Validating epiAgePublic on saliva samples

Following the development of the EpiAgePublic model, we aimed to validate its applicability to saliva, a noninvasive biological matrix different from blood. The rationale was based on the distinct cellular compositions of saliva, which differ significantly

between children and adults in terms of epithelial and immune cells.

Using datasets from GSE78874, GSE150643, GSE92767, and GSE99029, we processed samples from 609 healthy individuals ranging from 9 to 91 years, including 310 males and 294 females from diverse ethnic backgrounds such as Hispanic, Caucasian, African, and Asian using Illumina 450K and Epic array platforms (Table 1).

The comparative analysis of EpiAgePublic and other clocks demonstrated EpiAge's strong predictive

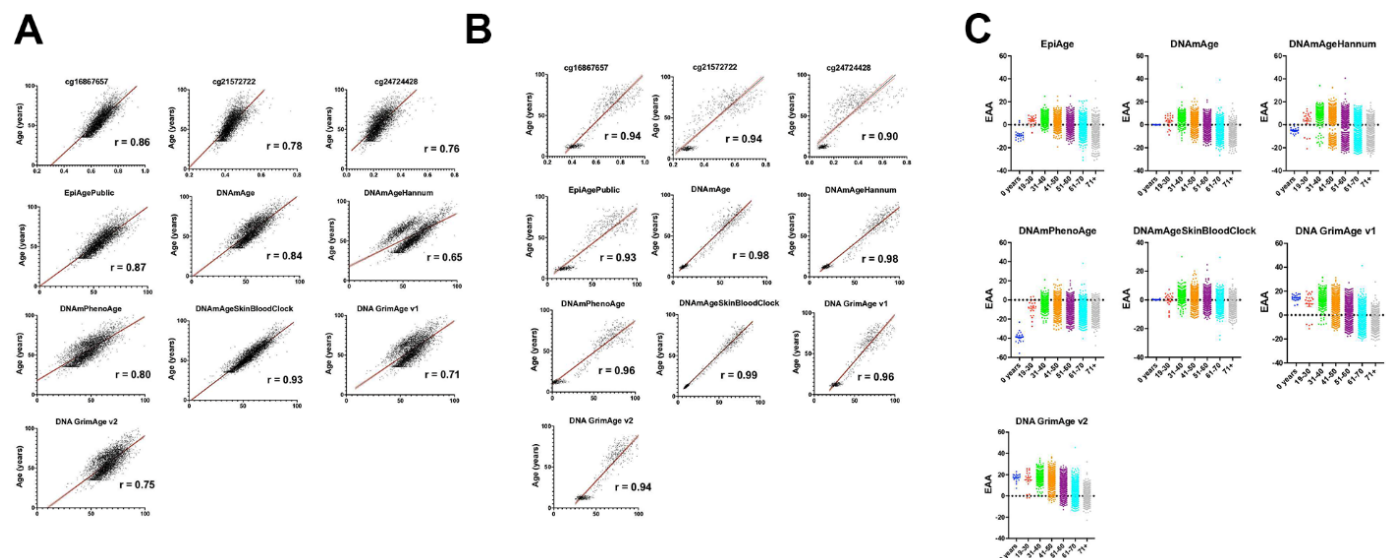


Figure 1. Comprehensive analysis of epigenetic aging across diverse datasets and demographics. (A) This figure illustrates the correlation between chronological age (y-axis) and measures of epigenetic age including EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2 (x-axis), as well as individual CpG sites cg16867657, cg21572722, and cg24724428. Data were aggregated from the datasets GSE55763, GSE157131, GSE40279, and GSE30870 (refer to Table 1), encompassing 4625 individuals with ages ranging from 0 to 103 years. The cohort exhibits a rich demographic diversity, including Caucasian-European, Hispanic Mexican, and African American ethnicities, comprising 2506 males, 2079 females, and 40 individuals with unspecified gender. The correlations were assessed using the Pearson r correlation coefficient, denoted by 'R' on each plot, highlighting the linear relationship between chronological and epigenetic age across the datasets. All plots achieved a significant p -value of < 0.0001 , indicating a strong and statistically significant correlation. Visualization includes a solid black line representing the mean correlation and flanking red lines depicting the 95% confidence interval, illustrating the precision of the correlation estimates and the degree of agreement between chronological and epigenetic age measures across the studied population. (B) This figure presents the correlation between chronological age (y-axis) and various measures of epigenetic age (x-axis), including EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2, alongside individual CpG sites cg16867657, cg21572722, and cg24724428. The data are derived from saliva samples collected from 609 healthy individuals aged 9 to 91 years, detailed in datasets GSE78874, GSE150643, GSE92767, and GSE99029 (referenced in Table 1). The study population includes 310 males and 294 females from diverse ethnic backgrounds—Hispanic, Caucasian, African, and Asian. Correlations are quantified using the Pearson r coefficient, denoted by 'R' on each plot, signifying the linear relationship between the two age measures. All correlations are marked by a significance level of $p < 0.0001$. Visuals include a solid black line indicating the average correlation and red lines showing the 95% confidence interval, emphasizing the reliability and consistency of epigenetic age measures with chronological age across the cohorts. (C) This figure presents scatter plots comparing Epigenetic Age Acceleration (EAA) across various age groups. Each dot represents an individual's EAA value, plotted against their chronological age group. The age groups are categorized as 0 years, 19-30, 31-40, 41-50, 51-60, 61-70, and 71+ years. The vertical axis indicates the EAA, while the horizontal axis delineates the age groups. A horizontal line at zero on the plot marks the threshold between age acceleration and deceleration; points above this line indicate epigenetic age acceleration, while points below indicate deceleration. This visualization highlights trends and patterns in EAA across the lifespan, offering insights into how biological aging progresses relative to chronological aging across different stages of life.

relationship with chronological age, achieving a correlation coefficient of 0.93. DNAmAge-SkinBloodClock exhibited the highest correlation at 0.99 (Figure 1B). These findings validate the utility of EpiAgePublic for non-invasive biological age estimation across different tissues, including saliva, which has obvious advantages as far as accessibility and compliance over blood.

Age-related epigenetic deceleration across the human lifespan

Our comprehensive dataset encompassed a diverse range of 4,625 individuals, spanning the entire human lifespan from birth to 103 years. The study of Epigenetic Age Acceleration (EAA) traditionally focuses on its clinical significance, often in relation to the acceleration or deceleration of aging and its associated health impacts. Considering that age acceleration may be influenced by various factors such as lifestyle, stress, and socioeconomic conditions, which vary across different age groups, we categorized the dataset into specific age brackets: Newborns, Young Adults (19-30 years), Thirties (31-40 years), Forties (41-50 years), Fifties (51-60 years), Sixties (61-70 years), and Seventies and Beyond (Figure 1C).

To analyze aging acceleration/deceleration patterns within each category, we examined all the epigenetic clocks, consistently observing a trend of epigenetic age deceleration with increasing chronological age among the majority of participants (Figure 1C). Further exploring this phenomenon, we calculated the Pearson r correlation between chronological age and EAA for each clock, assessing the strength and direction of their relationships.

The analysis revealed negative correlations across all clocks, indicating that the rate of epigenetic aging decreases as chronological age increases. Specifically, DNA GrimAge v1 and DNA GrimAge v2 displayed the most significant negative correlations with chronological age ($r=-0.6389$ and $r=-0.6697$, respectively), with tight 95% confidence intervals of -0.6556 to -0.6215 and -0.6853 to -0.6535, respectively, underscoring a strong inverse relationship. Other clocks also showed significant negative correlations, as follows: DNAmAge $r=-0.5943$ (CI: -0.6126 to -0.5753), EpiAgePublic $r=-0.4988$ (CI: -0.5201 to -0.4768), DNAmAgeHannum $r=-0.4424$ (CI: -0.4653 to -0.4189), DNAmAgeSkinBloodClock $r=-0.4420$ (CI: -0.4649 to -0.4185), and DNAmPhenoAge $r=-0.3220$ (CI: -0.3476 to -0.2959). These highly significant Pearson r correlations ($P<0.0001$ for all clocks) confirm a consistent inverse relationship between chronological age and EAA, reinforced by a substantial sample size ($N=4625$ for each clock).

These findings not only corroborate the anticipated pattern of biological aging but also highlight the effectiveness of EpiAgePublic, based on the single gene *ELOVL2*, compared to other clocks that incorporate hundreds or even thousands of CpG sites, as biomarkers for the aging process. The observed age-related trends across the lifespan provide insights into the complex interplay between biological and chronological aging.

Comparing EpiAgePublic to other epigenetic clocks across cohorts in the context of HIV-related accelerated aging

Early research has indicated a link between HIV and accelerated epigenetic aging, suggesting that HIV-positive individuals may exhibit an advanced biological age [33]. Inspired by findings that demonstrated accelerated aging due to CMV infection using the EpiAge model based on *Elovl2* regions in healthy older adults [18], we aimed to explore whether similar patterns could be observed with HIV. We examined the impact of HIV on epigenetic aging using four distinct cohorts: GSE53840, GSE67751, GSE117859, and GSE185391. Our objective was to assess whether the EpiAge model shows accelerated aging in HIV-positive individuals and to compare its performance with other established epigenetic clocks.

We compared EpiAge alongside other epigenetic clocks, including DNAmAge (Horvath's Clock), DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. Our analysis employed an unpaired parametric t -test, which passed the normality test, as most data points are normally distributed. Specifically, we compared HIV-positive individuals from the cohorts GSE53840, GSE67751, GSE117859, and GSE185391 against HIV-negative controls from GSE67751. Within the GSE185391 cohort, our focus was on one-time point w0, marking the week immediately preceding the commencement of HIV treatment.

Our comprehensive analysis across a spectrum of epigenetic clocks revealed a significant impact of HIV infection on epigenetic aging. Each clock offers a distinctive view of the aging acceleration that accompanies HIV. Notably, EpiAge displayed pronounced acceleration in epigenetic age among HIV-positive individuals compared to HIV-negative controls, with an average age advancement of 12.04 years ($P<0.0001$), strongly suggesting a link between HIV and expedited epigenetic aging. Similarly, DNAmAge (Horvath's Clock) and DNAmAgeHannum showed significant age acceleration in HIV-positive subjects, with increases of 6.612 years ($P<0.0001$) and 6.230 years ($P<0.0001$), respectively. The DNAmPhenoAge

clock indicated a significant acceleration of 12.79 years ($P < 0.0001$).

Further analysis highlighted that the DNAmAgeSkinBloodClock and DNAGrimAge v1 also demonstrated significant aging acceleration, with observed differences of 7.664 years ($P < 0.0001$) and an exceptional 19.97 years ($P < 0.0001$), respectively. These findings are consistent with the hypothesis that HIV infection accelerates biological aging. The consistency of epigenetic clocks, including EpiAge, in revealing accelerated aging in HIV-positive individuals further supports this hypothesis (Figure 2A).

Assessing the discriminatory power of EpiAge and other epigenetic clocks for HIV status using ROC analysis

To further assess the discriminatory power of EpiAge and other epigenetic clocks in distinguishing between

HIV-negative and HIV-positive individuals, we conducted a Receiver Operating Characteristic (ROC) analysis. This method provides a nuanced understanding of each clock's ability to accurately classify individuals based on their HIV status beyond the significance revealed by t-tests.

In our ROC analysis, EpiAge demonstrated an impressive Area Under the Curve (AUC) of 0.9109, with a sensitivity of 56.37% and specificity of 100% ($P < 0.0001$), indicating appreciable discriminatory power. This suggests that EpiAge is highly capable of differentiating between the two groups based on epigenetic aging patterns.

By comparison, DNAmAge showed an AUC of 0.7835, with a sensitivity of 14.57%, but maintained the same specificity of 100% ($P < 0.0001$), indicating moderate discriminatory ability. DNAmAgeHannum improved upon DNAmAge's performance, achieving an AUC of

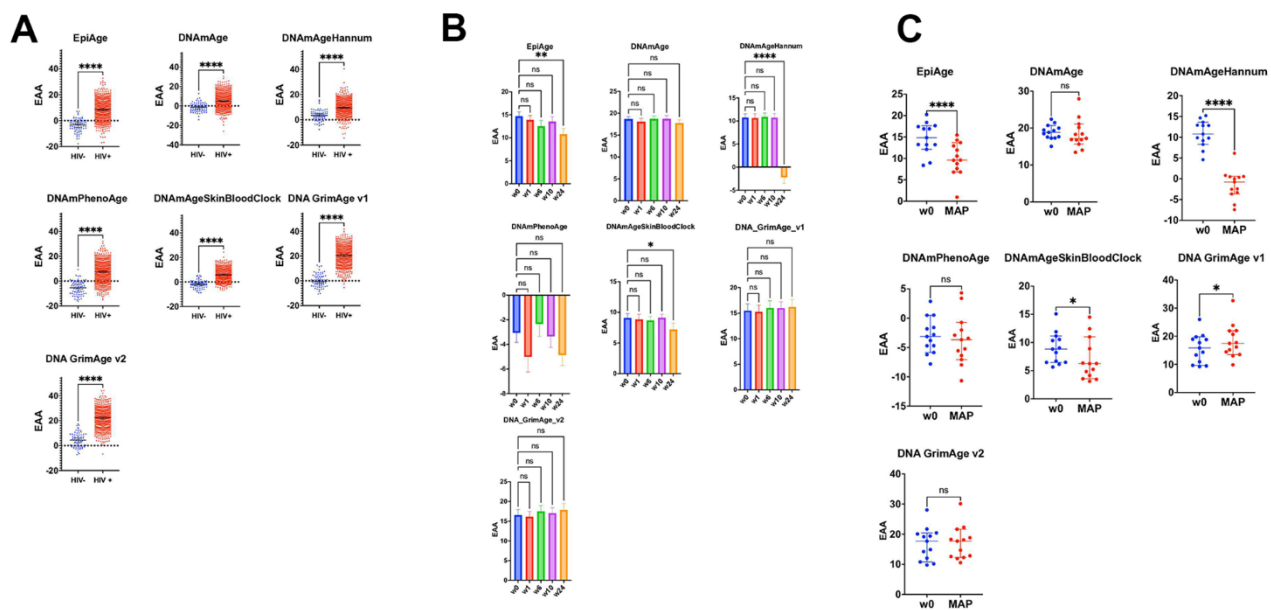


Figure 2. Impact of HIV on epigenetic age acceleration across multiple cohorts and treatment Phases. (A) Scatter plots comparing epigenetic age acceleration (EAA) across multiple epigenetic clocks: EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2, between HIV-negative individuals from the GSE67751 cohort ($n=69$) and HIV-positive individuals from cohorts GSE67751 ($n=23$), GSE117859 ($n=609$), GSE53840 ($n=120$), and GSE185391 ($n=86$). The plots display median EAA values, with error bars denoting 95% Confidence Intervals (CI). The significance of differences in EAA was tested using unpaired, two-tailed t-tests. A horizontal line at zero on the plot marks the threshold between age acceleration and deceleration; points above this line indicate epigenetic age acceleration, while points below indicate deceleration. (B) These scatter plots track changes in EAA from baseline (week 0) through weeks 1, 6, 10, and 24 of a combined HIV treatment strategy, which includes therapeutic vaccination and a latency-reversing agent using multiple epigenetic clocks. We employed repeated measures ANOVA for the analysis. To address potential deviations from the assumption of sphericity, we applied the Geisser-Greenhouse correction as determined by Mauchly's test. This correction ensures accurate and reliable results in repeated measures analysis, particularly when the equality of variances across the differences between all pairs of groups is not met. (C) These scatter plots focus on comparing epigenetic age acceleration (EAA) from baseline (week 0) to the monitored antiretroviral pause (MAP) phase of the BCN02 clinical trial, using various epigenetic clocks: EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. Differences were assessed using paired two-tailed t-tests, chosen for their appropriateness given the normal distribution of data.

0.8275 with a sensitivity of 14.44% and maintaining 100% specificity ($P < 0.0001$).

DNAmPhenoAge demonstrated strong performance with an AUC of 0.9041, sensitivity of 40.40%, and specificity of 100% ($P < 0.0001$), nearly matching EpiAge's discriminatory capacity. DNAmAgeSkinBloodClock exhibited a slightly higher AUC of 0.9168, with a sensitivity of 54.30% and perfect specificity ($P < 0.0001$).

Remarkably, DNAGrimAge v1 displayed the highest AUC among the clocks at 0.9795, with an impressive sensitivity of 82.39% and specificity of 100% ($P < 0.0001$), indicating exceptional performance in discriminating between HIV-negative and HIV-positive statuses. DNAGrimAge v2 also showed strong results with an AUC of 0.9595, sensitivity of 72.54%, and unchanged specificity ($P < 0.0001$) (Figure 2A).

Influence of an HIV treatment strategy on epigenetic aging

Building on these findings, we explored the dynamics of epigenetic aging among HIV-positive individuals to assess how a specific treatment strategy influenced these patterns.

We evaluated the impact of a combined HIV treatment strategy—comprising therapeutic HIV-1 vaccination and the latency-reversing agent romidepsin—using EpiAge, DNAmAge (Horvath's Clock), DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. This regimen also incorporated a monitored pause in antiretroviral therapy (MAP), as outlined by Oriol-Tordera et al. [34]. We tracked epigenetic aging from baseline (week 0, before treatment initiation) through several key intervals: weeks 1, 6, 10, and 24, as shown in Figure 2B.

EpiAge uniquely responded to the treatment, showing a significant deceleration in epigenetic aging by week 24, with an adjusted P-value of 0.0061, corresponding to an epigenetic age reduction of approximately 3.93 years from baseline.

DNAmAgeHannum also demonstrated a notable response by week 24, revealing a significant reduction in epigenetic age of 12.96 years (adjusted P-value < 0.0001). DNAmAgeSkinBloodClock displayed a moderate but significant deceleration of 1.88 years over the same period (adjusted P-value = 0.0215). The responses of other clocks did not reach statistical significance (Figure 2B).

Further, we specifically assessed changes from baseline to the Monitored Antiretroviral Pause (MAP) phase, employing paired t-tests to determine if the observed changes in epigenetic age were directly induced by the treatment regimen. This analysis, illustrated in Figure 2C, focused on the unique impacts of the therapeutic vaccination and latency-reversing agent during the pause in antiretroviral therapy. Unlike the continuous treatment period captured in Figure 2B, the MAP phase analysis aimed to isolate the effects of the intervention components without the confounding influence of ongoing antiretroviral therapy (ART).

Among the clocks evaluated, EpiAge and DNAmAgeHannum showed marked responses to the combined therapy despite the MAP. EpiAge demonstrated a significant deceleration in epigenetic aging with a mean difference of -4.811 years ($P < 0.0001$), while DNAmAgeHannum also indicated substantial deceleration with a mean difference of -12.23 years ($P < 0.0001$). DNAmAgeSkinBloodClock responded more modestly, showing a deceleration with a mean difference of -2.036 years ($P = 0.0129$). Surprisingly, DNAGrimAge v1 indicated an unexpected significant acceleration of aging, with a mean difference of +2.628 years ($P < 0.0001$).

These results underscore the distinct responses of epigenetic clocks to the HIV treatment regimen. EpiAge and DNAmAgeHannum, in particular, stood out for their sensitivity to treatment-induced changes, offering potential as reliable markers for assessing the impact of HIV therapies on biological aging.

Impact of COVID-19 on epigenetic aging: comparative analysis across multiple epigenetic clocks

Our findings indicated that the EpiAge clock showed an age acceleration that varied as a function of the COVID-19 severity score. As such, in our analysis of DNA methylation data from the GSE167202 dataset [35], we found that the EpiAge clock demonstrated a significant age acceleration in COVID-19 severity score 1 compared to negative controls ($P = 0.0002$), with an acceleration of 5.766 years (Figure 3A). Lesser or no significant differences were observed in higher severity scores or in response to an infection by other viruses.

Similarly, DNAmAge indicated significant age acceleration in patients with other infections and mild COVID-19 cases, with age accelerations of 2.514 years ($P = 0.0166$) and 3.935 years ($P = 0.0035$), respectively. DNAmAgeSkinBloodClock mirrored this pattern, showing significant age accelerations for other infections (2.312 years, $P = 0.0010$) and mild COVID-19

severity (2.268 years, $P=0.0225$), while no significant changes were observed for higher severity levels.

Hannum's clock only found COVID-19 severity to be significantly associated with age acceleration (3.506 years, $P=0.0064$), with other severity levels and infections showing no significant differences. PhenoAge displayed significant acceleration for other infections (3.535 years, $P=0.0031$) and remarkably for COVID-19 severity levels 3 (4.284 years, $P=0.0059$) and 4 (6.249 years, $P=0.0205$), but not for level 1.

GrimAge v1 did not show significant differences across groups. GrimAge v2 revealed that COVID-3 severity had a significant age acceleration (3.892 years, $P=0.0402$) compared to negative control, indicating nuanced responses across different epigenetic clocks to COVID-19 severity and other infections (Figure 3A).

In conjunction with our investigation, we analyzed dataset GSE168739, which was also utilized in a study by Cao et al. [36] This previous study examined the influence of COVID-19 on epigenetic aging in blood samples using a range of established epigenetic clocks, such as Horvath's DNAmAge, Hannum's DNAmAge, PhenoAge, SkinBloodClock, and GrimAge. The findings indicated that individuals with COVID-19 experienced a notable increase in DNAm age across all clocks compared to healthy individuals. Extending these observations, our application of the EpiAge clock to these samples demonstrated a significant acceleration in epigenetic aging among COVID-19 patients, with an average advancement of nearly four years ($p < 0.0001$) (Figure 3B). Similarly, DNAmAge revealed a moderate yet significant age acceleration ($p = 0.0114$, difference = 1.457 years). The SkinBloodClock also showed significant acceleration ($p < 0.0001$, difference = 2.648 years), while other clocks displayed no significant changes or, surprisingly, opposite effects such as those seen with GrimAge v1 and PhenoAge.

Based on our analyses, the EpiAge clock demonstrated the most pronounced acceleration in epigenetic aging in response to COVID-19 infection (COVID-19), markedly exceeding the age advancements detected by other clocks. This stark contrast highlights the EpiAge clock's sensitivity and effectiveness in capturing the impact of COVID-19 on biological aging processes, thus affirming its utility in clinical and epidemiological settings focused on the implications of infectious diseases on aging.

Current stress, not cumulative stress, is associated with epigenetic age acceleration

To investigate the link between stress and epigenetic age acceleration (EAA), we analyzed whole blood

samples from African American individuals using the GSE72680 dataset [37–39].

The original study employed a DNA methylation-based age prediction method, utilizing DNAmAge developed by Horvath [26]. They found that cumulative lifetime stress, as opposed to childhood maltreatment or current stress alone, predicted accelerated epigenetic aging in an urban African American cohort [39].

In our study, we expanded the analysis by using EpiAge, DNAmAge (Horvath's Clock), DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAmGrimAge v1, and DNAGrimAge v2. Contrary to the original findings, we did not detect any significant correlation between epigenetic age acceleration and cumulative life stress after correction for blood cell-type composition and lifestyle parameters, such as age, sex, body mass index (BMI), alcohol use, tobacco use, childhood sexual or physical abuse, childhood trauma, cocaine use, heroin use, marijuana use, posttraumatic stress disorder symptom scale, Beck Depression Inventory total score, and treatment for anxiety disorder, bipolar disorder, depression, and posttraumatic stress disorder.

Our investigation into the impact of stress on epigenetic age acceleration (EAA) across different epigenetic clocks reveals a complex relationship (Figure 4A). Initially, simple correlations between current stress and various clocks were established, with notable results as follows:

For Horvath's DNAmAge, a correlation coefficient (r) of 0.1517 and a p -value of 0.0064 were observed. After adjusting for confounders, a significant regression coefficient (β) of 0.5171 was noted, highlighting an association between higher current stress levels and increased EAA ($p = 0.030$). Cumulative life stress, however, did not exhibit a significant association ($p = 0.425$).

Hannum's clock displayed an initial r of 0.1556 and a p -value of 0.0051. Post-correction analysis revealed a β of 0.4831 ($p = 0.030$), reaffirming the significance of current stress, while cumulative life stress again showed no significant effect ($p = 0.399$).

PhenoAge presented an r of 0.117 and a p -value of 0.0347 between current stress and EAA. However, the relationship did not remain significant after correction.

DNAmAgeSkinBloodClock demonstrated an initial r of 0.1204 and a p -value of 0.0308 of the association between current stress and EAA, but this significance was not maintained post-correction.

The EpiAge clock showed a stronger initial correlation with current stress ($r = 0.1975$, p -value of 0.0004), which remained significant after correction ($p \leq 0.037$).

For GrimAge v1, we found an initial r of 0.2367 with a p -value of <0.0001 association with current stress, which persisted as significant after correction ($p = 0.011$).

GrimAge v2 revealed an r of 0.235 and a p -value of <0.0001 , with the association remaining significant post-correction ($p = 0.005$).

Cumulative life stress did not demonstrate a significant association with EAA across any of the clocks, both before and after applying corrections for confounding factors. This analysis underscores the specific impact of

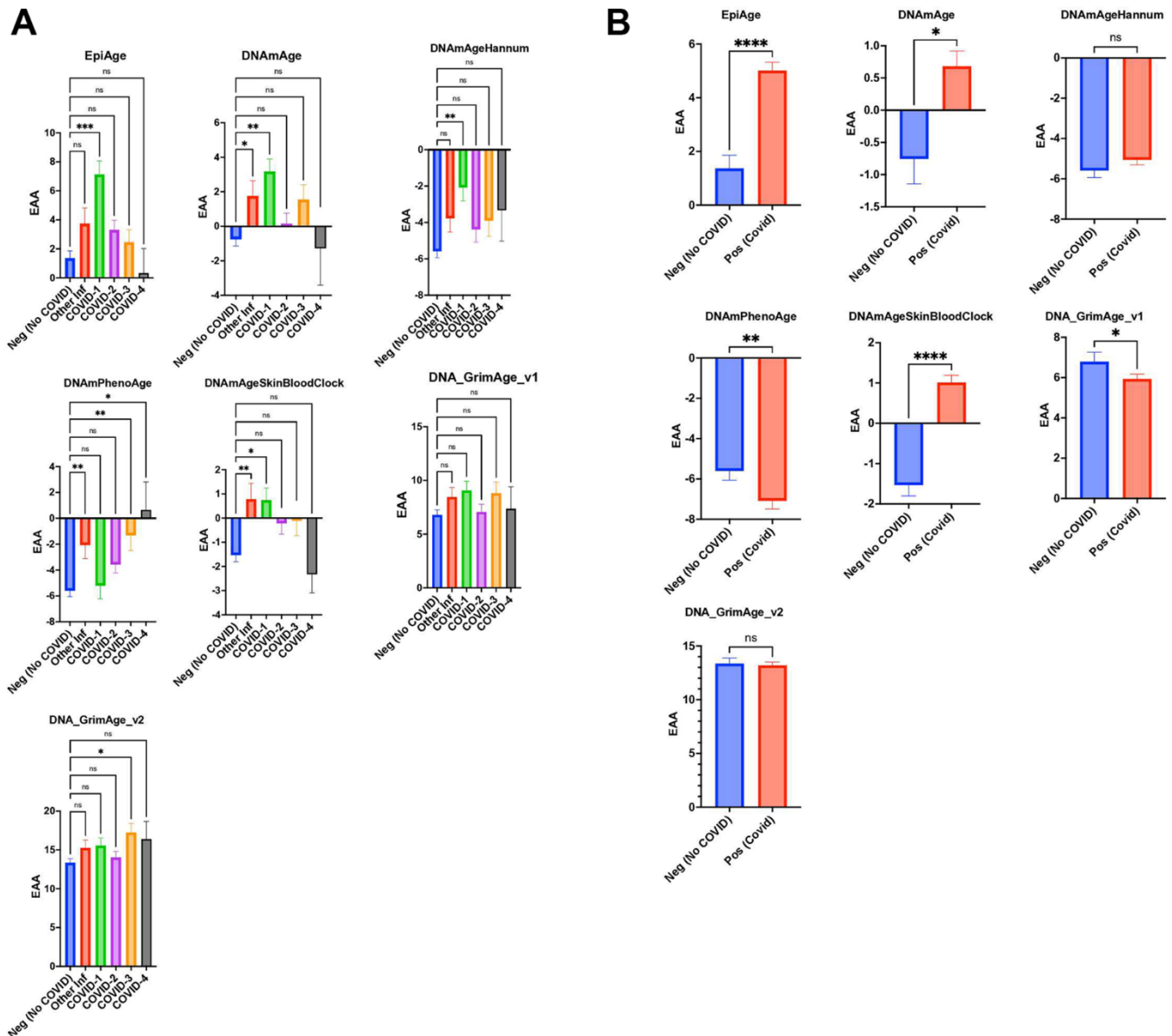


Figure 3. (A) Scatter plot analysis of epigenetic age acceleration (EAA) across individuals without COVID-19 (Neg, $n=296$) from GSE167202 and COVID-19 patients (Pos, $n=407$) from GSE168739. EAA is calculated by subtracting chronological age from EpiAge estimates. Median EAA values for each group are plotted. Non-parametric two-tailed Mann-Whitney U test reveals significant EAA in COVID-19 patients compared to non-COVID-19 individuals ($P < 0.0001$), with a median difference of 3.974 years, indicating accelerated biological aging in infected patients. (B) Scatter plot analysis comparing epigenetic age between healthy controls (GSE167202, $n=296$) and COVID-19 patients (GSE168739, $n=407$), utilizing clocks including EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. Due to the non-normal distribution of data, a Mann-Whitney U test was employed. Significance levels are denoted as ns (not significant), * ($p < 0.05$), ** ($p < 0.01$), *** ($p < 0.001$), **** ($p < 0.0001$).

current stress on epigenetic aging, which remained after corrections for blood cell-type composition and lifestyle parameters. These findings highlight the importance of current stress as a significant factor in epigenetic age acceleration across multiple epigenetic clocks.

We further categorized individuals from the GSE72680 dataset [37] into five groups based on the intensity or severity of stress: No Stress (Level 0), Low Stress (Levels 1-4) for mild stress experiences, Moderate Stress (Levels 5-7), High Stress (Levels 8-10), and Severe Stress (Levels 11-15). We examined the association of these stress levels with Epigenetic Age Acceleration (EAA) using the EpiAge clock (Figure 4B). The analysis revealed a gradual increase in EAA from the No Stress Group to the Severe Stress Group. Specifically, we noted a borderline significant difference between the Low Stress and No Stress

groups ($P = 0.1811$) and the Moderate Stress Group ($P = 0.0604$), a significant difference between the High Stress and No Stress groups ($P = 0.0268$), and the most pronounced difference was observed between the Severe Stress and No Stress groups ($P = 0.0143$).

A similar trend was observed across various clocks: For DNAmAge, the comparisons yielded P-values of 0.4334 (Low vs. No Stress), 0.2114 (Moderate vs. No Stress), 0.195 (High vs. No Stress), and 0.0443 (Severe vs. No Stress). Hannum's clock followed suit with P-values of 0.3713, 0.2711, 0.1466, and 0.0476, respectively. The trend persisted with PhenoAge and DNAmAgeSkinBloodClock, with PhenoAge showing a borderline significant difference at high-stress levels ($P=0.0428$) compared to no stress and DNAmAgeSkinBloodClock revealing a significant difference at severe stress levels ($P=0.0352$). DNA_GrimAge_v1

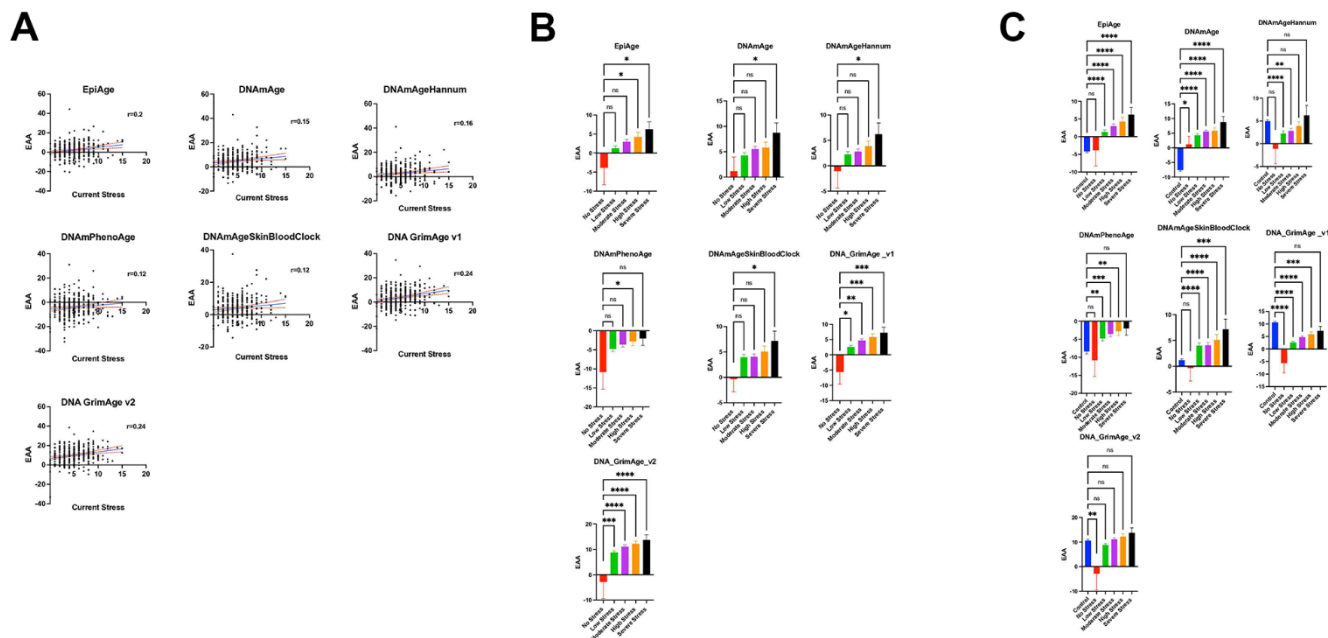


Figure 4. (A) Relationship between current stress levels (Axe X) and epigenetic age acceleration (EAA) (axe Y) across various epigenetic clocks (EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2) from the study GSE72680, analyzed using Pearson's r . Stress level severity is represented by increasing numbers where a score of 0 means No Stress, and 15 is the most severe stress. Each plot in this figure represents a different clock, highlighting the strength and significance of their relationships with current stress. The central line in each plot represents the regression line, accompanied by two red lines which delineate the 95% confidence intervals, underlining the precision of the correlation estimates. (B) Scatter plots illustrate the relationship between varying stress levels and epigenetic age acceleration across different epigenetic clocks. The plot compares five groups categorized by stress intensity: No Stress (Level 0, $n=9$), Low Stress (Levels 1-4, $n=149$), Moderate Stress (Levels 5-7, $n=106$), High Stress (Levels 8-10, $n=48$), and Severe Stress (Levels 11-15, $n=13$). Each point represents an individual's epigenetic age acceleration, with group comparisons analyzed using ordinary one-way ANOVA and Dunnett's multiple comparisons test to assess statistical differences between the No Stress group and each of the stress intensity groups. (C) Scatter plots depicting the relationship between stress levels and epigenetic age acceleration (EAA) across various epigenetic clocks, incorporating control groups for a comprehensive analysis. This plot includes six groups: No Stress (Level 0, $n=9$), Low Stress (Levels 1-4, $n=149$), Moderate Stress (Levels 5-7, $n=106$), High Stress (Levels 8-10, $n=48$), Severe Stress (Levels 11-15, $n=13$), and an added Control group ($n=419$, combining subjects from datasets GSE128235 and GSE125105). Each point illustrates an individual's EAA. Group differences were statistically evaluated using ordinary one-way ANOVA with Dunnett's multiple comparisons test, comparing each stress level group as well as the Control group against the No Stress group to identify significant variations in EAA.

and v2 showed increasingly significant differences with rising stress levels, underscoring a clear association between stress and epigenetic age acceleration. Specifically, DNA_GrimAge_v1 and v2 displayed significant differences at even low-stress levels ($P=0.0110$ and $P=0.0009$, respectively), becoming more pronounced at moderate to severe stress levels, culminating in $P<0.0001$ for severe stress.

These analyses, conducted using one-way ANOVA with Dunnett's multiple comparisons test, highlight a consistent trend across all clocks, indicating a significant relationship between stress levels and epigenetic age acceleration.

The GSE72680 dataset had a limited number of 9 samples in the no stress group, which might explain the loss of power to detect EAA differences in the lower stress groups. We therefore included additional control (normal stress) subjects from a different data set GSE128235 (209 subjects) and GSE125105 (210 subjects) datasets in the analysis, as illustrated in Figure 4C. Importantly, there was no significant difference in EpiAge between these newly added control samples and the no-stress group included in the GSE72680 data (Adjusted P-value >0.9999), which suggested that these control groups could be combined in our analysis. When the analysis included a larger sample of "normal" controls, a marked increase in Epigenetic Age Acceleration (EAA) was associated with stress from low to severe stress levels, which was significant even in the low stress group (Adjusted P-value <0.0001).

For DNAmAge, the no-stress group exhibited significantly higher EAA than the control group ($p = 0.0197$), with a more pronounced EAA observed across all stress levels in comparison to the control group.

The DNAmAgeHannum analysis indicated a borderline acceleration in EAA within the no-stress group compared to the added control group (Adjusted P-value $=0.0720$). However, EAA was notably higher in the control group than in both the low and moderate stress groups, with no significant differences noted when compared to the high and severe stress levels.

The PhenoAge analysis aligned with that of EpiAge, showing no distinction between the control and no-stress groups. Yet, there were significant escalations in EAA from low to high-stress levels in comparison to the control group. Intriguingly, PhenoAge suggested an overall deceleration of EAA across all groups, regardless of stress level. Similarly, DNAmAgeSkinBloodClock analysis showed no variance between the control and no-stress groups but

revealed significant increases in EAA from low to severe stress levels.

Notably, DNA GrimAge v1 indicated a significantly higher EAA across all groups except for the severe stress group when compared to the control group. DNA GrimAge v2 exhibited significantly higher EAA in the control group than in the no-stress group but did not show marked differences from other stress levels.

Altogether, these findings indicate that EpiAge showed the clearest sensitivity in detecting the stress-induced epigenetic age acceleration, whereby the differences in EAA were only observed between stressed and both controls and non-stressed controls. Furthermore, although all clocks were sensitive to stress, there was observed variability across different epigenetic clocks. These may be attributed to the varied number of CpGs analyzed in the different clocks and the absence of CpGs included in some clocks in the datasets that used different array designs. For example, there was a consistent issue with missing probes in the GSE72680 dataset for DNAmGrimAge2 (ranging typically from 57 to 64 out of 1331 probes) and DNAmFitAge (consistently 48 out of 789 probes). Furthermore, the unexpected results seen with GrimAge v1 and v2 in control groups, derived from diverse studies, may reflect batch effects. It is advised against attempting to correct or process the array data in ways that might lead to the exclusion of CpGs. These observations emphasize the consistency and reliability offered by utilizing fewer CpGs in the EpiAge analysis, which appears to yield more stable results across different conditions.

Epigenetic age acceleration in Down syndrome

To further assess the value of the EpiAge clock in detecting epigenetic age acceleration in disease states, in this study, we analyzed 22 healthy individuals and 20 individuals with Down syndrome (DS) using the Epic array. Down syndrome is the most common chromosomal anomaly in humans that causes mild to significant developmental, physical, and intellectual disabilities. Due to the trisomy of chromosome 21, which harbours the gene encoding for the amyloid precursor protein (APP), people with Down syndrome progressively develop Alzheimer's disease neuropathology starting early in life.

The two groups showed no significant chronological age difference, with the mean age for the healthy group at 40.95 and for the Down syndrome group at 40.65, and a p-value of 0.9146 as revealed using an unpaired parametric T-test (applied due to normal distribution). We then calculated epigenetic age using the EpiAgePublic model, among others and assessed epigenetic age acceleration (EAA) (Figure 5A).

Significant age acceleration in the Down syndrome group was observed across all clocks using T-tests and Receiver Operating Characteristic (ROC) analysis to examine the differences. Specifically, for EpiAgePublic,

the p-value was 0.0002, with a mean difference \pm SEM of 10.26 ± 2.539 and an ROC of 0.8205. DNAmAge showed a p-value of 0.0004, with a mean difference of 9.613 ± 2.481 and an ROC of 0.8341.

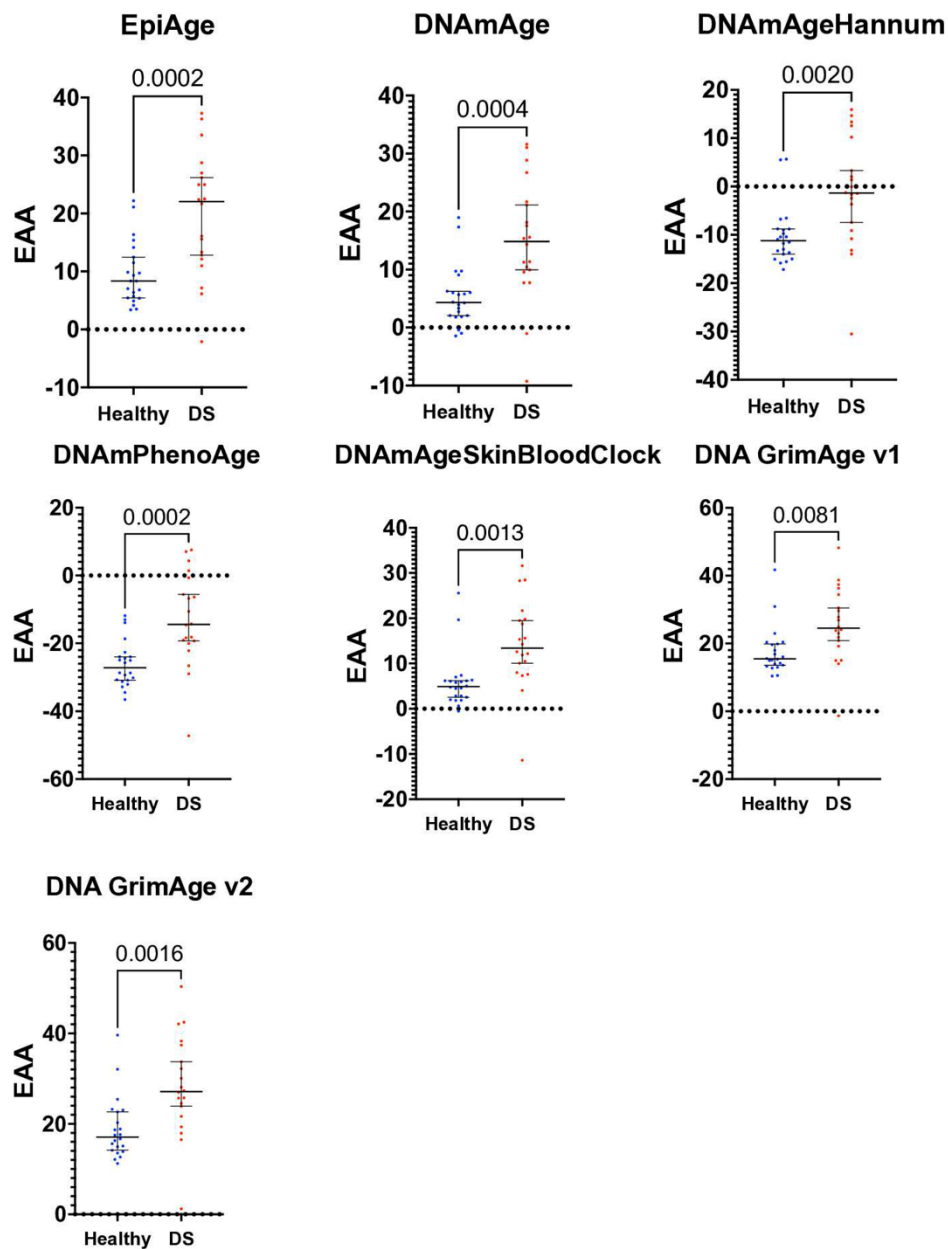


Figure 5. Scatter plot of epigenetic age acceleration (EAA) in healthy individuals and those with Down syndrome. This scatter plot illustrates the Epigenetic Age Acceleration (EAA) for each participant in our study, contrasting healthy controls (n=22) with individuals diagnosed with Down syndrome (DS, n=20). EAA, determined by the discrepancy between epigenetic age and chronological age, is plotted for each individual, utilizing various epigenetic clocks: EpiAgePublic, DNAmAge, DNAmAgeHannum, PhenoAge, DNAmAgeSkinBloodClock, GrimAge v1, and GrimAge v2. Each point represents one individual's EAA, with separate color codes for healthy controls and DS individuals. Unpaired t-tests were conducted to examine the differences in EAA between the groups, showing significant increase in EAA in DS: EpiAgePublic (p=0.0002), DNAmAge (p=0.0004), DNAmAgeHannum (p=0.002), PhenoAge (p=0.0002), DNAmAgeSkinBloodClock (p=0.0013), GrimAge v1 (p=0.0081), and GrimAge v2 (p=0.0016). ROC analysis, discussed in the results section, was performed to further assess the discriminative capability of each clock between the healthy and DS groups. This plot highlights the significant epigenetic age acceleration observed in individuals with Down syndrome in comparison to healthy controls.

DNAmAgeHannum had a p-value of 0.002, with a mean difference of 9.177 ± 2.771 and an ROC of 0.7932. PhenoAge's p-value was 0.0002, with a mean difference of 13.16 ± 3.272 and an ROC of 0.8318. The DNAmAgeSkinBloodClock had a p-value of 0.0013, with a mean difference of 8.439 ± 2.446 and an ROC of 0.8523. GrimAge v1 had a p-value of 0.0081, with a mean difference of 7.799 ± 2.798 and an ROC of 0.7818, while GrimAge v2 showed a p-value of 0.0016, with a mean difference of 9.404 ± 2.779 and an ROC of 0.8227. (The mean difference \pm SEM represents the difference in the mean epigenetic age acceleration (EAA) between the Down syndrome and healthy control groups, along with the standard error of that difference.)

Overall, the EpiAgePublic clock performs comparably to other established epigenetic clocks in detecting age acceleration in individuals with Down syndrome.

Epigenetic age and Alzheimer's disease

In all previous analyses, we utilized array technology, which is known for its limitations. This technology requires normalization methods such as BMIQ and others to adjust for variability in methylation data. However, a major concern with array technologies, such as the Illumina Infinium Methylation BeadChip often used in epigenome-wide association studies (EWAS), is the presence of batch effects. These effects can reduce experimental power and potentially lead to false positive results due to variations in the day of processing, the individual glass slide, and the array's position on the slide. Despite employing batch-effect removal tools like ComBat and Harman, residual batch effects persist and require significant correction [40].

In contrast, targeted next-generation sequencing offers advantages over array technologies, including higher resolution and greater specificity for detecting methylation changes across diverse biological samples. This method is less susceptible to batch effects and does not rely on pre-designed probes, allowing for a more comprehensive analysis of the methylome.

To validate the EpiAge model using targeted NGS, we conducted an observational longitudinal clinical study where we analyzed DNA extracted from the lymphocytes of peripheral blood mononuclear cells (PBMCs) from 55 individuals, categorized into Alzheimer's disease patients and control participants. One participant was excluded from the analysis due to technical issues, leaving a total of 54 participants. Among these, there were 10 males and 16 females in the control group and 7 males and 21 females in the Alzheimer's disease (AD) group. The mean age of the

control group was 71.53 years (SD = 9.36), while the AD group had a mean age of 72.57 years (SD = 8.89). A parametric t-test showed no significant difference in age between the control and AD groups ($p = 0.6793$).

For this study, we expanded the region analyzed in the *ELOVL2* gene to include not only the well-established CpG sites—cg16867657, cg21572722, and cg24724428—from the Illumina array but also an additional ten CpG sites, totaling 13 sites. This broader scope was intended to enhance the technical robustness of the EpiAge model.

First, we examined the replication consistency across our samples. Due to limitations in the amount of DNA obtained, only 41 out of the total 54 individuals had sufficient DNA to perform analyses in four replicates.

In our analysis, the coefficient of variation (CV) was calculated for each of the 13 CpG sites to evaluate the consistency of methylation measurements obtained via next-generation sequencing. The CV was determined by dividing the standard deviation by the mean methylation level for each site and expressed as a percentage. The results, depicted in box plots (Figure 6A), indicate that the CVs for the first nine CpG sites were relatively low, ranging from 0.28% to 6.7%, as these CpGs exhibited high percentages of methylation. In contrast, the CVs for sites 10 through 13, which are characterized by lower methylation levels (around 10% or less), showed greater variability, ranging from 1% to 25% across individuals. This increased variability at lower methylation levels aligns with statistical expectations, where measurements near the extremes tend to have lower relative variability, while those closer to the midpoint or at lower levels tend to fluctuate more [41, 42].

In our study, we assessed the technical accuracy of EpiAge in a cohort of 41 participants (Figure 6B), each tested four times using targeted next-generation sequencing (NGS). We first evaluated the Confidence Interval (CI) Range as a measure of variability in the epigenetic age estimation across repeated measurements:

CI Range <1: Observed in 14 individuals, suggesting a high precision in epigenetic age estimation for approximately 34% of the participants.

CI Range >1 and <1.5: Found in 15 individuals, indicating moderate variability in epigenetic age estimates, affecting about 37% of the study population.

CI Range >1.5 and <2: Present in 7 individuals, reflecting a greater variability in measurements, impacting around 17% of the participants.

CI Range >2 and <3: Noted in 4 individuals, demonstrating significant variability, which was evident in approximately 10% of the cohort.

CI Range >3: This highest variability was observed in 1 individual, accounting for about 2% of the participants.

Additionally, we summarized key descriptive statistics, including mean, standard deviation (SD), and coefficient of variation (CV) for each participant in Supplementary Table 2. The results showed that the CVs ranged from 0.57% to 6.10%, demonstrating varying levels of technical accuracy.

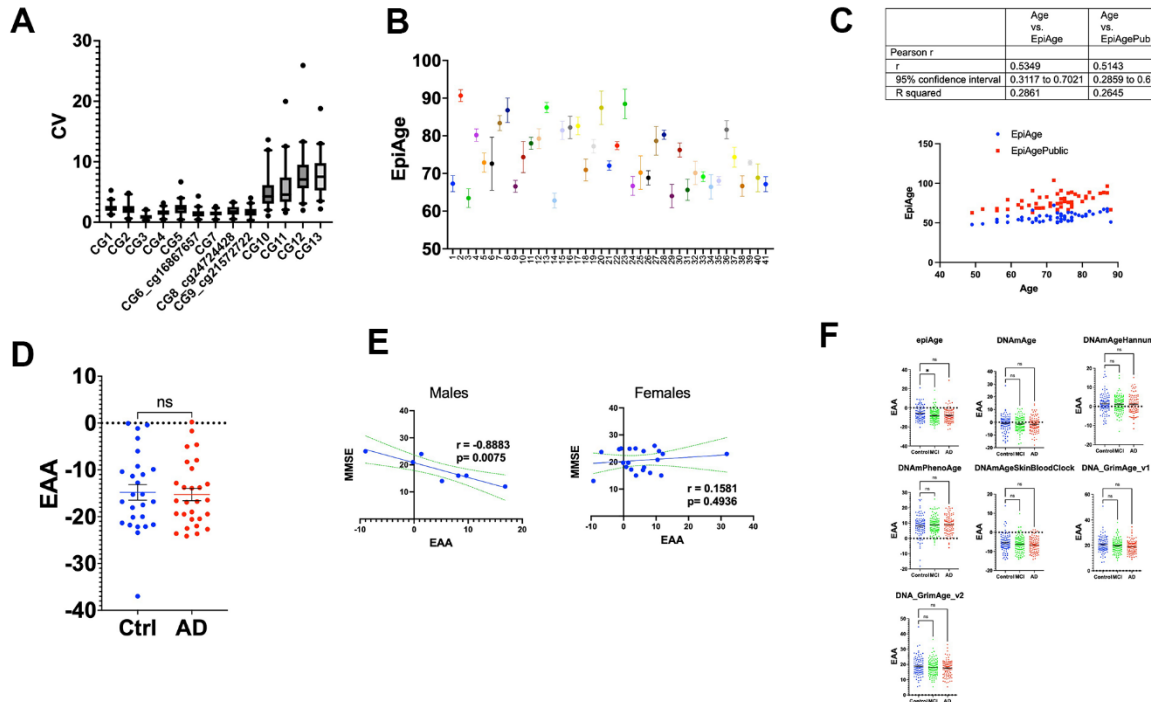


Figure 6. (A) Box plots of coefficient of variation (CV) for methylation levels across 13 CpG sites. This figure illustrates the distribution of the coefficient of variation (CV) for each of the 13 CpG sites within the ELOVL2 gene, analyzed using next-generation sequencing. The CV was calculated by dividing the standard deviation of methylation measurements by the mean for each site, expressed as a percentage. The box plots show the interquartile range (25th to 75th percentile) with whiskers extending to the 5th and 95th percentiles. Sites 1-9 exhibit lower CVs, ranging from 0.28% to 6.7%, indicative of high methylation consistency. Sites 10-13 display higher CVs, ranging from 1% to 25%, reflecting increased variability in regions of lower methylation. This variability highlights the influence of methylation levels on the precision of epigenetic age assessments. (B) The figure displays a scatter plot of epigenetic age measurements using our newly developed EpiAge next-generation sequencing assay for 41 study participants, using the linear regression model developed for EpiAgePublic in samples that had four technical replicates. Each point on the x-axis corresponds to the average EpiAge calculated for a blood sample (buffy coat) of an individual participant. The y-axis indicates the epigenetic age calculated for each replicate. Error bars represent the 95% confidence intervals for the mean epigenetic age of each individual. (C) Comparison analysis between EpiAgePublic (red dots) and EpiAge calculated using the HKG epiTherapeutics proprietary model (blue dots) relative to chronological age (Axe X). Correlation analysis was performed using Pearson's R. The table provides a comparison of the correlation coefficient (r), 95% confidence interval, and R-squared values. (D) Epigenetic Age Acceleration (EAA) Comparison between control and AD patients calculated using EpiAge next-generation sequencing assay. This scatter plot compares the EAA between control participants (n=26) and Alzheimer's Disease (AD) patients (n=28). Each dot represents an individual, and the plot shows the mean with SEM. The epigenetic age was calculated using the HKG epiTherapeutics proprietary model. Differences in EAA between the groups were analyzed using a two-tailed parametric t-test to assess statistical significance. (E) The left panel displays the Pearson correlation between Epigenetic Age Acceleration (EAA) and Mini-Mental State Examination MMSE in 7 male Alzheimer's patients, while the right panel displays the correlation in 21 female Alzheimer's patients. The middle line in each plot represents the linear regression fit, while the two lines surrounding it represent the 95% confidence bands, which indicate the variability of the correlation. (F) Comparative Analysis of Epigenetic Age Acceleration Across Alzheimer's Disease, Mild Cognitive Impairment, and Control Groups Using Multiple Epigenetic Clocks. This figure presents the comparison of Epigenetic Age Acceleration (EAA) using multiple clocks: EpiAge, DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, DNAGrimAge v1, and DNAGrimAge v2. The data were derived from 96 control individuals, 111 with mild cognitive impairment (MCI), and 93 with Alzheimer's Disease from dataset GSE144858, using DNA from human blood. The plot shows means with Standard Error of the Mean (SEM). For statistical analysis, we employed an ordinary one-way ANOVA to compare AD (Alzheimer's) to controls and MCI to controls. Parametric ANOVA was used due to the normal distribution of the data. 'Ns' stands for not significant; * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$; and **** for $p < 0.0001$.

- These results show that the targeted NGS-based EpiAge exhibited high reliability, with most participants displaying low CI ranges and low coefficients of variation.
- The CVs indicate that the method's precision varies across individuals but generally falls within acceptable limits for biological measurements.
- NGS offers improved accuracy compared to traditional array-based methods, but some variability remains, potentially due to noise in the replicates.
- The findings highlight the accuracy of EpiAge by NGS, although further optimization could improve consistency and reduce noise in repeated measurements.

These findings demonstrate that while most study participants exhibited relatively low to moderate variability in epigenetic age estimates, a smaller group showed significant variability. This distribution underscores the importance of considering technical variability in such analyses. While targeted next-generation sequencing provides high-resolution and specific detection of methylation changes, there is inherent variability that must be accounted for in clinical and research settings. The analysis of CI Ranges provides important insights into the reliability of epigenetic age measurements and their potential application in longitudinal studies and interventions aimed at aging and related neurodegenerative disorders.

Furthermore, these results highlight the importance of evaluating confidence intervals when interpreting epigenetic age to ensure robust conclusions. This is particularly vital in clinical or research contexts where precision is paramount. Relying solely on single measurements without technical replicates, which is common with less accurate array technologies, overlooks the importance of accuracy and the limitations of study design. Without this consideration, it becomes challenging to discern whether observed changes in epigenetic age across multiple time points are due to actual biological effects, batch effects, or merely technical noise.

We subsequently calculated the epigenetic age using the bisulfite-targeted next-generation sequencing data and the EpiAgePublic model (including 3 CpGs: cg16867657, cg21572722, and cg24724428), which was trained on GSE55763, GSE157131, GSE40279, and GSE30870 (Table 1), as discussed above. We also calculated EpiAge using a proprietary model

developed by HKG epiTherapeutics, as discussed in the Methods section, which was originally developed using saliva samples (Figure 6C). A comparison analysis between the two models revealed a slightly better performance for the proprietary model ($r = 0.5349$) compared to EpiAgePublic ($r = 0.5143$). The Epigenetic Age Acceleration (EAA) was determined by subtracting the EpiAge value from the proprietary model from the chronological age for 26 control participants and 28 patients with Alzheimer's Disease (AD). Our analysis revealed no significant difference between the two groups (parametric t-test, $p = 0.8188$) (Figure 6D).

We further analyzed whether Mini-Mental State Examination (MMSE) is associated with EAA separately in females and males. Interestingly, a strong negative correlation between MMSE T0 and epigenetic age acceleration was observed in males ($r = -0.8883$, $p = 0.0075$) (Figure 6D), while no significant correlation was observed in females (Figure 6E).

To examine EAA in AD patients in other studies, we analyzed the Epic Array dataset GSE144858 [43], derived from the EU-funded AddNeuroMed Cohort, a large cross-European AD biomarker study using human blood DNA [44, 45]. This dataset included 96 control individuals, 111 with mild cognitive impairment (MCI), and 93 with Alzheimer's Disease (AD).

We applied the EpiAgePublic model and other epigenetic clocks (Figure 6E). Consistent with the bisulfite next-generation sequencing EpiAge assay, we did not observe any significant EAA differences between AD and controls, based on the Kruskal-Wallis test, since the data did not pass the normality test. Notably, EpiAgePublic was the only model showing significant age deceleration in MCI compared to controls (adjusted p-value = 0.0339).

DISCUSSION

This study introduces epiAgePublic, a novel epigenetic aging model utilizing only three CpG sites within the *ELOVL2* gene, traditionally known for its strong association with aging [24, 25].

The simplicity and precision of epiAgePublic, designed for compatibility with next-generation sequencing (NGS) technologies, mark a significant step forward in the field of epigenetic research. Our findings demonstrate that epiAgePublic can effectively estimate biological age with an accuracy comparable to more complex, established epigenetic clocks. This could potentially streamline and reduce the cost of biological age assessments in clinical settings.

In developing the epiAgePublic model, we focused on three CpG sites within the *ELOVL2* gene—cg16867657, cg21572722, and cg24724428—known for their links to aging, supported by a body of research (Bell et al., 2012; Christiansen et al., 2016; Hao et al., 2021; Horvath & Raj, 2018; Li et al., 2022; Marioni et al., 2015). Employing linear regression, we trained a model utilizing a training cohort compiled from four public databases (GSE55763, GSE157131, GSE40279, and GSE30870), capturing a wide demographic range (Table 1). The diversity of this training cohort was crucial for developing an inclusive EpiAgePublic model, which achieved an R-squared value of 0.7512.

The model was validated on several cohorts (Table 1 and Figure 1) and was highly correlated with chronological age. Furthermore, our comparative analysis with established epigenetic clocks revealed EpiAgePublic's robustness and its ability to assess biological aging. epiAgePublic remains significantly correlated with chronological age even when potential confounding factors like blood cell composition and sex are included in the model. Age consistently emerged as a significant predictor across all models, underscoring its fundamental role in epigenetic aging, while sex exhibited a nuanced influence, revealing subtle but significant associations with all models except epiAgePublic.

The deceleration of epigenetic aging relative to chronological age, as individuals grow older suggests a nuanced interaction between genetics, environment, and aging processes. This phenomenon (Figure 1C) might reflect adaptive mechanisms or biological resilience in response to environmental stressors and lifestyle factors across the lifespan. This deceleration of epigenetic aging in long-livers also aligns with findings from other studies [46].

After developing the EpiAgePublic model, a crucial validation step was undertaken to confirm its utility not just for blood samples but also for saliva, thereby expanding its application range. This validation is pivotal, serving both as an independent test on fresh datasets and as an extension of the model's utility beyond its initial scope, moreover increasing the feasibility and availability of the test as a tool for monitoring public health without requiring invasive blood draw.

The cellular composition of saliva, as well as blood, could vary across individuals, reflecting varied proportions of epithelial and immune cells. Specifically, saliva in children comprises approximately 35% epithelial cells and 65% immune cells [47], whereas adult saliva contains around 80% epithelial cells and 20% immune cells [48, 49]. Nevertheless, our data show a high correlation across ages between the EpiAge

predicted in blood and saliva. This is consistent with previous studies. The *ELOVL2* gene has been consistently identified as a predictive marker of age in saliva samples [50, 51]. The EpiAgePublic model, originally developed for blood, could be effectively applied to saliva, enabling a highly accessible biological tool for measuring biological age.

Sex confounding effects; EpiAge versus other clocks

Unlike other epigenetic clocks that exhibit a significant correlation with sex, EpiAge's lack of association with sex may indicate a more refined capacity to isolate the biological essence of aging from sex-based epigenetic variation. This positions EpiAge as a more universally applicable tool for assessing biological age. Therefore, EpiAge could offer a clearer, more focused lens through which to study the aging process, unencumbered by the variation introduced by sex. This advantage underscores the importance of developing and utilizing epigenetic clocks that can accurately reflect the aging process in a manner that is as inclusive and representative as possible of the general population.

Thus, EpiAgePublic, despite being nimble, relying on a single small genomic region is nevertheless an effective measure of the aging process comparable or even superior to other clocks that incorporate hundreds or even thousands of CpG sites.

Understanding HIV-related accelerated epigenetic aging

The investigation into HIV-related accelerated epigenetic aging represents a critical aspect of our study, shedding light on the intricate relationship between HIV infection and biological aging. Building upon previous research linking HIV to expedited epigenetic aging processes, our study delved into the specific impacts of HIV infection on epigenetic aging using multiple cohorts and compared the performance of various epigenetic clocks, including EpiAge, in elucidating these patterns.

Our findings corroborate earlier observations [52] suggesting a significant association between HIV infection and accelerated epigenetic aging. Notably, EpiAge demonstrated remarkable sensitivity in detecting age acceleration among HIV-positive individuals, with a substantial average advancement of 12.04 years compared to HIV-negative controls. This aligns with previous studies highlighting the accelerated aging phenomenon in HIV-infected populations and underscores the utility of EpiAge as a reliable marker for assessing epigenetic changes associated with HIV infection.

Furthermore, our comparative analysis across multiple epigenetic clocks revealed consistent evidence of aging acceleration in HIV-positive individuals, as indicated by DNAmAge, DNAmAgeHannum, DNAmPhenoAge, DNAmAgeSkinBloodClock, and DNAGrimAge v1. This acceleration is possibly due to inflammatory processes associated with HIV infection and, therefore, might be related to other inflammatory diseases as well. These findings underscore the robustness of epigenetic clocks in capturing the biological consequences of HIV infection on aging processes, thereby providing valuable insights into the pathophysiology of HIV-related accelerated aging.

In addition to identifying accelerated epigenetic aging patterns, our study evaluated the discriminatory power of epigenetic clocks in distinguishing between HIV-negative and HIV-positive individuals. ROC analysis highlighted the superior performance of EpiAge in accurately classifying individuals based on their HIV status, further underscoring its potential as a sensitive biomarker for HIV-related accelerated aging.

Moreover, our investigation into the influence of HIV treatment strategies on epigenetic aging dynamics yielded intriguing results. The observed deceleration in epigenetic aging following specific pharmacological treatment interventions underscores the potential reversibility of age-related epigenetic changes in HIV-positive individuals. Notably, EpiAge and DNAmAgeHannum exhibited significant responses to treatment, suggesting their utility as sensitive indicators of treatment-induced changes in biological aging. An open question remains in future studies to establish the specific effects of the different drug classes used in current antiretroviral therapy on epigenetic aging [33] and their correlation with changes in CD4+ cell counts and the CD4/CD8 ratio as well as with long-term clinical outcomes in individuals with HIV.

Overall, our study provides comprehensive insights into the complex interplay between HIV infection, epigenetic aging, and treatment interventions. By elucidating the underlying mechanisms driving HIV-related accelerated aging and evaluating the efficacy of epigenetic clocks in capturing these dynamics, our findings contribute to the growing body of knowledge aimed at improving clinical outcomes and therapeutic strategies for HIV-infected individuals.

Impact of COVID-19 on biological aging as measured by the different epigenetic clocks

The relationship between COVID-19 infection and epigenetic age acceleration (EAA) is complex and continues to be a subject of significant scientific debate.

While some studies have found no notable epigenetic acceleration in COVID-19 patients [53], others have reported marked changes [36, 54]. Our study shows the effects of COVID-19 on biological aging, with variations in response depending on the severity of the disease and the specific epigenetic clock used.

Different epigenetic clocks, including the EpiAge metric described here, show different sensitivity to changes induced by the virus. This variability can be attributed to the clocks' distinct molecular foundations and their differential responsiveness to the biological pathways affected by COVID-19.

For instance, the EpiAge metric, designed to capture age-related changes through specific CpG sites within the ELOVL2 gene, detects accelerated aging in less severe COVID-19 cases. This suggests that EpiAge may be particularly sensitive to early biological changes that other clocks might miss. Surprisingly, as COVID-19 severity increased, the EAA did not show significant differences. This could be attributed to severe cases overwhelming the immune system, thereby masking such changes.

Conversely, clocks like DNAmAgeSkinBloodClock and DNAGrimAge, which incorporate a broader array of CpG sites, showed significant changes primarily in cases with higher disease severity, reflecting their potential to capture larger biological disturbances.

Our analysis using multiple epigenetic clocks underscores the complex landscape in which COVID-19 infection and its severity levels affect biological aging. Notably, our findings are in line with other recent studies reporting accelerated epigenetic aging in COVID-19 patients [36], as further evidenced by our analysis of the GSE168739 dataset. These results support the utility of EpiAge and similar metrics in shedding light on the biological consequences of COVID-19, underscoring the importance of epigenetic clocks in understanding the broader implications of this and other infectious diseases on human health and aging. In future long-term observational studies, it would also be interesting to evaluate the potential relationship between Epigenetic Age Acceleration and cognitive dysfunction in individuals with post-COVID-19 syndrome.

Stress and epigenetic age acceleration

Existing research on the association between stress and epigenetic age acceleration has shown conflicting or inconclusive results [55]. This inconsistency may stem from the selection of different genes in various epigenetic models, which might not be responsive to

stress factors. The EpiAge model, which incorporates the *ELOVL2* gene, demonstrates, however, a clear association between current stress and epigenetic age acceleration. This suggests that the choice of stress-responsive genes, like *ELOVL2*, is crucial for accurately assessing the impact of stress on biological aging.

In our analysis of the EpiAge model, we've uncovered a noteworthy link between current stress and epigenetic age acceleration. Current stress, arising from immediate and ongoing challenges in an individual's daily life and reflects the current state of an individual's personal, professional, and social life, encompassing recent events or situations that are directly impacting their well-being now, shows a unique association with accelerated epigenetic aging. This contrasts with the lack of similar associations with other stress types, like cumulative life stress, which represents the lifelong accumulation of stressors [56], network life stress arising from social network events [57], and personal life stress, which focuses on individual-specific life circumstances [58]. These findings, although counterintuitive, highlight the distinct impact of current, day-to-day stressors on biological aging, as opposed to the more chronic or relational stresses captured by other categories. This indicates a unique impact of current stressors, possibly related to their immediate biological effects and to the possibility that the effect of stress on aging is potentially reversible once the burden of current stress is relieved.

This distinct association between current stress and accelerated epigenetic aging opens new avenues for understanding the biological underpinnings of stress responses. It also raises questions about the reversibility of such epigenetic changes with stress management or resolution. Future research might explore the mechanisms behind this phenomenon and investigate whether interventions targeting current stress can effectively decelerate epigenetic aging. This could have profound implications for stress management strategies and their role in healthy aging and stress-related disorders. Zannas et al. have demonstrated that aging and stress can epigenetically synergize in stress-related disorders such as Major Depressive Disorder (MDD) [37]. However, it is currently unknown whether Epigenetic Age Acceleration occurs in severe depressive phenotypes.

Interestingly, our observation that no significant correlation was found between epigenetic age acceleration and cumulative life stress after adjusting for blood cell-type composition and lifestyle parameters, as previously reported by Zannas et al., [39] is partially supported by the recent research conducted by Poganic et al. [59]. They suggest that the effects of

acute stress are more pronounced and directly associated with measures of epigenetic age acceleration. This finding is in line with our results, highlighting the predominant influence of current stress over cumulative stress on epigenetic age acceleration, even after controlling blood cell-type composition and lifestyle parameters.

Epigenetic age acceleration in Down syndrome: insights, implications, and clinical applications

Epigenetic age acceleration in Down syndrome has been a subject of interest in various studies. Down syndrome, characterized by the presence of all or part of a third copy of chromosome 21, has been associated with accelerated epigenetic aging. Studies demonstrated that trisomy 21 significantly increased the biological age of blood and brain tissue by an average of 6.6 years [60]. Furthermore, it has been noted that age acceleration can be observed in both blood and brain tissue in Down syndrome (Milicic et al., 2022). These findings collectively suggest that Down syndrome is associated with accelerated epigenetic aging, as evidenced by various studies focusing on epigenetic age acceleration in individuals with this condition. Additionally, moderate acceleration of epigenetic aging has been described in Down syndrome in another study (Cypris et al., 2020). Other studies have also found evidence of epigenetic age acceleration in newborns with Down syndrome (Xu et al., 2022). The research highlights the importance of understanding the epigenetic mechanisms underlying aging in Down syndrome and its implications for health outcomes in affected individuals. Research has shown that the use of pan-tissue DNA methylation clocks revealed epigenetic age acceleration in segmental progeroid syndromes like Down syndrome (Ashapkin et al., 2019). Furthermore, recent studies have indicated that epigenetic age acceleration can occur in segmental progeria conditions such as Down syndrome (Lee, 2023).

In our study, we further explored epigenetic age acceleration in individuals with Down syndrome using various epigenetic clocks, including the EpiAgePublic model. Our findings corroborate earlier research indicating that Down syndrome, characterized by an extra copy of chromosome 21, is associated with accelerated biological aging. This acceleration was consistently observed across multiple epigenetic clocks, affirming the robustness of these tools in capturing biological age differences.

The comparable performance of the EpiAgePublic clock with other established models suggests its utility as a simpler, potentially more accessible tool for assessing biological aging in clinical settings. The significant

differences in epigenetic age between the Down syndrome and healthy control groups highlight the profound impact of trisomy 21 on aging processes. These insights could inform better management strategies for age-related conditions in individuals with Down syndrome, such as cognitive deficits and the early onset of Alzheimer's disease.

Our results underscore the relevance of using epigenetic clocks to understand aging in Down syndrome. Continued research into the specific epigenetic changes associated with Down syndrome could lead to targeted interventions that might help mitigate the accelerated aging process, ultimately improving health outcomes. This study supports the growing recognition of epigenetic clocks as valuable tools in aging research and their potential application in improving clinical care for populations with unique aging trajectories like those seen in individuals with Down syndrome.

Epigenetic age in Alzheimer's disease: insights from EpiAgePublic and other epigenetic clocks

Differences in DNA methylation have been reported both in blood and brain tissue [44, 45] of Alzheimer's disease cases compared to healthy controls, especially in gene regions associated with AD pathology, including apolipoprotein E (ApoE) and the amyloid precursor protein (APP) but also in other genes [61–68]. However, the direction of the changes is not consistent across studies [44, 45, 69, 70]. Accordingly, the limited reports on epigenetic age in AD have also yielded inconsistent findings [69], and their utility in detecting AD pathology and cognitive status remains a matter of debate.

In the present study, we calculated the epigenetic age using bisulfite-targeted next-generation sequencing data and the EpiAgePublic model (including 3 CpGs: cg16867657, cg21572722, and cg24724428), which was trained on GSE55763, GSE157131, GSE40279, and GSE30870 (Table 1). The Epigenetic Age Acceleration (EAA) was determined for 26 control participants and 28 patients with AD. Our analysis revealed no significant difference between the two groups (parametric t-test, $p = 0.9875$) (Figure 6C).

We further analyzed whether the Mini-Mental State Examination (MMSE) score is associated with EAA separately in females and males. Interestingly, a strong negative correlation between MMSE T0 and epigenetic age acceleration was observed in males ($r = -0.8883$, $p = 0.0075$) (Figure 6D), while no significant correlation was observed in females (Figure 6D).

To examine EAA in AD patients in other studies, we analyzed the Epic Array dataset GSE144858

(Roubroeks et al., 2020), derived from the EU-funded AddNeuroMed Cohort, a large cross-European AD biomarker study using human blood DNA (Fransquet et al., 2021; Levine et al., 2015). This dataset included 96 control individuals, 111 with mild cognitive impairment (MCI), and 93 with Alzheimer's Disease (AD).

We applied the EpiAgePublic model and other epigenetic clocks (Figure 6E). Consistent with the bisulfite next-generation sequencing EpiAge assay, we did not observe any significant EAA differences between AD and controls, based on the Kruskal-Wallis test, since the data did not pass the normality test. Notably, EpiAgePublic was the only model showing significant age deceleration in MCI compared to controls (adjusted p -value = 0.0339).

These findings highlight the complexity of epigenetic changes in AD, as well as the need for further research to clarify the relationship between epigenetic age acceleration and cognitive dysfunction. The observed deceleration of epigenetic aging in AD patients may be explained in part by the older age of the cohort [71], but future research is necessary to support this hypothesis.

Limitations of array-based clocks; batch effects

The choice of analytical platform is critical for epigenetic studies, as highlighted by recent comparative analyses between the Illumina Infinium HumanMethylation450K and MethylationEPIC BeadChip arrays. These studies have shown substantial variability in methylation measurements, particularly at sites with low methylation variance, underscoring a fundamental limitation of array technologies that can be exacerbated by batch effects and dependency on pre-designed probes (Cheung, Burgers, Young, Cockell, & Reynard, 2020).

Batch effects are a significant concern in genomic technologies, especially in epigenome-wide association studies using Illumina Infinium Methylation BeadChips, where they can diminish experimental power and lead to false positives. Despite employing batch-effect removal tools like ComBat and Harman on various datasets, residual batch effects persist, particularly related to the processing day, glass slide, and array position, affecting thousands of probes and complicating the interpretation of epigenetic studies (Ross et al., 2022). Additionally, there is a growing recommendation in the field of epigenetic age calculation using array technologies that data should not be normalized nor corrected for batch effects. This advice stems from the need to preserve the integrity of longitudinal studies where epigenetic age is monitored

over time, a major application of this metric beyond single-timepoint clinical studies. The capability to consistently measure epigenetic age across multiple time points is crucial for assessing the effects of lifestyle changes and other interventions. However, the persistent batch effects and the necessity for normalization challenge the utility of array-based platforms for such longitudinal monitoring, underscoring the advantages of using targeted NGS for ongoing epigenetic age analysis.

In contrast, targeted next-generation sequencing (NGS) provides a more robust alternative. It offers higher resolution by sequencing individual DNA molecules, which allows for the detection of methylation profiles of the DNA molecules rather than just averages. NGS platforms are typically less prone to batch effects due in part to the sequencing process itself, which does not rely on the hybridization conditions that can vary in array-based platforms.

However, since next-generation sequencing captures profiles of each DNA molecule independently and takes into consideration the heterogeneity of DNA methylation profiles of single DNA molecules even in the same tissue, deep reads are required to have sufficient statistical power to accurately represent the diversity of DNA methylation profiles. Clinical tests need to be cost-effective and robust. Targeted amplification and Next-generation sequencing of thousands or even hundreds of regions with high depth and accuracy is highly technical, demanding and extremely costly and makes it unfeasible as a widespread robust clinical test. It is, therefore, imperative that a clinically widely used next-generation sequencing-based test be parsimonious and nimble and utilize the lowest number of regions amplified and sequenced. We, therefore, performed in this study a thorough analysis to determine whether it is possible to develop a clock that is based on the smallest number of DNA regions without compromising its correlation with age and its sensitivity to clinical states that affect aging. Our analysis revealed that the *ELOVL2* region on its own performs as well or is even superior to other clocks that use hundreds or even thousands of sites. We then developed a robust next-generation sequencing assay of a region that contains the CpGs included in the EpiAgePublic model as well as the other 10 CpGs.

We observed remarkably low coefficients of variation (CVs) across the first nine CpG sites, with CVs ranging from 0.28% to 6.7%, demonstrating NGS's robustness in accurate methylation profiling. Even for CpG sites with higher CVs, which ranged from 1% to 25% at sites of lower methylation levels, NGS shows a more consistent performance compared to arrays. Illustrated

in Figure 6A, these results highlight that NGS not only offers higher resolution and greater specificity but also exhibits reduced susceptibility to batch effects. This independence from hybridization conditions and pre-designed probes allows for a more comprehensive and flexible analysis, making NGS particularly valuable for precise and reproducible methylation profiling across diverse biological samples and time points. This technological superiority enhances our ability to monitor epigenetic changes with greater accuracy and reliability, which is crucial for longitudinal studies and the assessment of lifestyle or other interventions over time.

Advantages and limitations of simplifying epigenetic clocks with EpiAgePublic

While traditional epigenetic clocks rely on analyzing hundreds to thousands of CpG sites, EpiAgePublic achieves comparable predictive accuracy with significantly fewer sites. This reduction in complexity could lead to fewer errors associated with probe variability and hybridization inefficiencies inherent in array platforms. However, it is essential to acknowledge that the broader genomic coverage of traditional clocks may capture a more comprehensive epigenetic signature of aging. Therefore, while EpiAgePublic offers an efficient alternative, it should be seen as complementary to existing methods rather than a replacement.

The primary limitation of this study is the potential oversimplification of the aging process using only three CpG sites. Aging is a multifactorial process influenced by numerous genetic, environmental, and lifestyle factors. Therefore, further studies are needed to validate the effectiveness of EpiAgePublic across larger and more varied populations. Additionally, longitudinal studies would help in understanding how EpiAgePublic responds to changes over time and under different physiological or pathological conditions.

Moreover, the integration of EpiAgePublic with other biological markers of aging, such as telomere length, oxidative stress markers, and inflammatory cytokines, could provide a more holistic view of the aging process. This integrative approach could lead to the development of a multi-dimensional aging model that combines genetic, epigenetic, and biochemical indicators of age.

AUTHOR CONTRIBUTIONS

D.C. contributed to conceptualization, investigation, formal analysis, writing (original draft, review & editing), and visualization. S.D.C. performed DNA extractions, contributed to establishing the collaboration and the original study design, provided feedback on the project's

progression, and edited the manuscript. F.C. participated in the original concept of the study, revised the manuscript, and, together with M.G., recruited participants for the study (AD and DS individuals at Oasi Research Institute). M.G. collected the plasma samples and clinical data. A.C.C. contributed to the original concept of the investigation and engaged in discussions with Drs. Szyf, Caraci, and Do Carmo, as well as occasionally with D.C., revised the manuscript's early drafts. M.S. was involved in conceptualization, investigation, supervision, and writing (review & editing).

CONFLICTS OF INTEREST

DC is an inventor on a patent application (PCT/IB2020/055146) titled "EPIAGING; NOVEL ECOSYSTEM FOR MANAGING HEALTHY AGING". The patent application includes the tests disclosed in the paper. MS is a shareholder in epiMedTech Global (Singapore). All other authors declare no competing interest.

ETHICAL STATEMENT AND CONSENT

These samples were obtained following ethical guidelines, and the study protocols were approved by the Institutional Review Board (IRB) of the Oasi Research Institute—IRCCS, Troina, Italy (Protocol Numbers: 2016/1.0/122/CE-IRCCS-OASI and 2019/03/18/CE-IRCCS-OASI/18, respectively). Participants or their legal representatives gave written informed consent, as approved by the Ethics Committee of the Oasi Research Institute-IRCCS.

FUNDING

This research was supported by the Oasi Research Institute—IRCCS Troina, Italy. The additional funding was provided by EpiMedTech Global Ltd. The EpiAge measurement for the Alzheimer's cohort was founded by HKG epiTherapeutics.

REFERENCES

1. Fulop T, Larbi A, Dupuis G, Le Page A, Frost EH, Cohen AA, Witkowski JM, Franceschi C. Immunosenescence and Inflamm-Aging As Two Sides of the Same Coin: Friends or Foes? *Front Immunol*. 2018; 8:1960. <https://doi.org/10.3389/fimmu.2017.01960> PMID:29375577
2. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AM, Gentilini D, Vitale G, Collino S, Rezzi S, Castellani G, Capri M, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell*. 2012; 11:1132–4.
3. Spólnicka M, Pośpiech E, Peptłońska B, Zbieć-Piekarska R, Makowska Ż, Pięta A, Karłowska-Pik J, Ziemkiewicz B, Wężyk M, Gasperowicz P, Bednarczuk T, Barcikowska M, Żekanowski C, et al. DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups. *Int J Legal Med*. 2018; 132:1–11. <https://doi.org/10.1007/s00414-017-1636-0> PMID:28725932
4. Chao DL, Skowronska-Krawczyk D. ELOVL2: Not just a biomarker of aging. *Transl Med Aging*. 2020; 4:78–80 <https://doi.org/10.1016/j.tma.2020.06.004> PMID:33043173
5. Haerter JO, Lövkvist C, Dodd IB, Sneppen K. Collaboration between CpG sites is needed for stable somatic inheritance of DNA methylation states. *Nucleic Acids Res*. 2014; 42:2235–44. <https://doi.org/10.1093/nar/gkt1235> PMID:24288373
6. Landau DA, Clement K, Ziller MJ, Boyle P, Fan J, Gu H, Stevenson K, Sougnez C, Wang L, Li S, Kotliar D, Zhang W, Ghandi M, et al. Locally disordered methylation forms the basis of intratumor methylome variation in chronic lymphocytic leukemia. *Cancer Cell*. 2014; 26:813–25. <https://doi.org/10.1016/j.ccell.2014.10.012> PMID:25490447
7. Lapato DM, Roberson-Nay R, Kirkpatrick RM, Webb BT, York TP, Kinser PA. DNA methylation associated with postpartum depressive symptoms overlaps findings from a genome-wide association meta-analysis of depression. *Clin Epigenetics*. 2019; 11:169. <https://doi.org/10.1186/s13148-019-0769-z> PMID:31779682
8. Pauwels S, Symons L, Vanautgaerden EL, Ghosh M, Duca RC, Bekaert B, Freson K, Huybrechts I, Langie SA, Koppen G, Devlieger R, Godderis L. The Influence of the Duration of Breastfeeding on the Infant's Metabolic Epigenome. *Nutrients*. 2019; 11:1408. <https://doi.org/10.3390/nu11061408> PMID:31234503
9. Higgins-Chen AT, Thrush KL, Wang Y, Minter CJ, Kuo PL, Wang M, Niimi P, Sturm G, Lin J, Moore AZ, Bandinelli S, Vinkers CH, Vermetten E, et al. A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. *Nat Aging*. 2022; 2:644–61. <https://doi.org/10.1038/s43587-022-00248-2> PMID:36277076
10. Sun Z, Cunningham J, Slager S, Kocher JP. Base resolution methylome profiling: considerations in platform selection, data preprocessing and analysis. <https://doi.org/10.1111/accel.12005> PMID:23061750

- Epigenomics. 2015; 7:813–28.
<https://doi.org/10.2217/epi.15.21> PMID:[26366945](#)
11. Murata Y, Fujii A, Kanata S, Fujikawa S, Ikegame T, Nakachi Y, Zhao Z, Jinde S, Kasai K, Bundo M, Iwamoto K. Evaluation of the usefulness of saliva for DNA methylation analysis in cohort studies. *Neuropsychopharmacol Rep*. 2019; 39:301–05.
<https://doi.org/10.1002/npr2.12075> PMID:[31393092](#)
 12. Jin Z, Liu Y. DNA methylation in human diseases. *Genes Dis*. 2018; 5:1–8.
<https://doi.org/10.1016/j.gendis.2018.01.002> PMID:[30258928](#)
 13. Yoshizawa JM, Schafer CA, Schafer JJ, Farrell JJ, Paster BJ, Wong DT. Salivary biomarkers: toward future clinical and diagnostic utilities. *Clin Microbiol Rev*. 2013; 26:781–91.
<https://doi.org/10.1128/CMR.00021-13> PMID:[24092855](#)
 14. Sullivan AD, Bozack AK, Cardenas A, Comer JS, Bagner DM, Forehand R, Parent J. Parenting Practices May Buffer the Impact of Adversity on Epigenetic Age Acceleration Among Young Children With Developmental Delays. *Psychol Sci*. 2023; 34:1173–85.
<https://doi.org/10.1177/09567976231194221> PMID:[37733001](#)
 15. Moccia C, Popovic M, Isaevska E, Fiano V, Trevisan M, Rusconi F, Polidoro S, Richiardi L. Birthweight DNA methylation signatures in infant saliva. *Clin Epigenetics*. 2021; 13:57.
<https://doi.org/10.1186/s13148-021-01053-1> PMID:[33741061](#)
 16. Raffington L, Schnepel L, Mallard T, Fisher J, Vinnik L, Hollis-Hansen K, Notterman DA, Tucker-Drob EM, Mitchell C, Harden KP. Salivary Epigenetic Measures of Body Mass Index and Social Determinants of Health Across Childhood and Adolescence. *JAMA Pediatr*. 2023; 177:1047–54.
<https://doi.org/10.1001/jamapediatrics.2023.3017> PMID:[37669030](#)
 17. Han Y, Eipel M, Franzen J, Sakk V, Dethmers-Ausema B, Yndriago L, Izeta A, de Haan G, Geiger H, Wagner W. Epigenetic age-predictor for mice based on three CpG sites. *Elife*. 2018; 7:e37462.
<https://doi.org/10.7554/eLife.37462> PMID:[30142075](#)
 18. Poloni C, Szyf M, Cheishvili D, Tsoukas CM. Are the Healthy Vulnerable? Cytomegalovirus Seropositivity in Healthy Adults Is Associated With Accelerated Epigenetic Age and Immune Dysregulation. *J Infect Dis*. 2022; 225:443–52.
<https://doi.org/10.1093/infdis/jiab365> PMID:[34255838](#)
 19. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011; 27:1571–2.
<https://doi.org/10.1093/bioinformatics/btr167> PMID:[21493656](#)
 20. Tian Y, Morris TJ, Webster AP, Yang Z, Beck S, Feber A, Teschendorff AE. ChAMP: updated methylation analysis pipeline for Illumina BeadChips. *Bioinformatics*. 2017; 33:3982–4.
<https://doi.org/10.1093/bioinformatics/btx513> PMID:[28961746](#)
 21. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, Beck S. ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics*. 2014; 30:428–30.
<https://doi.org/10.1093/bioinformatics/btt684> PMID:[24336642](#)
 22. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*. 2013; 29:189–96.
<https://doi.org/10.1093/bioinformatics/bts680> PMID:[23175756](#)
 23. Gim JA. Integrative Approaches of DNA Methylation Patterns According to Age, Sex and Longitudinal Changes. *Curr Genomics*. 2023; 23:385–99.
<https://doi.org/10.2174/1389202924666221207100513> PMID:[37920553](#)
 24. Refn MR, Andersen MM, Kampmann ML, Tfelt-Hansen J, Sørensen E, Larsen MH, Morling N, Børsting C, Pereira V. Longitudinal changes and variation in human DNA methylation analysed with the Illumina MethylationEPIC BeadChip assay and their implications on forensic age prediction. *Sci Rep*. 2023; 13:21658.
<https://doi.org/10.1038/s41598-023-49064-7> PMID:[38066081](#)
 25. Johansson A, Enroth S, Gyllenstein U. Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. *PLoS One*. 2013; 8:e67378.
<https://doi.org/10.1371/journal.pone.0067378> PMID:[23826282](#)
 26. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol*. 2013; 14:R115.
<https://doi.org/10.1186/gb-2013-14-10-r115> PMID:[24138928](#)
 27. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*. 2013; 49:359–67.
<https://doi.org/10.1016/j.molcel.2012.10.016>

PMID:[23177740](https://pubmed.ncbi.nlm.nih.gov/23177740/)

28. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, Hou L, Baccarelli AA, Stewart JD, Li Y, Whitsel EA, Wilson JG, Reiner AP, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* (Albany NY). 2018; 10:573–91.
<https://doi.org/10.18632/aging.101414>
PMID:[29676998](https://pubmed.ncbi.nlm.nih.gov/29676998/)
29. Horvath S, Oshima J, Martin GM, Lu AT, Quach A, Cohen H, Felton S, Matsuyama M, Lowe D, Kabacik S, Wilson JG, Reiner AP, Maierhofer A, et al. Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging* (Albany NY). 2018; 10:1758–75.
<https://doi.org/10.18632/aging.101508>
PMID:[30048243](https://pubmed.ncbi.nlm.nih.gov/30048243/)
30. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, Hou L, Baccarelli AA, Li Y, Stewart JD, Whitsel EA, Assimes TL, Ferrucci L, Horvath S. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* (Albany NY). 2019; 11:303–27.
<https://doi.org/10.18632/aging.101684>
PMID:[30669119](https://pubmed.ncbi.nlm.nih.gov/30669119/)
31. Lu AT, Binder AM, Zhang J, Yan Q, Reiner AP, Cox SR, Corley J, Harris SE, Kuo PL, Moore AZ, Bandinelli S, Stewart JD, Wang C, et al. DNA methylation GrimAge version 2. *Aging* (Albany NY). 2022; 14:9484–549.
<https://doi.org/10.18632/aging.204434>
PMID:[36516495](https://pubmed.ncbi.nlm.nih.gov/36516495/)
32. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, Irizarry RA. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014; 30:1363–9.
<https://doi.org/10.1093/bioinformatics/btu049>
PMID:[24478339](https://pubmed.ncbi.nlm.nih.gov/24478339/)
33. Esteban-Cantos A, Rodríguez-Centeno J, Barruz P, Alejos B, Saiz-Medrano G, Nevado J, Martín A, Gayá F, De Miguel R, Bernardino JI, Montejano R, Mena-Garay B, Cadiñanos J, et al, and NEAT001/ANRS143 Study Group. Epigenetic age acceleration changes 2 years after antiretroviral therapy initiation in adults with HIV: a substudy of the NEAT001/ANRS143 randomised trial. *Lancet HIV*. 2021; 8:e197–205.
[https://doi.org/10.1016/S2352-3018\(21\)00006-0](https://doi.org/10.1016/S2352-3018(21)00006-0)
PMID:[33794182](https://pubmed.ncbi.nlm.nih.gov/33794182/)
34. Oriol-Tordera B, Esteve-Codina A, Berdasco M, Rosàs-Umbert M, Gonçalves E, Duran-Castells C, Català-Moll F, Llano A, Cedeño S, Puertas MC, Tolstrup M, Sogaard OS, Clotet B, et al. Epigenetic landscape in the kick-and-kill therapeutic vaccine BCN02 clinical trial is associated with antiretroviral treatment interruption (ATI) outcome. *EBioMedicine*. 2022; 78:103956.
<https://doi.org/10.1016/j.ebiom.2022.103956>
PMID:[35325780](https://pubmed.ncbi.nlm.nih.gov/35325780/)
35. Konigsberg IR, Barnes B, Campbell M, Davidson E, Zhen Y, Pallisard O, Boorgula MP, Cox C, Nandy D, Seal S, Crooks K, Sticca E, Harrison GF, et al. Host methylation predicts SARS-CoV-2 infection and clinical outcome. *Commun Med (Lond)*. 2021; 1:42.
<https://doi.org/10.1038/s43856-021-00042-y>
PMID:[35072167](https://pubmed.ncbi.nlm.nih.gov/35072167/)
36. Cao X, Li W, Wang T, Ran D, Davalos V, Planas-Serra L, Pujol A, Esteller M, Wang X, Yu H. Accelerated biological aging in COVID-19 patients. *Nat Commun*. 2022; 13:2135.
<https://doi.org/10.1038/s41467-022-29801-8>
PMID:[35440567](https://pubmed.ncbi.nlm.nih.gov/35440567/)
37. Zannas AS, Jia M, Hafner K, Baumert J, Wiechmann T, Pape JC, Arloth J, Ködel M, Martinelli S, Roitman M, Röh S, Haehle A, Emeny RT, et al. Epigenetic upregulation of FKBP5 by aging and stress contributes to NF-κB-driven inflammation and cardiovascular risk. *Proc Natl Acad Sci USA*. 2019; 116:11370–9.
<https://doi.org/10.1073/pnas.1816847116>
PMID:[31113877](https://pubmed.ncbi.nlm.nih.gov/31113877/)
38. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Röh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, Menke A, Lange JF, Brückl T, et al. Correction to: Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. *Genome Biol*. 2018; 19:61.
<https://doi.org/10.1186/s13059-018-1441-1>
PMID:[29792225](https://pubmed.ncbi.nlm.nih.gov/29792225/)
39. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Röh S, Ressler KJ, Nemeroff CB, Smith AK, Bradley B, Heim C, Menke A, Lange JF, Brückl T, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. *Genome Biol*. 2015; 16:266.
<https://doi.org/10.1186/s13059-015-0828-5>
PMID:[26673150](https://pubmed.ncbi.nlm.nih.gov/26673150/)
40. Ross JP, van Dijk S, Phang M, Skilton MR, Molloy PL, Oytam Y. Batch-effect detection, correction and characterisation in Illumina HumanMethylation450 and MethylationEPIC BeadChip array data. *Clin Epigenetics*. 2022; 14:58.
<https://doi.org/10.1186/s13148-022-01277-9>
PMID:[35488315](https://pubmed.ncbi.nlm.nih.gov/35488315/)
41. Teschendorff AE, Jones A, Fiegl H, Sargent A, Zhuang JJ, Kitchener HC, Widschwendter M. Epigenetic variability in cells of normal cytology is associated with the risk of future morphological transformation. *Genome Med*. 2012; 4:24.
<https://doi.org/10.1186/gm323>
PMID:[22453031](https://pubmed.ncbi.nlm.nih.gov/22453031/)

42. Teschendorff AE, Liu X, Caren H, Pollard SM, Beck S, Widschwendter M, Chen L. The dynamics of DNA methylation covariation patterns in carcinogenesis. *PLoS Comput Biol.* 2014; 10:e1003709. <https://doi.org/10.1371/journal.pcbi.1003709> PMID: [25010556](#)
43. Roubroeks JA, Smith AR, Smith RG, Pishva E, Ibrahim Z, Sattlecker M, Hannon EJ, Kłoszewska I, Mecocci P, Soininen H, Tsolaki M, Vellas B, Wahlund LO, et al. An epigenome-wide association study of Alzheimer's disease blood highlights robust DNA hypermethylation in the HOXB6 gene. *Neurobiol Aging.* 2020; 95:26–45. <https://doi.org/10.1016/j.neurobiolaging.2020.06.023> PMID: [32745807](#)
44. Fransquet PD, Lacaze P, Saffery R, Shah RC, Vryer R, Murray A, Woods RL, Ryan J. Accelerated Epigenetic Aging in Peripheral Blood does not Predict Dementia Risk. *Curr Alzheimer Res.* 2021; 18:443–51. <https://doi.org/10.2174/1567205018666210823100721> PMID: [34429046](#)
45. 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; 7:1198–211. <https://doi.org/10.18632/aging.100864> PMID: [26684672](#)
46. Guvatova ZG, Kobelyatskaya AA, Pudova EA, Tarasova IV, Kudryavtseva AV, Tkacheva ON, Strazhesko ID, Moskalev AA. Decelerated Epigenetic Aging in Long Livers. *Int J Mol Sci.* 2023; 24:16867. <https://doi.org/10.3390/ijms242316867> PMID: [38069189](#)
47. Middleton LY, Dou J, Fisher J, Heiss JA, Nguyen VK, Just AC, Faul J, Ware EB, Mitchell C, Colacino JA, M Bakulski K. Saliva cell type DNA methylation reference panel for epidemiological studies in children. *Epigenetics.* 2022; 17:161–77. <https://doi.org/10.1080/15592294.2021.1890874> PMID: [33588693](#)
48. Theda C, Hwang SH, Czajko A, Loke YJ, Leong P, Craig JM. Quantitation of the cellular content of saliva and buccal swab samples. *Sci Rep.* 2018; 8:6944. <https://doi.org/10.1038/s41598-018-25311-0> PMID: [29720614](#)
49. Wong YT, Tayeb MA, Stone TC, Lovat LB, Teschendorff AE, Iwaszow R, Craig JM. A comparison of epithelial cell content of oral samples estimated using cytology and DNA methylation. *Epigenetics.* 2022; 17:327–34. <https://doi.org/10.1080/15592294.2021.1950977> PMID: [34254878](#)
50. Hamano Y, Manabe S, Morimoto C, Fujimoto S, Tamaki K. Forensic age prediction for saliva samples using methylation-sensitive high resolution melting: exploratory application for cigarette butts. *Sci Rep.* 2017; 7:10444. <https://doi.org/10.1038/s41598-017-10752-w> PMID: [28874809](#)
51. Hao T, Guo J, Liu J, Wang J, Liu Z, Cheng X, Li J, Ren J, Li Z, Yan J, Zhang G. Predicting human age by detecting DNA methylation status in hair. *Electrophoresis.* 2021; 42:1255–61. <https://doi.org/10.1002/elps.202000349> PMID: [33629357](#)
52. Horvath S, Levine AJ. HIV-1 Infection Accelerates Age According to the Epigenetic Clock. *J Infect Dis.* 2015; 212:1563–73. <https://doi.org/10.1093/infdis/jiv277> PMID: [25969563](#)
53. Franzen J, Nüchtern S, Tharmapalan V, Vieri M, Nikolić M, Han Y, Balfanz P, Marx N, Dreher M, Brümmendorf TH, Dahl E, Beier F, Wagner W. Epigenetic Clocks Are Not Accelerated in COVID-19 Patients. *Int J Mol Sci.* 2021; 22:9306. <https://doi.org/10.3390/ijms22179306> PMID: [34502212](#)
54. Corley MJ, Pang AP, Dody K, Mudd PA, Patterson BK, Seethamraju H, Bram Y, Peluso MJ, Torres L, Iyer NS, Premeaux TA, Yeung ST, Chandar V, et al. Genome-wide DNA methylation profiling of peripheral blood reveals an epigenetic signature associated with severe COVID-19. *J Leukoc Biol.* 2021; 110:21–6. <https://doi.org/10.1002/JLB.5HI0720-466R> PMID: [33464637](#)
55. Protsenko E, Wolkowitz OM, Yaffe K. Associations of stress and stress-related psychiatric disorders with GrimAge acceleration: review and suggestions for future work. *Transl Psychiatry.* 2023; 13:142. <https://doi.org/10.1038/s41398-023-02360-2> PMID: [37130894](#)
56. Epel ES, Crosswell AD, Mayer SE, Prather AA, Slavich GM, Puterman E, Mendes WB. More than a feeling: A unified view of stress measurement for population science. *Front Neuroendocrinol.* 2018; 49:146–69. <https://doi.org/10.1016/j.yfrne.2018.03.001> PMID: [29551356](#)
57. Poon CY, Knight BG. Emotional reactivity to network stress in middle and late adulthood: the role of childhood parental emotional abuse and support. *Gerontologist.* 2012; 52:782–91. <https://doi.org/10.1093/geront/gns009> PMID: [22419249](#)
58. Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT. The stress process. *J Health Soc Behav.* 1981;

22:337–56.

PMID:[7320473](#)

59. Poganik JR, Zhang B, Baht GS, Tyshkovskiy A, Deik A, Kerepesi C, Yim SH, Lu AT, Haghani A, Gong T, Hedman AM, Andolf E, Pershagen G, et al. Biological age is increased by stress and restored upon recovery. *Cell Metab.* 2023; 35:807–20.e5.
<https://doi.org/10.1016/j.cmet.2023.03.015>
PMID:[37086720](#)
60. Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio AM, Giuliani C, Tung S, Vinters HV, Franceschi C. Accelerated epigenetic aging in Down syndrome. *Aging Cell.* 2015; 14:491–5.
<https://doi.org/10.1111/accel.12325> PMID:[25678027](#)
61. Foraker J, Millard SP, Leong L, Thomson Z, Chen S, Keene CD, Bekris LM, Yu CE. The APOE Gene is Differentially Methylated in Alzheimer's Disease. *J Alzheimers Dis.* 2015; 48:745–55.
<https://doi.org/10.3233/JAD-143060> PMID:[26402071](#)
62. Poon CH, Tse LS, Lim LW. DNA methylation in the pathology of Alzheimer's disease: from gene to cognition. *Ann N Y Acad Sci.* 2020; 1475:15–33.
<https://doi.org/10.1111/nyas.14373> PMID:[32491215](#)
63. Wei X, Zhang L, Zeng Y. DNA methylation in Alzheimer's disease: In brain and peripheral blood. *Mech Ageing Dev.* 2020; 191:111319.
<https://doi.org/10.1016/j.mad.2020.111319>
PMID:[32721406](#)
64. Wen KX, Milić J, El-Khodori B, Dhana K, Nano J, Pulido T, Kraja B, Zaciragic A, Bramer WM, Troup J, Chowdhury R, Ikram MA, Dehghan A, et al. The Role of DNA Methylation and Histone Modifications in Neurodegenerative Diseases: A Systematic Review. *PLoS One.* 2016; 11:e0167201.
<https://doi.org/10.1371/journal.pone.0167201>
PMID:[27973581](#)
65. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A Systematic Review and Meta-analysis of Environmental, Lifestyle, and Health Factors Associated With DNA Methylation Age. *J Gerontol A Biol Sci Med Sci.* 2020; 75:481–94.
<https://doi.org/10.1093/gerona/glz099>
PMID:[31001624](#)
66. De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, Eaton ML, Keenan BT, Ernst J, McCabe C, Tang A, Raj T, Replogle J, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. *Nat Neurosci.* 2014; 17:1156–63.
<https://doi.org/10.1038/nn.3786> PMID:[25129075](#)
67. Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. *Nat Neurosci.* 2014; 17:1138–40.
<https://doi.org/10.1038/nn.3792> PMID:[25157507](#)
68. Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, Troakes C, Al-Sarraj S, Burrage J, Macdonald R, Condliffe D, Harries LW, Katsel P, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci.* 2014; 17:1164–70.
<https://doi.org/10.1038/nn.3782> PMID:[25129077](#)
69. Milicic L, Porter T, Vacher M, Laws SM. Utility of DNA Methylation as a Biomarker in Aging and Alzheimer's Disease. *J Alzheimers Dis Rep.* 2023; 7:475–503.
<https://doi.org/10.3233/ADR-220109> PMID:[37313495](#)
70. Pellegrini C, Pirazzini C, Sala C, Sambati L, Yusipov I, Kalyakulina A, Ravaioli F, Kwiatkowska KM, Durso DF, Ivanchenko M, Monti D, Lodi R, Franceschi C, et al. A Meta-Analysis of Brain DNA Methylation Across Sex, Age, and Alzheimer's Disease Points for Accelerated Epigenetic Aging in Neurodegeneration. *Front Aging Neurosci.* 2021; 13:639428.
<https://doi.org/10.3389/fnagi.2021.639428>
PMID:[33790779](#)
71. El Khoury LY, Gorrie-Stone T, Smart M, Hughes A, Bao Y, Andrayas A, Burrage J, Hannon E, Kumari M, Mill J, Schalkwyk LC. Systematic underestimation of the epigenetic clock and age acceleration in older subjects. *Genome Biol.* 2019; 20:283.
<https://doi.org/10.1186/s13059-019-1810-4>
PMID:[31847916](#)

SUPPLEMENTARY MATERIALS

Supplementary Tables

Supplementary Table 1. Regression analysis results for chronological age across various epigenetic clocks and CpG sites including confounder adjustments.

Model description	R ²	Adj. R ²	F-Stat	p(F-stat)	Age Coef	Std Err	t-Value	p-Value	95% CI Lower	95% CI Upper	AIC	BIC	Observations
cg16867657 (ELOVL2 site)	0.757	0.757	1801	<0.0001	0.0047	4.50E-05	105.465	<0.0001	0.005	0.005	-17790-17730		4625
cg21572722 (ELOVL2 site)	0.702	0.702	1362	<0.0001	0.0025	3.04E-05	82.951	<0.0001	0.002	0.003	-21430-21370		4625
cg24724428 (ELOVL2 site)	0.626	0.625	965.1	<0.0001	0.0039	4.87E-05	79.2	<0.0001	0.004	0.004	-17060-17010		4625
EpiAgePublicBlood	0.779	0.778	2030	<0.0001	0.7612	0.007	111.887	<0.0001	0.748	0.775	28630 28690		4625
DNAmAge (Horvath)	0.842	0.841	3067	<0.0001	0.78	0.005	145.819	<0.0001	0.77	0.791	26410 26470		4625
DNAmAge Hannum	0.883	0.883	4357	<0.0001	0.8554	0.006	153.678	<0.0001	0.845	0.866	26780 26830		4625
DNAmPhenoAge	0.787	0.786	2127	<0.0001	0.8966	0.008	116.929	<0.0001	0.882	0.912	29740 29800		4625
DNAmAgeSkinBloodClock	0.898	0.898	5078	<0.0001	0.8855	0.005	184.542	<0.0001	0.876	0.895	25400 25460		4625
DNA GrimAge v1	0.806	0.806	2395	<0.0001	0.689	0.006	121.39	<0.0001	0.678	0.7	26960 27010		4625
DNA GrimAge v2	0.765	0.765	1882	<0.0001	0.6501	0.006	108.32	<0.0001	0.638	0.662	27470 27530		4625

Supplementary Table 2. Summary table for replicates.

ID	Mean	SD	CV (%)	95% CI	Minimum	Maximum	Median	25th Percentile	75th Percentile
1	67.3	1.34	1.99%	[65.17, 69.43]	66.09	68.72	67.2	66.13	68.58
2	90.67	0.99	1.09%	[89.10, 92.24]	89.38	91.77	90.77	89.7	91.55
3	63.47	1.56	2.46%	[60.99, 65.95]	61.65	65.33	63.44	61.97	64.99
4	80.19	1.04	1.29%	[78.54, 81.84]	79.51	81.72	79.76	79.54	81.26
5	72.89	1.6	2.19%	[70.35, 75.43]	71.5	74.44	72.8	71.5	74.35
6	72.61	4.43	6.10%	[65.56, 79.67]	66.57	76.47	73.71	67.96	76.18
7	83.37	1.25	1.50%	[81.38, 85.36]	82.04	84.84	83.3	82.2	84.61
8	86.78	2.04	2.35%	[83.54, 90.03]	84.93	89.21	86.49	85.02	88.83
9	66.6	0.98	1.48%	[65.03, 68.16]	65.58	67.73	66.54	65.68	67.57
10	74.35	2.6	3.49%	[70.22, 78.48]	71.28	77.41	74.36	71.84	76.85
11	78	1.01	1.29%	[76.40, 79.60]	76.91	79.11	78	77.04	78.97
12	79.25	1.65	2.08%	[76.62, 81.87]	77.56	81.29	79.07	77.75	80.92
13	87.51	0.85	0.98%	[86.15, 88.87]	86.59	88.64	87.41	86.76	88.37
14	62.81	1.24	1.98%	[60.83, 64.79]	61.1	63.94	63.1	61.51	63.82
15	81.45	1.52	1.86%	[79.03, 83.86]	79.65	83.28	81.43	80	82.91
16	82.22	1.88	2.29%	[79.23, 85.21]	79.59	84.02	82.62	80.29	83.74
17	82.62	1.47	1.78%	[80.28, 84.96]	81.62	84.81	82.03	81.72	84.11
18	70.94	1.8	2.54%	[68.07, 73.81]	69.39	73.07	70.64	69.42	72.75
19	77.23	1.14	1.48%	[75.41, 79.05]	75.67	78.25	77.5	76.03	78.16
20	87.44	2.78	3.18%	[83.02, 91.86]	84.12	90.63	87.5	84.72	90.09
21	72.08	0.79	1.09%	[70.83, 73.33]	71.4	73.21	71.85	71.51	72.87
22	77.4	0.67	0.87%	[76.33, 78.47]	76.86	78.35	77.19	76.9	78.11
23	88.47	2.45	2.77%	[84.57, 92.37]	86.01	91.31	88.28	86.23	90.9
24	66.71	1.58	2.37%	[64.19, 69.22]	65.26	68.7	66.44	65.36	68.34
25	70.23	2.79	3.98%	[65.78, 74.67]	67.92	73.78	69.61	67.96	73.12
26	68.85	1.19	1.73%	[66.95, 70.74]	67.78	70.23	68.69	67.82	70.03

27	78.66	2.4	3.05%	[74.85, 82.48]	75.61	81.41	78.81	76.31	80.86
28	80.28	0.76	0.94%	[79.08, 81.49]	79.59	81.14	80.2	79.62	81.03
29	64.02	1.97	3.08%	[60.88, 67.15]	62.06	66.69	63.66	62.35	66.05
30	76.23	1.19	1.56%	[74.35, 78.12]	75.15	77.8	75.99	75.24	77.47
31	65.66	1.76	2.68%	[62.86, 68.46]	64.11	67.2	65.66	64.12	67.19
32	70.15	1.95	2.78%	[67.05, 73.25]	68.06	72.41	70.07	68.32	72.07
33	69.16	0.82	1.19%	[67.85, 70.47]	67.97	69.79	69.44	68.29	69.75
34	66.46	2.03	3.05%	[63.24, 69.69]	64.71	69.29	65.93	64.87	68.59
35	68.06	0.72	1.06%	[66.92, 69.20]	67.25	68.72	68.13	67.36	68.69
36	81.62	1.52	1.86%	[79.20, 84.03]	80.2	83.75	81.26	80.42	83.18
37	74.34	1.66	2.24%	[71.69, 76.98]	72.53	75.93	74.44	72.73	75.84
38	66.67	1.71	2.57%	[63.95, 69.39]	64.57	68.56	66.77	64.96	68.27
39	72.9	0.42	0.57%	[72.24, 73.56]	72.47	73.41	72.85	72.52	73.32
40	68.86	2.32	3.37%	[65.16, 72.55]	66.68	72.11	68.32	67	71.26
41	67.18	1.25	1.86%	[65.19, 69.17]	65.78	68.63	67.15	65.98	68.4