Applicability of epigenetic age models to next-generation methylation arrays

Leonardo D. Garma[1], Miguel Quintela-Fandino[1*]

[1]Breast Cancer Clinical Research Unit, Centro Nacional de Investigaciones Oncológicas–CNIO

Madrid, Spain. Melchor Fernández Almagro, 3, 28029 Madrid

[*]Corresponding author, mquintela@cnio.es

Abstract

Background

Epigenetic clocks based on DNA methylation data are routinely used to obtain surrogate measures of biological age and estimate epigenetic age acceleration rates. These tools are mathematical models that rely on the methylation state of specific sets of CpG islands quantified using microarrays. The set of CpG islands probed in the microarrays differed between the models. Thus, as new methylation microarrays are developed and older models are discontinued, existing epigenetic clocks might become obsolete. Here, we explored the effects of the changes introduced in the new DNA methylation array from Illumina (EPICv2) on existing epigenetic clocks.

Methods

We compiled a whole-blood DNA methylation dataset of 10835 samples to test the performance of four epigenetic clocks on the probe set of the EPICv2 array. We then used the same data to train a new epigenetic age prediction model compatible across the 450k, EPICv1 and EPICv2 microarrays. We compiled a validation dataset of 2095 samples to compare our model with a state-of-the-art epigenetic clock. Using two datasets with repeated samples from the same subjects, we computed an estimate of the contribution of technical noise and intra-subject variation to the variation of epigenetic age predictions from each of the models tested. We used a dataset of cancer survivors who had undergone different types of therapy , a dataset of breast cancer patients and controls, and a dataset from an exercise-based interventional study to test the ability of our model to detect alterations in epigenetic age acceleration.

Results

We found that the results of the four epigenetic clocks tested are significantly distorted by the absence of specific probes in the EPICv2 microarray, causing an average difference of up to 25 years. We developed an epigenetic age prediction model compatible with the 450k, EPICv1 and EPICv2 microarrays. Our model produced highly accurate chronological age predictions that were comparable to those of a state-of-the-art epiclock. We obtained estimates for the variation of epigenetic age acceleration on normal, non-pathological populations associated with each of the models tested. These parameters provide thresholds to evaluate the relevance of epigenetic age alterations. In all cases, the estimated technical noise and intra-subject variability were smaller than the population-based epigenetic age prediction variability. Finally, we used our new models to reproduce previous results showing increased

epigenetic age acceleration in cancer patients and in survivors who had been treated with radiation therapy, as well as a lack of changes as a result of exercise-based interventions.

Conclusion

Our work demonstrated that existing epigenetic clocks need to be updated to be applicable to data generated with the new EPICv2 microarray, which has phased out the 450k and EPICv1 models. To overcome this technical hurdle, we developed a new model that translates the capabilities of state-of-the-art epigenetic clocks to the new EPICv2 platform and is cross-compatible with the 450k and EPICv1 microarrays. Our characterization of the variation of epigenetic age predictions provides useful metrics to contextualize the biological relevance of epigenetic age alterations. The analysis of data from subjects influenced by radiation, cancer and exercise-based interventions shows that despite being good predictors of chronological age, neither a pathological state like breast cancer, a hazardous environmental factor (radiation) or exercise (a beneficial intervention) caused significant changes in the values of the “epigenetic age” determined by these first-generation models.

Keywords

Epigenetics, epigenetic clock, DNA methylation

Introduction

In the rapidly evolving field of epigenomics, the development of epigenetic clocks has revolutionized our ability to gauge biological aging through DNA methylation patterns. Changes in the methylation state of CpG sites have proven to be highly correlated with chronological age[1,2]. Thus, DNA methylation patterns are being increasingly used to gain insights into aging and associated pathologies[2] and have led to the development of several epigenetic clocks, which are predictive models that estimate the age of subjects based on methylation markers[3]. In these epigenetic clocks, the model prediction is interpreted as the "biological" or "epigenetic" age. Subjects whose biological age is greater than their chronological age are considered to have epigenetic age acceleration[4] (EAA). EAAs have been linked to increased risk for various health issues and mortality, independent of traditional risk factors[5–8]. Epigenetic clocks have thus not only underscored the potential of DNA methylation as a biomarker of aging[9] but also represent a valuable tool to gain insight into the complexities of multiple pathologies[4]. In cancer, EAAs have been linked to an increased risk of developing breast and colon cancer[10,11]. Moreover, breast tissue from breast cancer patients has been shown to exhibit EAA[12]; likewise, cancer survivors do present an accelerated epigenetic age compared to noncancer patients, with an acceleration rate dependent on the treatment intensity received[13].

From the initial model of Bockland *et al*. published in 2011[1], the field quickly evolved toward more robust models, such as the Hannum[14] and the Horvath[15] clocks, and later toward "second generation" approaches, which included features beyond DNA methylation with the promise of providing a "biological age" prediction that could better reflect age-related health status[16–18]. The field continues to

evolve, with the implementation of new predictive models[19–21] and the development of species-[22,23] and tissue-specific[24–27] epigenetic clocks. Most recently, Bernabeu *et al*. conducted a large-scale study to refine the predictive ability of both first (methylation-based) and second (including other features) generation epigenetic clocks, which yielded a significant improvement over existing models in the ability to predict chronological age from DNA methylation data across multiple cohorts with subjects of all ages[28].

The evolution of epigenetic clocks has run in parallel with that of methylation microarrays, whose coverage of DNA methylation sites increased from 27578 in 2008[29,30] to 866836 in 2015[31]. In 2023, Illumina released the newest iteration of its series of methylation arrays, EPICv2, which targets 936866 methylation sites[32]. Although the vast majority of probes have been conserved across different microarray models over the years[32–34], some of them have been lost. The probes on the latest (and currently the only commercially available) methylation microarray version cover more than 80% of the probes used by the most popular epigenetic clocks (PhenoAge[16], Horvath (2013)[15], Horvath (2018)[35], Hannum[14], DunedinPACE[17], etc.), but they do not offer complete coverage for any of them[36]. Given that earlier microarray models are now discontinued while there is still a need to use epigenetic clocks, it is necessary to evaluate whether the existing models can offer reliable results when using EPICv2 data.

Recent studies have shown that DNA methylation is a dynamic molecular feature, exhibiting changes which can affect epigenetic age predictions even within a 24-hour timeframe[37–39]. Apsley et al. demonstrated that methylation values of probes in the EPICv1 array do exhibit significant changes using repeated sampling on the same subjects on 4 different timepoints within a 5-hour interval. The effect was observed on multiple probes used by common epigenetic clocks, although the changes in the epigenetic age predictions were not directly reported[38]. More recently, Koncevičius et al. reported oscillations in the epigenetic age predictions produced by 17 epigenetic age models on 48 samples from the same subject obtained during a 72-hour period, which they attributed to circadian oscillations[37]. Therefore, it is necessary to quantify the range of spontaneous variation in epigenetic age predictions to be able to distinguish truly biologically relevant alterations.

In this study, we examined the performance of existing epigenetic clocks using the CpG probes available in the new EPICv2 methylation array in human blood samples. We observed that the biological age predictions of these models were distorted due to missing probes in the EPICv2 array. To overcome this hurdle, we trained a model that can be applied to data obtained with 450k, EPICv1 or EPICv2 chips and whose performance is at least on par with that of state-of-the-art models. We then estimated the variation of EAA in normal, non-pathological populations for our model and for four existing epiclocks. We used two datasets with repeated samples from the same subjects to obtain estimates of the contributions of technical and biological, intra-subject noise to the variation observed in the EAA distributions. The extent of these variations can set a threshold to distinguish normal from biologically relevant EAA changes. Finally, our model reproduced previously reported EAA differences attributed to radiation therapy and to sporadic breast cancer, and the negative results from an interventional study using physical exercise.

Results

DNA methylation datasets

We explored public repositories to assemble a large and diverse DNA methylation dataset to examine the performance of existing epigenetic clocks and to train and test new models. We used public data to generate two separate datasets: 1) a main dataset for evaluating existing models and training new ones and 2) a validation dataset to test the generalizability of new models.

The main dataset included whole blood DNA methylation data of human subjects from 24 previous studies (Supplementary Table 1). We filtered the data to retain exclusively subjects labeled as controls, with reported age and sex values, resulting in a total of 11825 individuals. To be able to split the data in a stratified fashion into training and test sets, we kept only the samples with an age value present at least twice in the dataset, leaving a total of 10835 subjects (Figure 1A). The subjects in this main dataset had ages between 8 and 96 years, with a mean of 47.68 years and an almost even distribution of sexes (5456 females and 5379 males) (Figure 1B).

The validation dataset was composed of data from 6 additional studies (Supplementary Table 1). We selected only control subjects with reported age and sex values, which resulted in a total of 2095 individuals. The age range in this validation set was between 14 and 94 years, and the sex distribution was slightly biased, with 828 females and 1267 males (Figure 1C-D).

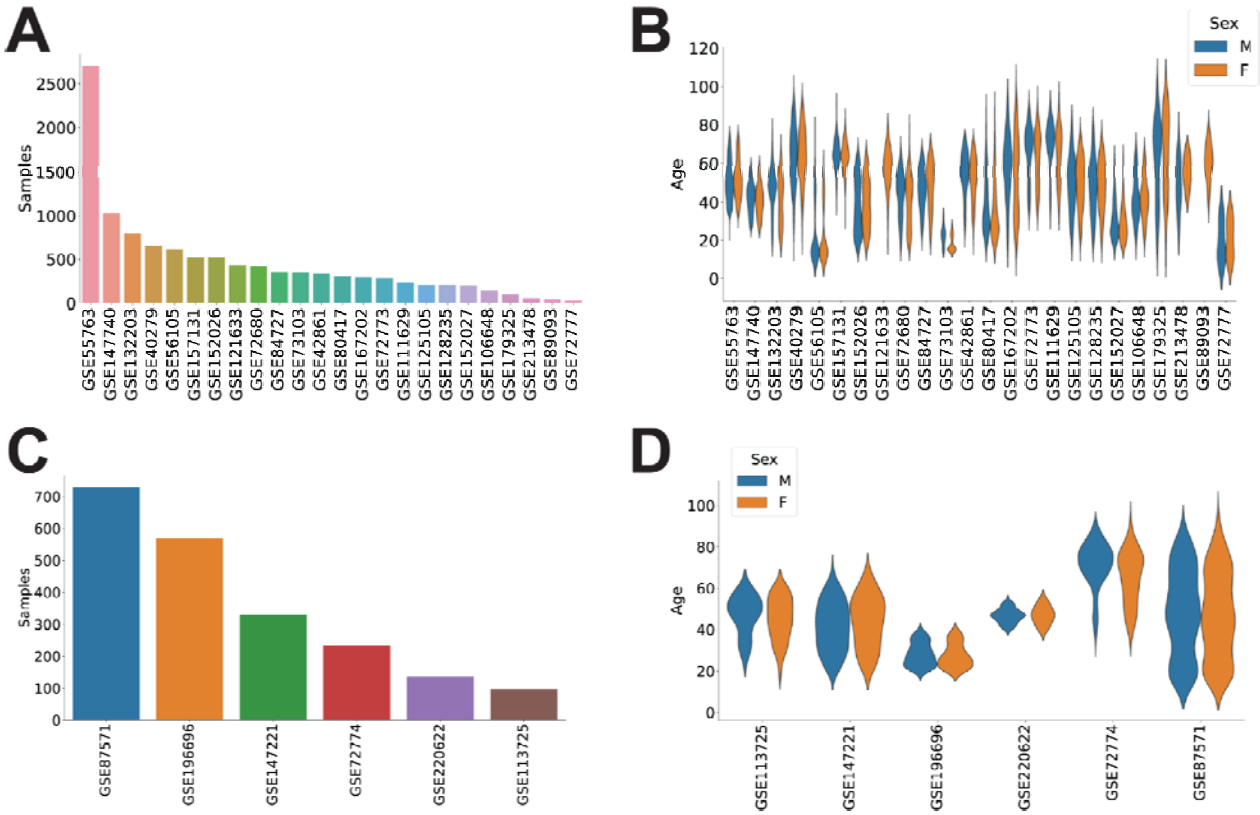


Figure 1 - Public DNA methylation datasets compiled for the study. **A**. Number of subjects present in each of the datasets included in the main set. **B.** Distribution of ages and sex in the datasets included in the main set. **C** and **D**. Distribution of the number of subjects and their age and sex in the datasets used for the validation set.

Existing epigenetic clocks generate distorted results from EPICv2 data

Although more than 80% of the probes used by existing epiclocks are conserved in the EPICv2 methylation array[36], applying these models to data limited to the probes present in the EPICv2 array results in differences in epigenetic age predictions. We applied four different models (Horvath[15], Hannum[14], PhenoAge[16] and the recent chronological age prediction model (cAge) from Bernabeu *et al.*[28]) to the main dataset of 10835 individuals, using either all the required probes (*complete* models) or only those present in the new EPICv2 array (*truncated* models). The proportion of probes relevant for these models in the EPICv2 ranged from 84.5% (Hannum) to 97.1% (cAge).

We observed a high degree of correlation between chronological age and the predicted biological age in all cases, with cAge obtaining the highest value (r=0.979). After the models were truncated to the EPICv2 probeset, they maintained this correlation, which was still strong (r>0.88) in all cases (Figure 2A).

However, limiting the number of probes distorted the epigenetic age predictions in all the models. In all four epiclocks, we observed a significant difference ($p<0.001$, one sample t-test) in the biological age predicted by the complete and the truncated models (Figure 2B). This effect was greater on the Hannum clock model, with an average difference of 25.45 years. The larger models PhenoAge (513 parameters) and cAge (4058 parameters) had average differences of -1.11 and -1.67 years respectively, whereas the average difference on the Horvath clock was just -0.82 years.

To characterize the performance of these models in predicting chronological age, we computed the Pearson correlation coefficient between the predictions and chronological age (r), the mean squared error (MSE) and median absolute error (MAE) for each of them. We computed the mean (μ_{EAA}) and standard deviations (σ_{EAA}) of the EAA distributions associated to each epigenetic (Table 1).

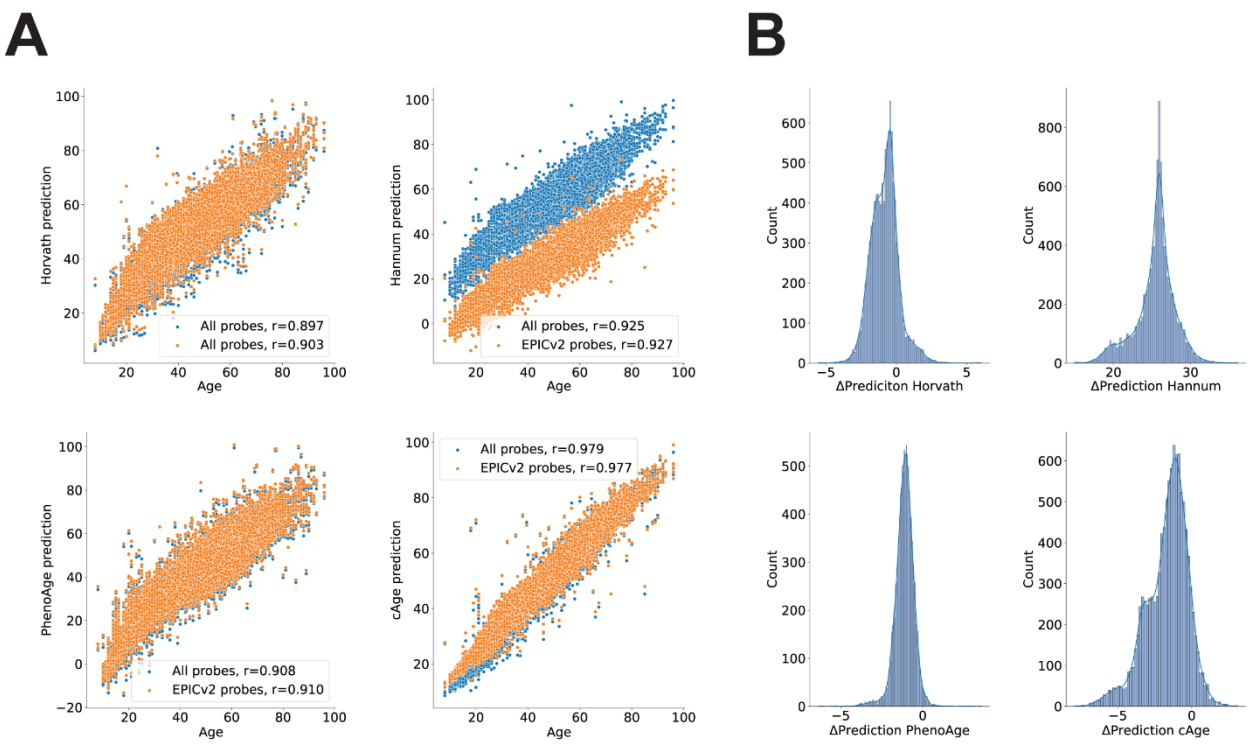


Figure 2 – Epigenetic clock age prediction drifts caused by the loss of probes in the EPICv2 microarray. **A**. Epigenetic age predictions by four existing epigenetic clocks in 10835 subjects using either their complete sets of CpG probes (blue) or only those available in the EPICv2 array (orange). **B**. Distribution of differences between the values predicted using the full sets of probes or only those present in the EPICv2 array (drift) for the same four epigenetic clocks.

A new epigenetic clock compatible with the EPICv2 array

The changes in the CpG probes present in the EPICv2 array produce significant alterations in the results of existing epigenetic clocks. To overcome this limitation and to produce robust biological age predictions using EPICv2 data, we trained a new epiclock model using CpG probes common to the 450k, EPICv1 and EPICv2 methylation arrays. Based on the epigenome-wide association study (EWAS) from Bernabeu *et al.*[28], there are 42728 probes with a significant (p-adjusted<0.05) linear association with chronological age that are present in the three different methylation arrays. Similarly, according to the EWAS, there are 63324 CpG probes with significant (p-adjusted<0.05) quadratic associations with age which are present in the three microarray models (Figure 3A).

Following the strategy employed by Bernabeu *et al.*[28], we ranked the probes by their p-value on the EWAS and preselected the first 10000 with linear and the first 300 with quadratic associations to chronological age to train an epiclock using elastic net regression[40] (Supplementary Table 2).

We split the main dataset with 10835 subjects into training and test sets using a 70/30 split, resulting in 7259 and 3576 samples, respectively, and trained a general age-prediction model. The model used an L1 ratio of 0.5[28], and the alpha value was optimized to 0.001 through 5-fold cross-validation. This general

model obtained a low prediction error (mean squared error (MSE)=3.14 years) and a high correlation between the predicted and chronological ages (r=0.982) in the test set (Figure 3B).

Previous works have suggested that the relationship between methylation state and age is nonlinear[14,28,41–44]. Therefore, we decided to stratify our training data into three age groups and train separate predictors for "young", "middle-aged", and "old" subjects. The predictors were intended to work on nonoverlapping age groups, but we did use overlapping age ranges during training to limit the inconsistencies between models assigned to contiguous groups. We produced a combined model in which the general model assigned a sample to one of the three age groups, and then the corresponding age-stratified model generated the age prediction (Figure 3C). This combined model reduced the age prediction error significantly ($p<0.05$, one-sided Wilcoxon rank-sum test) with respect to the general model (Figure 3B,D), lowering the MSE to 3.06 years in the test set and increasing the correlation between the predicted and chronological ages to 0.983. Although the combined model reduced the error in all age groups, the difference was statistically significant ($p<0.05$, one-sided Wilcoxon rank-sum test) only in the younger subjects (Figure 3E). Weights for the general and age-specific models are provided in Supplementary Table 3. The overlap of EPICv2 probes used by our general and combined models and by the rest of the epiclocks examined is presented schematically in Supplementary Figure 1A and B. To underscore the numerical differences between models, we present the coefficients used for the overlapping probes in our models, cAge and the Horvath clock in Supplementary Figure 1C.

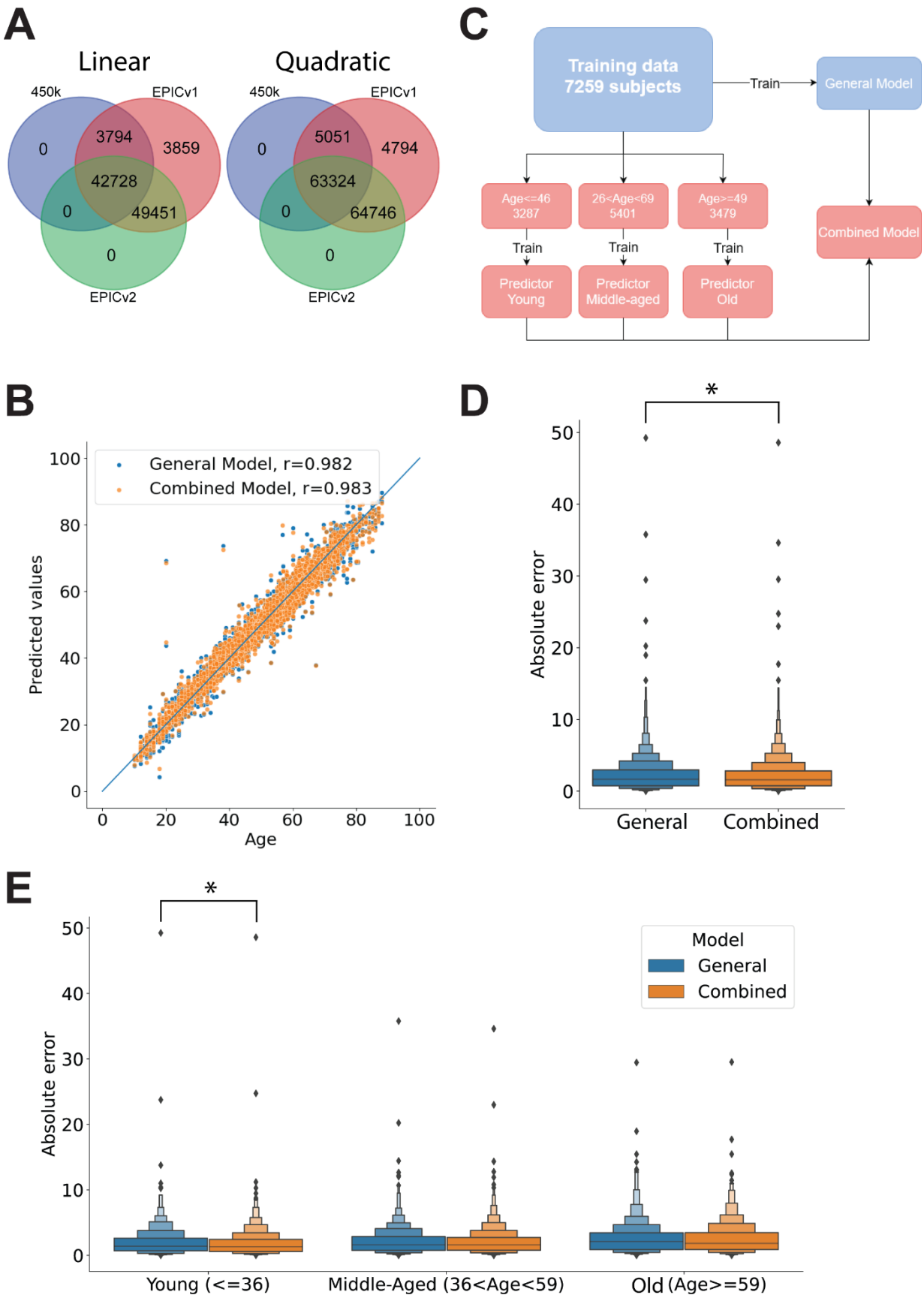

A
Linear
450k
EPICv1
EPICv2
0
3794
3859
42728
0
49451
0
Quadratic
450k
EPICv1
EPICv2
0
5051
4794
63324
0
64746
0
C
Training data
7259 subjects
Train
General Model
Age<=46
3287
26<Age<69
5401
Age>=49
3479
Train
Train
Train
Predictor
Young
Predictor
Middle-aged
Predictor
Old
Combined Model
B
General Model, r=0.982
Combined Model, r=0.983
Predicted values
Age
D
*
Absolute error
General
Combined
E
*
Absolute error
Model
General
Combined
Young (<=36)
Middle-Aged (36<Age<59)
Old (Age>=59)

Figure 3 – Results of the new epigenetic clocks in the test dataset. A. Methylation probes with significant linear (left) and quadratic (right) associations with chronological age in the EWAS from Bernabeu *et al.* present in the 450k, EPICv1 and EPICv2 arrays. **B**. Prediction results on the test set by the General and Combined models. The blue line indicates the 1:1 correspondence. **C**. Schematic overview of the training of the different age prediction models. **D**. Absolute error of the ages predicted in the test set by the general and combined models. **E**. Absolute error from the predictions on the test set by each model, broken down by age group. *Indicates significant differences ($p<0.05$, one-sided Wilcoxon rank-sum test) between groups.

The performance of the new clock is on par with that of a state-of-the-art model

To examine the performance of our model, we compared it against cAge, a refined model that has shown significant improvement over previous epigenetic clocks[28]. We tested the complete cAge and our general and combined models on a validation dataset with 2095 samples from 6 studies. We included all the probes used for cAge and those required by our models and predicted the age of each sample based on the methylation data. In all cases, the predicted ages were strongly correlated with the chronological ages (Figure 4A).

Our general model performed as well as cAge, with no significant difference ($p>0.05$, one-sided Wilcoxon rank-sum test) in the distribution of errors between the two, whereas the error from the predictions of our combined (age-segregated) model was significantly lower ($p>0.05$, one-sided Wilcoxon test) (Figure 4B). By breaking down the data by age group, we could observe differences in the prediction error. In the young (<=36) and old (>=59) groups, the error from the combined model was significantly lower ($p>0.05$, one-sided Wilcoxon test) than that of the cAge and the general models. In the group of middle-aged subjects, the error of the general model was significantly greater ($p>0.05$, one-sided Wilcoxon test) than that of the other two models (Figure 4C).

To be able to perform a more general comparison, we computed again the values of r, MSE, MAE, μ_{EAA} and σ_{EAA} for our models and of the other four epigenetic clocks using the validation dataset (Table 1, Supplementary Figure 2). These metrics were largely compatible with those obtained on the main dataset, with the exception of cAge, which had a large decrease in MAE, MSE, μ_{EAA} and σ_{EAA}.

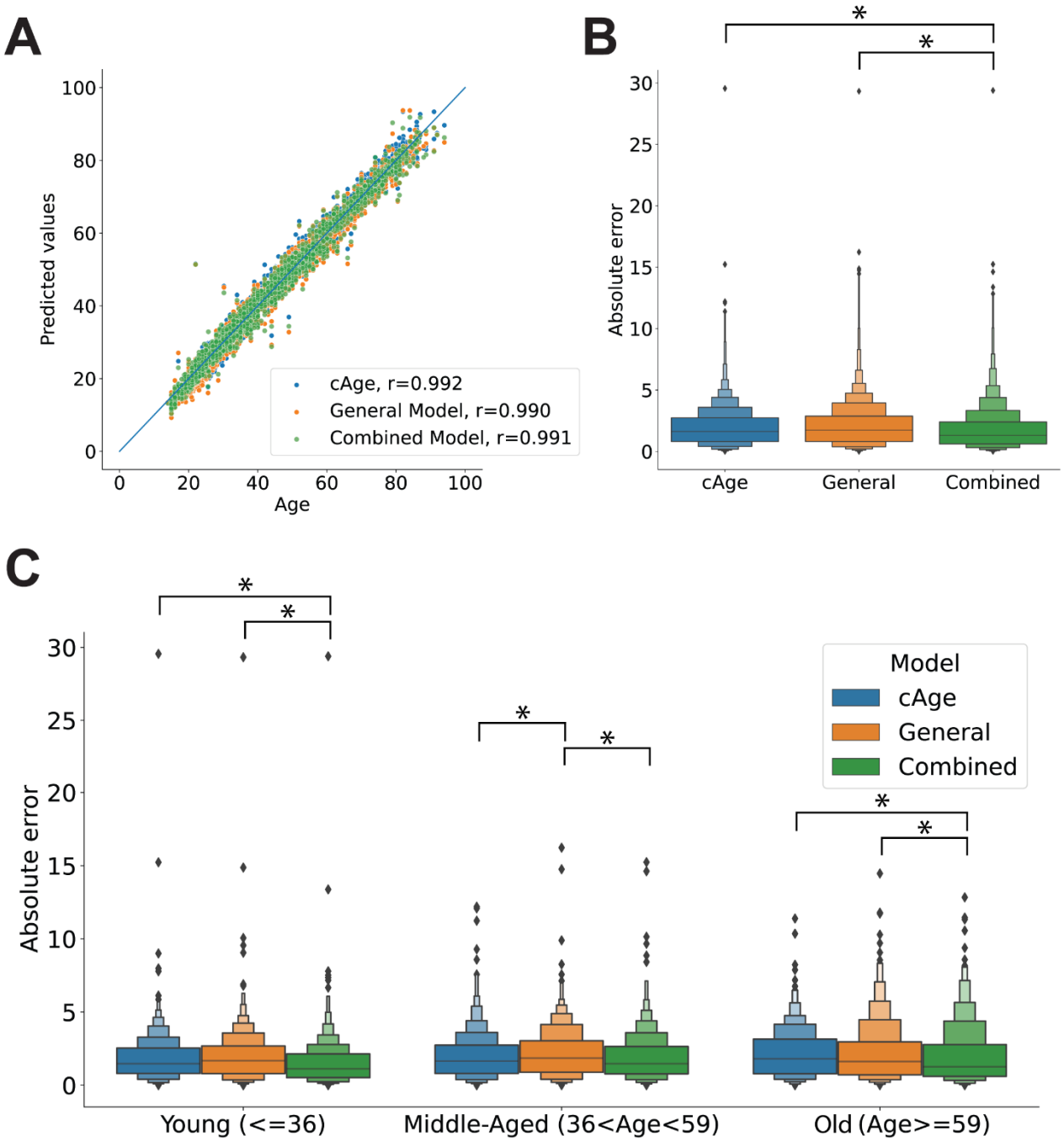


Figure 4 – **Age predictions on the validation set**. **A**. Chronological age and age values predicted by cAge and the general and combined models in the validation set. **B**. Distribution of the absolute prediction error for each of the three models. **C**. Distribution of the absolute error of each model in the validation set broken down by age group. *Indicates significant differences ($p<0.05$, one-sided Wilcoxon rank-sum test) between groups.

Technical and intra-subject EAA variations

It has recently been shown that epigenetic clocks are affected by spontaneous variation in DNA methylation values, leading to changes in epigenetic age predictions[37,38]. Additionally, the experimental process to obtain methylation values might introduce another source of variation. In order to establish the magnitude of these variations, we applied our model and the four epiclocks discussed above to data obtained from repeated sampling of the same subjects.

First, we examined the data from the study of Koncevičius et al., which reported DNA methylation of blood samples of a 52-year-old male subject taken every 3 hours over a period of 72 hours[37] (GSE247197). This dataset included technical replicates, which we used to estimate the EAA deviations associated with technical noise, i.e. variations across multiple experimentally determined beta values on the same sample. All the (complete) models tested exhibited variation across technical replicates (Figure 5A). To characterize this variation, we subtracted the mean value obtained on each set of replicates and then computed the standard deviation across all the mean-centered values. We interpreted the resulting measure as an estimate of the technical noise (σ_{noise}). The predictions of our general model had the lowest σ_{noise} (0.78 years), whereas Horvath's model had the highest with, reaching 1.26 years (Figure 5B, Table 1). We then computed the standard deviation across all samples in the dataset as a first estimate for the intra-subject variability ($\sigma_{subject}$). The values of $\sigma_{subject}$ were in all cases larger than σ_{noise}, and in this case the largest and lowest values corresponded to our combined model (0.83 years) and to PhenoAge (2.36 years) respectively (Figure 5B, Table 1).

We then applied all the models to the data from Apsley et al., discarding the samples from subjects under the stress test. This data includes DNA methylation data from blood drawn from 31 subjects at 4 different time points over a period of 4 hours and 45 minutes under stress or control conditions[38] (GSE227809). Using data from the 14 subjects in control condition, we computed the standard deviation across the mean-subtracted predictions of all subjects for all the models. We interpreted this measure as another estimate of $\sigma_{subject}$. The results were largely compatible with the estimate of $\sigma_{subject}$ from Koncevičius' data; the largest difference was observed on Horvath's model, for which $\sigma_{subject}$ rose from 1.99 to 3.31 years. In all the other models, $\sigma_{subject}$ changed by less than 1 year. Also in this dataset, our combined model's predictions suffered the least variation, whereas those from Horvath's model changed the most (Figure 5C, Table 1).

According to these results, the sum of technical and intra-subject variation of the EAA predictions represent ~50 to 60% of the population-wide σ_{EAA} depending on the model. Taken together, the metrics in Table 1 provide a means to judge whether the magnitude of EAA findings puts them beyond the range of normal variations.

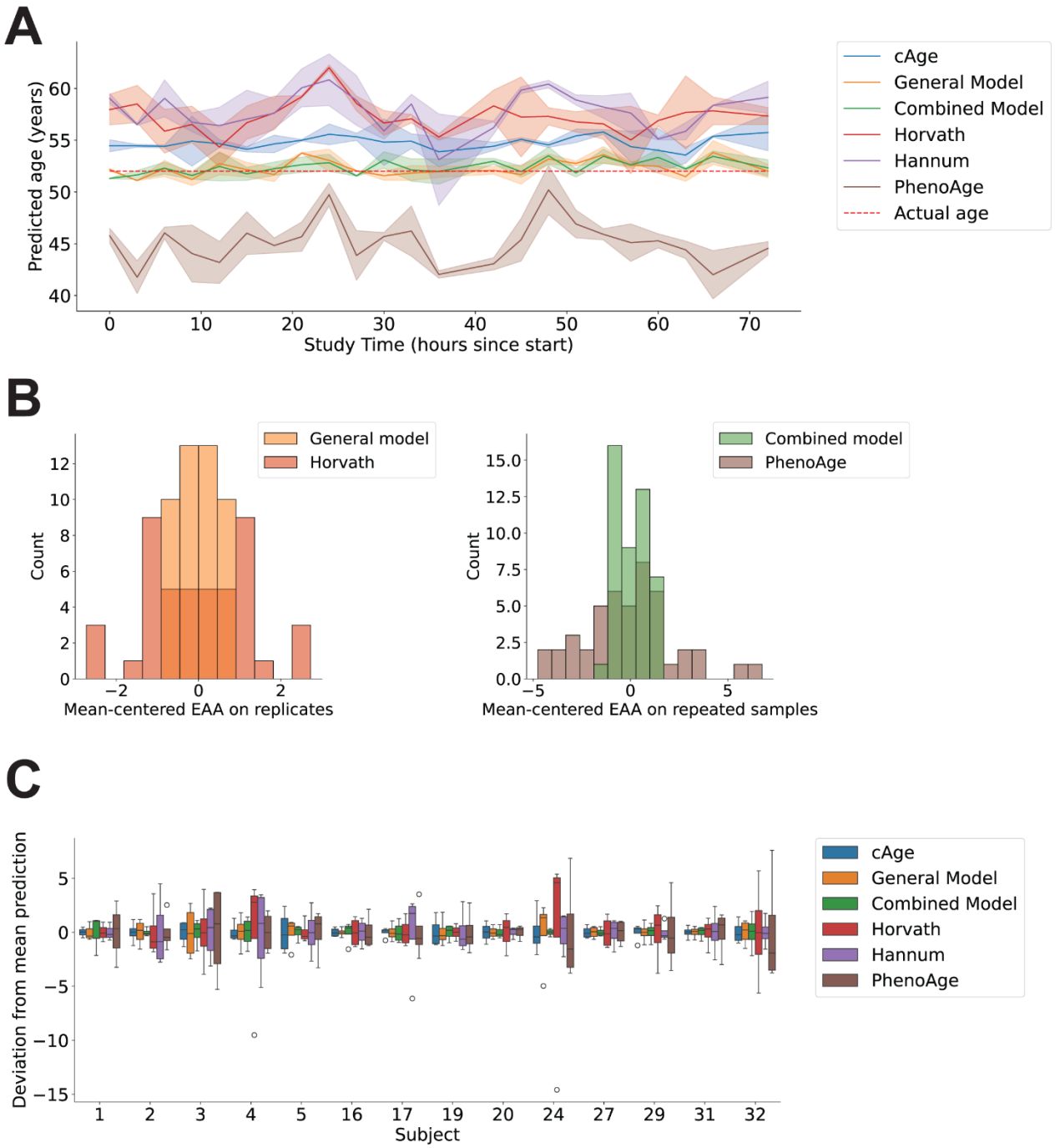


Figure 4 – **Age predictions technical replicates and repeated samples**. **A**. Age values predicted by the different models on the dataset from Koncevičius et al., with multiple samples of the same subject (52-year-old male) obtained at different times and with replicates. Solid lines indicate the mean values, whereas the colored area marks the 95% confidence intervals. **B**. Distribution of mean-centered EAA values on technical replicates (left) and repeated samples from the same subject (right) on the Koncevičius et al. data for the models with the highest and lowest standard deviations. **C.** Distribution of mean-centered age predictions for the 14 subjects from the Apsley et al.´s dataset.

Table 1. Metrics from all the complete (C) and truncated (T) models on the different datasets used in the study.

		$Hannum_C$	$Hannum_T$	$Horvath_C$	$Horvath_T$	$PhenoAge_C$	$PhenoAge_T$	$cAge_C$	$cAge_T$	General Model	Combined Model
Main dataset	**r**	0,92	0,93	0,90	0,90	0,91	0,91	0,98	0,98	N/A	N/A
	MAE (years)	7,58	19,08	6,37	6,65	7,05	6,44	2,55	3,57	N/A	N/A
	MSE (years)	81,58	410,80	65,85	69,57	78,43	66,92	12,82	20,88	N/A	N/A
	μ_{EAA} (years)	6,38	-19,06	3,32	4,14	-5,25	-4,14	1,07	2,73	N/A	N/A
	σ_{EAA} (years)	6,39	6,89	7,40	7,24	7,13	7,06	3,42	3,67	N/A	N/A
Validation dataset	**r**	0,96	0,97	0,94	0,94	0,94	0,94	0,99	0,99	0,99	0,99
	MAE (years)	7,04	18,93	6,10	6,59	7,90	7,16	1,94	2,21	2,08	1,74
	MSE (years)	68,45	397,10	55,64	63,44	89,60	74,98	6,41	8,32	7,55	5,96
	μ_{EAA} (years)	6,21	-18,92	3,08	4,24	-6,85	-5,86	0,01	1,37	-0,54	-0,21
	σ_{EAA} (years)	5,47	6,27	6,80	6,74	6,53	6,38	2,53	2,54	2,70	2,43
Koncevičius et al.	**σ_{noise} (years)**	1,18	1,01	1,26	1,19	1,13	1,09	0,78	0,86	0,46	0,50
	$\sigma_{subject}$ (years)	2,20	1,74	1,99	1,85	2,37	2,31	0,98	1,04	0,89	0,83
Apsley et al.	**$\sigma_{subject}$ (years)**	2,06	1,77	3,31	3,28	2,61	2,58	0,99	1,05	1,30	0,88

The new epigenetic clock reflects the influence of radiation therapy and breast cancer

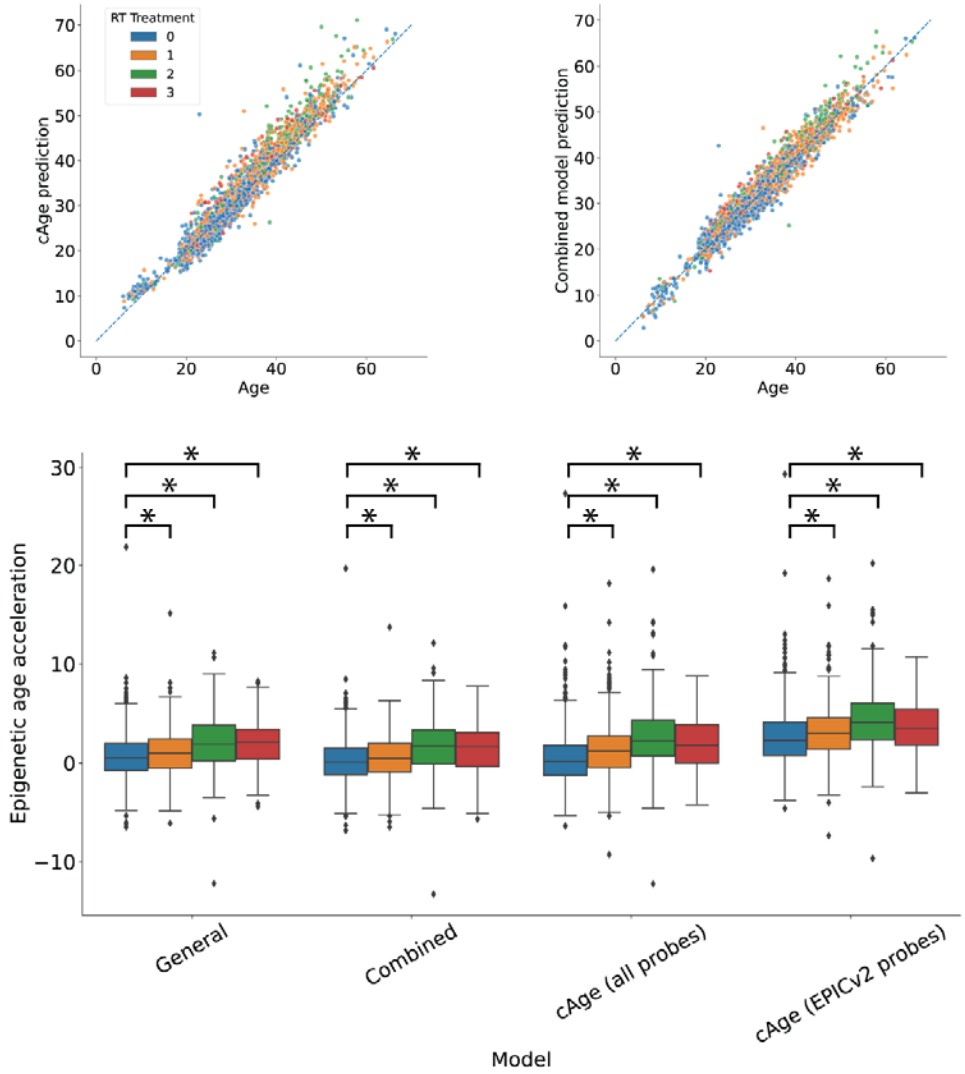
Several studies have employed different epigenetic clocks to assess the impact of pathologies and environmental factors on epigenetic age acceleration (EAA)[45]. Using PhenoAge[16], Qin *et al*. demonstrated that exposure to different anticancer treatments had a significant influence on the EAA of childhood cancer survivors[46]. To validate the performance of our model, we tested it in public datasets that have already shown increased EAA in cancer patients. Using data made public in a subsequent study by Dong *et al*.[47] (GEO accession number GSE197674) which contained data from 2138 childhood cancer survivors, we studied the influence of radiation therapy (RT) on the EAA determined by cAge and our models. In this dataset, all the models produced epigenetic age predictions highly correlated with chronological age, with r values above 0.9 in all cases (Figure 6A). When comparing subjects who had received RT in one or more body areas (chest, abdomen or pelvis, brain) to those who had not been exposed to RT, all models revealed that the latter group had a significantly lower ($p<0.05$, one-sided Wilcoxon test) EAA (Figure 6A). This is in line with the results reported in the original study[46], and similar observations have been made in other studies in which radiation exposure was associated with an increased EAA[48,49]. The rest of the complete and truncated models assign very large EAA values to virtually all samples and do not attribute significantly larger EAAs ($p>0.05$ in all cases, one-sided Wilcoxon test) to the groups with RT treatments (Supplementary Figure 3). In all cases, the magnitude of these EAA changes is smaller than the σ_{EAA} in the general population (Table 2). Therefore, it is questionable whether the observed EAA differences (here as well as in previous studies) are a true, biologically relevant, reflection of the RT treatment during childhood.

Next, we leveraged a dataset with DNA methylation data from peripheral blood leukocytes of sporadic breast cancer patients and control subjects[50] (GEO accession number GSE148663) to study the effect of the disease on the EAA determined by the different models. We computed the predicted epigenetic age for 22 sporadic breast cancer patients and 10 controls using our models and cAge. Although the correlation between the epigenetic age predictions and the chronological ages of the subjects remained high (0.914-0.951), our models predicted a significantly greater ($p<0.05$, one-sided Wilcoxon test) EAA in the cancer group (Figure 6B). As reported in the original study[50], the cancer patients had no previous cancer history and were sampled at the time of diagnosis, so the difference in EAA cannot be attributed to anticancer therapy but rather to the disease itself. The rest of the complete and truncated models do not assign a significantly greater EAA ($p>0.05$ in all cases, one-sided Wilcoxon test) to the breast cancer group (Supplementary Figure 4). In this case, the magnitude of the EAA differences between cancer patients and controls determined by our models are larger than the σ_{EAA}. Therefore, we can consider that these changes do reflect differences larger than the normal variations across individuals in the general population (Table 3).

Finally, we explored data an interventional study using resistance and aerobic training[51] (GEO accession number GSE213363). Besides its numerous benefits for overall health[52,53] and aging[54,55], physical activity has demonstrated influence on DNA methylation[56,57]. Yet, the study by Furtado et al. reported that after

a 16-week exercise intervention, they did not detect significant changes in the epigenetic age calculated by the Horvath model[51]. We applied all the models discussed here to the data from subjects before and after the intervention and found the same result: none of the epiclocks showed significant reductions in epigenetic age after the intervention ($p>0.05$ in all cases, one-sided Wilcoxon test), neither in the resistance nor in the aerobic training groups (Supplementary Figure 5). This suggests that although the first-generation models examined here can predict chronological age with considerably accuracy, they are less responsive in terms of ability to reflect other biological factors or the effect of certain interventions.

A


RT Treatment
0
1
2
3
cAge prediction
Age
Combined model prediction
Age
Epigenetic age acceleration
General
Combined
cAge (all probes)
cAge (EPICv2 probes)
Model

B

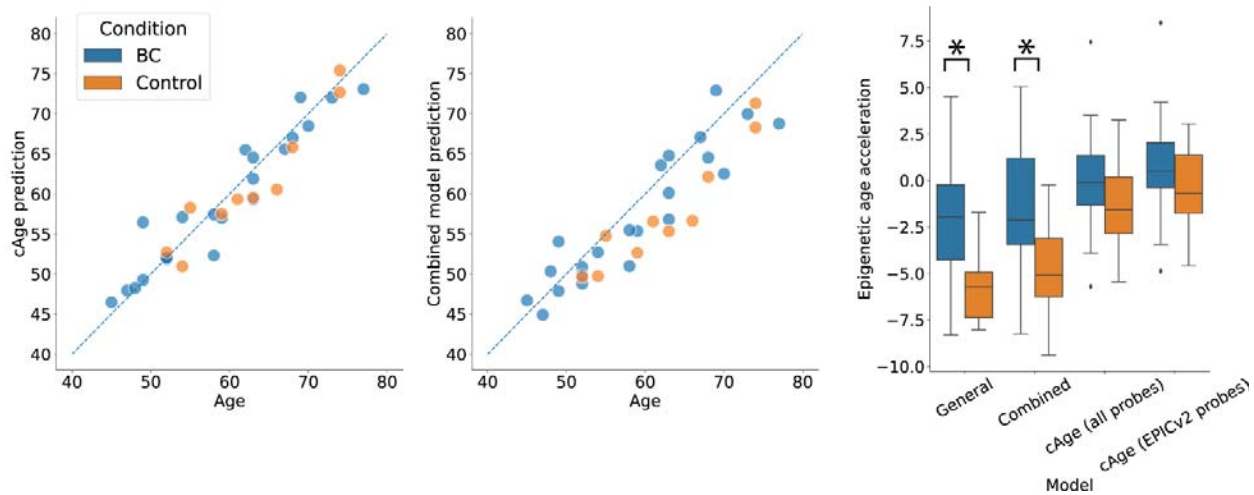

Condition
BC
Control
cAge prediction
Age
Combined model prediction
Age
Epigenetic age acceleration
General
Combined
cAge (all probes)
cAge (EPICv2 probes)
Model

Figure 6 - Effects of radiation therapy and spontaneous breast cancer in EAA. **A.** Chronological age and predicted age for childhood cancer survivors, colored by the number of body areas (brain, chest, abdomen/pelvis) in which they received radiation therapy (top). Distribution of the EAA for each group of subjects determined by each of the models (bottom). **B.** Chronological ages of spontaneous breast cancer patients (BC) and control subjects and the corresponding epigenetic ages determined by the combined model and cAge (left). Distribution of the EAA estimated for each group by the different models (right).

Table 2. Differences between average EAA of the RT>0 and RT=0 groups observed in the different complete (C) and truncated (T) models.

	$Hannum_C$	**$Hannum_T$**	**$Horvath_C$**	**$Horvath_T$**	**$PhenoAge_C$**	**$PhenoAge_T$**	**$cAge_C$**	**$cAge_T$**	**General Model**	**Combined Model**
$EAA_{RT1-RT0}$ (years)	-0,82	-1,23	-1,65	-1,65	-1,70	-1,70	0,96	0,54	0,39	0,39
$EAA_{RT2-RT0}$ (years)	-0,19	-1,07	-1,68	-1,68	-1,87	-1,87	2,37	1,73	1,34	1,48
$EAA_{RT3-RT0}$ (years)	0,35	-0,15	-0,65	-0,65	-0,35	-0,35	2,05	1,46	1,43	1,44

Table 3. Average EAA differences between cancer patients and healthy controls observed in the different complete (C) and truncated (T) models.

	$Hannum_C$	**$Hannum_T$**	**$Horvath_C$**	**$Horvath_T$**	**$PhenoAge_C$**	**$PhenoAge_T$**	**$cAge_C$**	**$cAge_T$**	**General Model**	**Combined Model**
EAA_{BC}-$EAA_{Control}$ (years)	2,41	2,46	-0,03	-0,03	-0,76	-0,76	1,30	1,27	3,37	3,12

Discussion

Epigenetic clocks are valuable tools for research on aging and pathological states. In this work, we evaluated the applicability of existing epigenetic clocks to the data generated by EPICv2, the newest methylation microarray model from Illumina. The EPICv2 will phase out previous microarray models (namely, 450k and EPIC), which were used for the development of some widely used epigenetic clocks (e.g., Hannum, Horvath). As the EPICv2 discontinued the use of some of the probes used by existing epiclocks[36], we felt urged to test whether this would affect the performance of the models.

We generated a large DNA methylation dataset by compiling data from whole blood samples obtained in 24 different studies. This dataset allowed us to quantify the effect of running 4 different epiclocks (Hannum[14], Horvath[15], PhenoAge[16] and the cAge model from Bernabeu *et al.*[28]) restricted to the set of methylation probes present in the EPICv2 methylation array. Our observations indicate that the results of these 4 models are significantly altered by probes absent from the EPICv2 array.

Epigenetic age models are routinely applied in epigenetic research and are also exploited commercially[58–60]. In the face of commercial discontinuation of the 450k and EPICv1 arrays, our findings suggest that both researchers and commercial vendors alike will need to update their epigenetic clock models to make them compatible with EPICv2 data. Future studies and commercial solutions using this new microarray will require readjusted or new epigenetic age models, which are currently lacking.

As our results show, *none* of the epiclocks tested work as intended on data generated by the EPICv2 microarrays, showing significant differences between the epigenetic ages predicted using data from different chip models. Therefore, we sought to produce a model compatible across the 450k, EPICv1 and EPICv2 microarray platforms, aiming to obtain results in accordance with those of existing methods rather than outperforming them. The approach that we used closely followed the one applied by Bernabeu *et al.* in the development of their cAge model[28], so that we would rely on established methodologies. Our results on the training and test datasets indicate that our epiclock offered very high performance in the prediction of ages from DNA methylation values. Similar to Bernabeu's model, these results highlight the benefit of using feature preselection, nonlinear terms, and age-based stratification. We followed a simple strategy regarding these aspects, as we relied on a prior EWAS study[28] to preselect features, we considered only linear and quadratic features, and we stratified our subjects into three arbitrary age groups. We are confident that follow-up work can improve the results presented here by, for example, conducting more thorough feature selection, deriving more features from the original beta values and/or exploring different data stratification strategies.

Despite its simplicity, our implementation outperformed cAge in the validation dataset, where it obtained significantly lower prediction errors. These results demonstrate that our model can be generalizable (as it is applicable to new data) and that its performance is on par with that of state-of-the-art models.

We tested the variability of epigenetic age predictions on two datasets with repeated samples from the same subjects obtained hours apart from each other. The results indicated that the predictions of cAge and our combined model had the lowest variations across replicates and across repeated samples. We claim that these spontaneous variations of epigenetic age predictions should be considered when examining EAA-related findings: differences smaller than those seen on repeated samples of the same subject obtained on the same day could hardly be considered biologically relevant. Consequently, the significance of previous reports of altered EAAs based on the models d here might need to be re-evaluated.

To demonstrate the ability of our model to reproduce the results of previous models in the detection of EAA alterations, we first applied it to methylation data from cancer survivors[46]. The results of our epiclock indicated an increased EAA induced by radiation therapy, in agreement with previous studies[48,49]. However, the magnitude of this increase (0.39 to 1.44 years) was smaller than the variation

we observed in the general population (σ_{EAA}). Thus, we cannot claim that the EAA changes detected are reflection of the RT effect on the subjects.

Next, we applied our model to data obtained from breast cancer patients. Our epiclock revealed a significant increase in the EAA of cancer patients compared to that of control subjects, suggesting that the epigenetic age predicted by the model could be sensitive to this particular pathology. In this case, the average EAA difference between the cancer patients and the control group (3.13-3.37 years) was larger than the σ_{EAA} from the non-pathological subjects in the validation dataset (2.43-2.7 years). Therefore, these changes could indeed be a reflection of the pathological state. Notably, the cAge model did not detect a significant difference in the EAA between the control and the cancer groups.

Finally, we analyzed data from an interventional study aimed at improving the health of patients suffering from polycystic ovary syndrome. As in the original study, none of the models we tested reported significant differences in the epigenetic age of the subjects before and after the intervention. Considering the large and extensively documented benefits of physical exercise on health, these results support the idea that despite being good predictors of chronological age, first-generation epigenetic clocks do not necessarily reflect biological factors such as pathological states, environmental exposures or interventions.

As a relevant limitation of our model, we would like to highlight that it is limited to methylation data obtained from whole blood samples, a constraint that is also shared by multiple other models[28,61,62]. Its applicability to data obtained from other tissues has not been assessed. Likewise, we would like to emphasize that its ability to reflect pathological states beyond the sporadic breast cancer cases we have shown here remains to be explored.

Taken together, our results demonstrate that our epigenetic clock is compatible with data generated using the 450k, EPICv1 and EPICv2 microarray platforms. Its epigenetic age predictions are highly correlated with chronological age in control subjects of all ages. Our model exhibited consistency across technical replicates and across repeated samples of the same subjects, with lower variation than the rest of the models tested. As a first-generation epigenetic clock, our model can predict chronological age from DNA methylation data with high accuracy, but its ability to reflect the effects of environmental exposures, pathological states or beneficial interventions is limited. Our work solves a technical barrier derived from technological development that has not yet been addressed and has important implications both for aging research and for biotechnology companies offering services in this field.

Methods

Data collection and preprocessing

Methylation data from previous studies were collected from the GEO database[63,64]. In all cases, the data were preprocessed: we used beta values provided by the original authors when available; otherwise, we computed the beta value from the methylated and unmethylated signals using the following formula:

$$\beta = \frac{M}{M+U+100},$$

where M and U are the intensities of the methylated and unmethylated signals respectively. The main dataset consisted of 11825 subjects with reported methylation, age and sex data. We discarded 147 with nonnumerical age values, one with a reported age of 891 years, and one with a negative epigenetic age according to the cAge model. To generate training and test sets with even age distributions, we kept only samples with an age value present at least twice in the dataset. This process left a total of 10835 samples on the main dataset, which were split into training and test sets with 7259 and 3576 subjects, respectively.

On the main dataset, we computed the average beta value of each probe across the whole dataset (Supplementary Table 4) and used this average to impute missing values. Among the probes used to train our model (see below) or for running other epigenetic clocks, we found 1.58% missing values. The number of probes missing per dataset is provided in Supplementary Table 5.

The validation dataset included data from 2098 subjects. Excluding subjects with missing or invalid sex and/or age annotations reduced the data to 2095 subjects. A total of 0.23% of the beta values were missing for the set of probes used to run our epiclcok and the cAge models. The missing values were imputed using the average values derived from the main dataset. Supplementary Table 5 reports the number of probes missing per dataset.

The data from studies with repeated samples (GSE247197, GSE227809, 1 and 31 subjects respectively), the cancer survivor dataset (GSE197674, 2138 subjects), the data from the breast cancer study (GSE148663, 32 subjects) and the data from the interventional study (GSE213363, 56 subjects) were also retrieved from the GEO database[63,64]. We used the average beta values derived from the main datasets to impute beta values for 25 missing probes in the cancer survivor dataset.

Epigenetic age prediction using existing models

We applied the Horvath[15], Hannum[14], PhenoAge[16] and cAge[28] epiclocks to the methylation data in the main dataset by using the methods and parameters reported by the authors for each case and validated using the implementations in the pyaging Python library[65]. In the case of the Horvath clock, this implies computing an initial value using a linear model and then applying additional transformations for the "young" (<=20) and "adult" (>20) subjects[15]. Similarly, the cAge model uses one model to predict *age* and another to predict log(*age*); if the *age* prediction is <=20, then it is replaced by the exponential of the log(*age*) prediction[28].**Training of the new epigenetic clock**

We trained a general epigenetic age prediction model using the beta values of 10,000 probes and the squared beta values of 300 probes to perform a regularized linear regression of chronological age using elastic net[40]. The linear and quadratic features were chosen based on an EWAS study[28]: we chose the 10,000 probes with a significant linear association with age with the lowest association p-value and the 300 probes with a significant quadratic association with age with the lowest association p-value.

To train our general epigenetic age prediction model, we separated the main dataset into training and test sets using a 70/30 split stratified by age (i.e., all ages present in the dataset were sampled in both the training and test sets). Elastic net combines L1 and L2 regularizations, which in the scikit-learn implementation[66] are mixed in proportions given by the parameter L1_ratio, which takes values between 0 and 1. This parameter multiplies the L1 penalty term whereas the L2 term is multiplied by 1-L1_ratio.

Basedon its performance in previous studies, we set the L1_ratio to 0.5[15,67,]. The value of the parameter alpha (i.e. 1/C), which multiplies both the L1 and L2 penalties, was optimized to 0.001 through 5-fold cross-validation.

Age-specific models were trained using the same approach on 3 different batches of data with subjects of specific ages: age<=46 (young), 26<age<69 (middle-aged), and age>=59 (old). The alpha value was set to 0.001 in all cases. The three age-specific models were then used to perform predictions on nonoverlapping *putative* age (age values assigned by the general model) groups: age<=36 (young), 36<age<59 (middle-aged), and age>=59 (old).

List of abbreviations

Epigenetic age acceleration (EAA)

Chronological age prediction model (cAge)

Epigenome-wide association study (EWAS)

Mean squared error (MSE)

Radiation therapy (RT)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed during the current study are available in the GEO repository. All accession numbers are reported in Supplementary Table 1.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors' contributions

L.G. and M.Q. designed the study. L.G. generated and analyzed the data. All authors read and approved the final manuscript.

Acknowledgements

We are grateful to all the researchers who made their data publicly available and enabled us to carry out this work.

References

1. Bocklandt, S. *et al.* Epigenetic predictor of age. *PloS one* **6**, e14821 (2011).
2. Jones, M. J., Goodman, S. J. & Kobor, M. S. DNA methylation and healthy human aging. *Aging cell* **14**, 924–932 (2015).
3. Ryan, C. P. "Epigenetic clocks": Theory and applications in human biology. *American Journal of Human Biology* **33**, e23488 (2021).
4. Oblak, L., van der Zaag, J., Higgins-Chen, A. T., Levine, M. E. & Boks, M. P. A systematic review of biological, social and environmental factors associated with epigenetic clock acceleration. *Ageing research reviews* **69**, 101348 (2021).
5. Bozack, A. K. *et al.* DNA methylation age at birth and childhood: performance of epigenetic clocks and characteristics associated with epigenetic age acceleration in the Project Viva cohort. *Clinical Epigenetics* **15**, 62 (2023).
6. Christiansen, L. *et al.* DNA methylation age is associated with mortality in a longitudinal Danish twin study. *Aging cell* **15**, 149–154 (2016).
7. Marioni, R. E. *et al.* DNA methylation age of blood predicts all-cause mortality in later life. *Genome biology* **16**, 1–12 (2015).
8. Perna, L. *et al.* Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clinical epigenetics* **8**, 1–7 (2016).
9. Marioni, R. E. *et al.* The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. *International journal of epidemiology* **44**, 1388–1396 (2015).

10. Valencia, C. I., Saunders, D., Daw, J. & Vasquez, A. DNA methylation accelerated age as captured by epigenetic clocks influences breast cancer risk. *Frontiers in Oncology* **13**, 1029 (2023).
11. Berstein, F. M. *et al.* Assessing the causal role of epigenetic clocks in the development of multiple cancers: a Mendelian randomization study. *Elife* **11**, e75374 (2022).
12. Rozenblit, M. *et al.* Evidence of accelerated epigenetic aging of breast tissues in patients with breast cancer is driven by CpGs associated with polycomb-related genes. *Clinical Epigenetics* **14**, 30 (2022).
13. Gehle, S. C. *et al.* Accelerated epigenetic aging and myopenia in young adult cancer survivors. *Cancer Medicine* **12**, 12149–12160 (2023).
14. Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular cell* **49**, 359–367 (2013).
15. Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**, 1–20 (2013).
16. Levine, M. E. *et al.* An epigenetic biomarker of aging for lifespan and healthspan. *Aging (albany NY)* **10**, 573 (2018).
17. Belsky, D. W. *et al.* DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* **11**, e73420 (2022).
18. Lu, A. T. *et al.* DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (albany NY)* **11**, 303 (2019).
19. Freire-Aradas, A. *et al.* A common epigenetic clock from childhood to old age. *Forensic Science International: Genetics* **60**, 102743 (2022).
20. Prosz, A. *et al.* Biologically informed deep learning for explainable epigenetic clocks. *Scientific Reports* **14**, 1306 (2024).
21. Tomusiak, A. *et al.* Development of a novel epigenetic clock resistant to changes in immune cell composition. *bioRxiv* 2023–03 (2023).
22. Caulton, A. *et al.* Development of epigenetic clocks for key ruminant species. *Genes* **13**, 96 (2021).

23. Zoller, J. A. *et al.* DNA methylation clocks for clawed frogs reveal evolutionary conservation of epigenetic aging. *GeroScience* **46**, 945–960 (2024).

24. Sala, C. *et al.* Where are we in the implementation of tissue-specific epigenetic clocks? *Frontiers in Bioinformatics* **4**, 1306244 (2024).

25. Pośpiech, E. *et al.* Epigenetic clock in the aorta and age-related endothelial dysfunction in mice. *GeroScience* 1–10 (2024).

26. Voisin, S. *et al.* An epigenetic clock for human skeletal muscle. *Journal of cachexia, sarcopenia and muscle* **11**, 887–898 (2020).

27. Coninx, E. *et al.* Hippocampal and cortical tissue-specific epigenetic clocks indicate an increased epigenetic age in a mouse model for Alzheimer's disease. *Aging (Albany NY)* **12**, 20817 (2020).

28. Bernabeu, E. *et al.* Refining epigenetic prediction of chronological and biological age. *Genome Medicine* **15**, 12 (2023).

29. Bibikova, M. *et al.* Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics* **1**, 177–200 (2009).

30. Weisenberger, D. J., Van Den Berg, D., Pan, F., Berman, B. & Laird, P. Comprehensive DNA methylation analysis on the Illumina Infinium assay platform. *Illumina, San Diego* (2008).

31. Pidsley, R. *et al.* Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome biology* **17**, 1–17 (2016).

32. Noguera-Castells, A., García-Prieto, C. A., Álvarez-Errico, D. & Esteller, M. Validation of the new EPIC DNA methylation microarray (900K EPIC v2) for high-throughput profiling of the human DNA methylome. *Epigenetics* **18**, 2185742 (2023).

33. Sandoval, J. *et al.* Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. *Epigenetics* **6**, 692–702 (2011).

34. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389–399 (2016).

35. Horvath, S. *et al.* Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies. *Aging (Albany NY)* **10**, 1758 (2018).

36. Kaur, D. *et al.* Comprehensive evaluation of the Infinium human MethylationEPIC v2 BeadChip. *Epigenetics Communications* **3**, 6 (2023).

37. Koncevičius, K. *et al.* Epigenetic age oscillates during the day. *Aging Cell* e14170 (2024).

38. Apsley, A. T. *et al.* Biological stability of DNA methylation measurements over varying intervals of time and in the presence of acute stress. *Epigenetics* **18**, 2230686 (2023).

39. Oh, G. *et al.* Circadian oscillations of cytosine modification in humans contribute to epigenetic variability, aging, and complex disease. *Genome biology* **20**, 1–14 (2019).

40. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology* **67**, 301–320 (2005).

41. Johnson, N. D. *et al.* Non-linear patterns in age-related DNA methylation may reflect CD4+ T cell differentiation. *Epigenetics* **12**, 492–503 (2017).

42. Okada, D., Cheng, J. H., Zheng, C., Kumaki, T. & Yamada, R. Data-driven identification and classification of nonlinear aging patterns reveals the landscape of associations between DNA methylation and aging. *Human Genomics* **17**, 8 (2023).

43. Carlsen, L., Holländer, O., Danzer, M. F., Vennemann, M. & Augustin, C. DNA methylation-based age estimation for adults and minors: considering sex-specific differences and non-linear correlations. *International Journal of Legal Medicine* **137**, 635–643 (2023).

44. Bell, C. G. *et al.* DNA methylation aging clocks: challenges and recommendations. *Genome biology* **20**, 1–24 (2019).

45. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature reviews genetics* **19**, 371–384 (2018).

46. Qin, N. *et al.* Epigenetic age acceleration and chronic health conditions among adult survivors of childhood cancer. *JNCI: Journal of the National Cancer Institute* **113**, 597–605 (2021).

47. Dong, Q. *et al.* Genome-wide association studies identify novel genetic loci for epigenetic age acceleration among survivors of childhood cancer. *Genome Medicine* **14**, 32 (2022).

48. Sehl, M. E., Carroll, J. E., Horvath, S. & Bower, J. E. The acute effects of adjuvant radiation and chemotherapy on peripheral blood epigenetic age in early stage breast cancer patients. *NPJ Breast Cancer* **6**, 23 (2020).

49. Xiao, C. *et al.* Epigenetic age acceleration, fatigue, and inflammation in patients undergoing radiation therapy for head and neck cancer: a longitudinal study. *Cancer* **127**, 3361–3371 (2021).

50. Cappetta, M. *et al.* Discovery of novel DNA methylation biomarkers for non-invasive sporadic breast cancer detection in the Latino population. *Molecular oncology* **15**, 473–486 (2021).

51. Miranda Furtado, C. L. *et al.* Resistance and aerobic training increases genome-wide DNA methylation in women with polycystic ovary syndrome. *Epigenetics* **19**, 2305082 (2024).

52. Ruegsegger, G. N. & Booth, F. W. Health benefits of exercise. *Cold Spring Harbor perspectives in medicine* **8**, a029694 (2018).

53. Mandolesi, L. *et al.* Effects of physical exercise on cognitive functioning and wellbeing: biological and psychological benefits. *Frontiers in psychology* **9**, 509 (2018).

54. DiPietro, L. Physical activity in aging: changes in patterns and their relationship to health and function. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* **56**, 13–22 (2001).

55. Paterson, D. H., Jones, G. R. & Rice, C. L. Ageing and physical activity: evidence to develop exercise recommendations for older adults. *Applied physiology, nutrition, and metabolism* **32**, S69–S108 (2007).

56. Grazioli, E. *et al.* Physical activity in the prevention of human diseases: role of epigenetic modifications. *BMC genomics* **18**, 111–123 (2017).

57. Ferioli, M. *et al.* Role of physical exercise in the regulation of epigenetic mechanisms in inflammation, cancer, neurodegenerative diseases, and aging process. *Journal of cellular physiology* **234**, 14852–14864 (2019).

58. Biological Age Test | Horvath's Clock | myDNAge. https://mydnage.com/.

59. TruDiagnostic.com. https://www.trudiagnostic.com/.

60. Elysium Health - Healthy Aging Supplements. https://www.elysiumhealth.com/.

61. Lee, Y. *et al.* Blood-based epigenetic estimators of chronological age in human adults using DNA methylation data from the Illumina MethylationEPIC array. *BMC genomics* **21**, 1–13 (2020).

62. Knight, A. K. *et al.* An epigenetic clock for gestational age at birth based on blood methylation data. *Genome biology* **17**, 1–11 (2016).

63. Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research* **41**, D991–D995 (2012).

64. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research* **30**, 207–210 (2002).

65. de Lima Camillo, L. P. pyaging: a Python-based compendium of GPU-optimized aging clocks. *Bioinformatics* btae200 (2024).

66. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* **12**, 2825–2830 (2011).

67. Zhang, Q. *et al.* Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing. *Genome medicine* **11**, 1–11 (2019).