

## RESEARCH

## Open Access



# Systematic underestimation of the epigenetic clock and age acceleration in older subjects

Louis Y. El Khoury<sup>1,2</sup>, Tyler Gorrie-Stone<sup>1</sup>, Melissa Smart<sup>3</sup>, Amanda Hughes<sup>4</sup>, Yanchun Bao<sup>3</sup>, Alexandria Andrayas<sup>1</sup>, Joe Burrage<sup>5</sup>, Eilis Hannon<sup>5</sup>, Meena Kumari<sup>3</sup>, Jonathan Mill<sup>5</sup> and Leonard C. Schalkwyk<sup>1\*</sup> 

## Abstract

**Background:** The Horvath epigenetic clock is widely used. It predicts age quite well from 353 CpG sites in the DNA methylation profile in unknown samples and has been used to calculate “age acceleration” in various tissues and environments.

**Results:** The model systematically underestimates age in tissues from older people. This is seen in all examined tissues but most strongly in the cerebellum and is consistently observed in multiple datasets. Age acceleration is thus age-dependent, and this can lead to spurious associations. The current literature includes examples of association tests with age acceleration calculated in a wide variety of ways.

**Conclusions:** The concept of an epigenetic clock is compelling, but caution should be taken in interpreting associations with age acceleration. Association tests of age acceleration should include age as a covariate.

**Keywords:** DNA methylation, Epigenetic clock, Age acceleration

## Background

Subject age is a piece of data available in almost every study in which DNA methylation profiles are obtained. There is thus a huge amount of cross-sectional data in which it can be seen that the methylation level of many CpG sites varies with subject age [1–4], which, amongst other processes, could reflect developmental changes, cellular aging, cumulative environmental effects, and changes in cell-type composition. Exploring these sources of variation could give insights into age-related processes. Predicted ages can also provide a valuable quality control and identity check on data in EWAS studies [5–8].

Horvath [8] used a large collection ( $n > 8000$ ) of publicly available Illumina HumanMethylation array data on multiple tissue types to train and test a model for age prediction from 353 CpG loci. This “epigenetic clock” continues to be widely used and is extremely valuable as a way of estimating ages of samples from unknown donors and possibly as an indicator of whether there are

alterations in the aging rate of certain tissues or individuals. Although the epigenetic clock developed by Horvath [8] provides an estimate of age, the testing data used in generating this clock did not have a large representation of tissue from elderly individuals and as such it is unclear if the clock is accurate in older age groups, or those with age-related diseases.

We have previously published an epigenome-wide association study (EWAS) in Alzheimer disease (AD), utilizing four brain tissues and pre-mortem blood, and demonstrated DNA methylation differences at specific loci in a tissue-specific manner [9]. This dataset offers a good opportunity to examine the properties of the Horvath [8] clock on different tissues in both elderly non-demented individuals and AD sufferers. We further explore the properties of the model using a cross-sectional population sample from the UK Household Longitudinal study (UKHLS), which has a wide range of ages [10].

\* Correspondence: [lschal@essex.ac.uk](mailto:lschal@essex.ac.uk)

<sup>1</sup>School of Life Sciences, University of Essex, Colchester, UK

Full list of author information is available at the end of the article



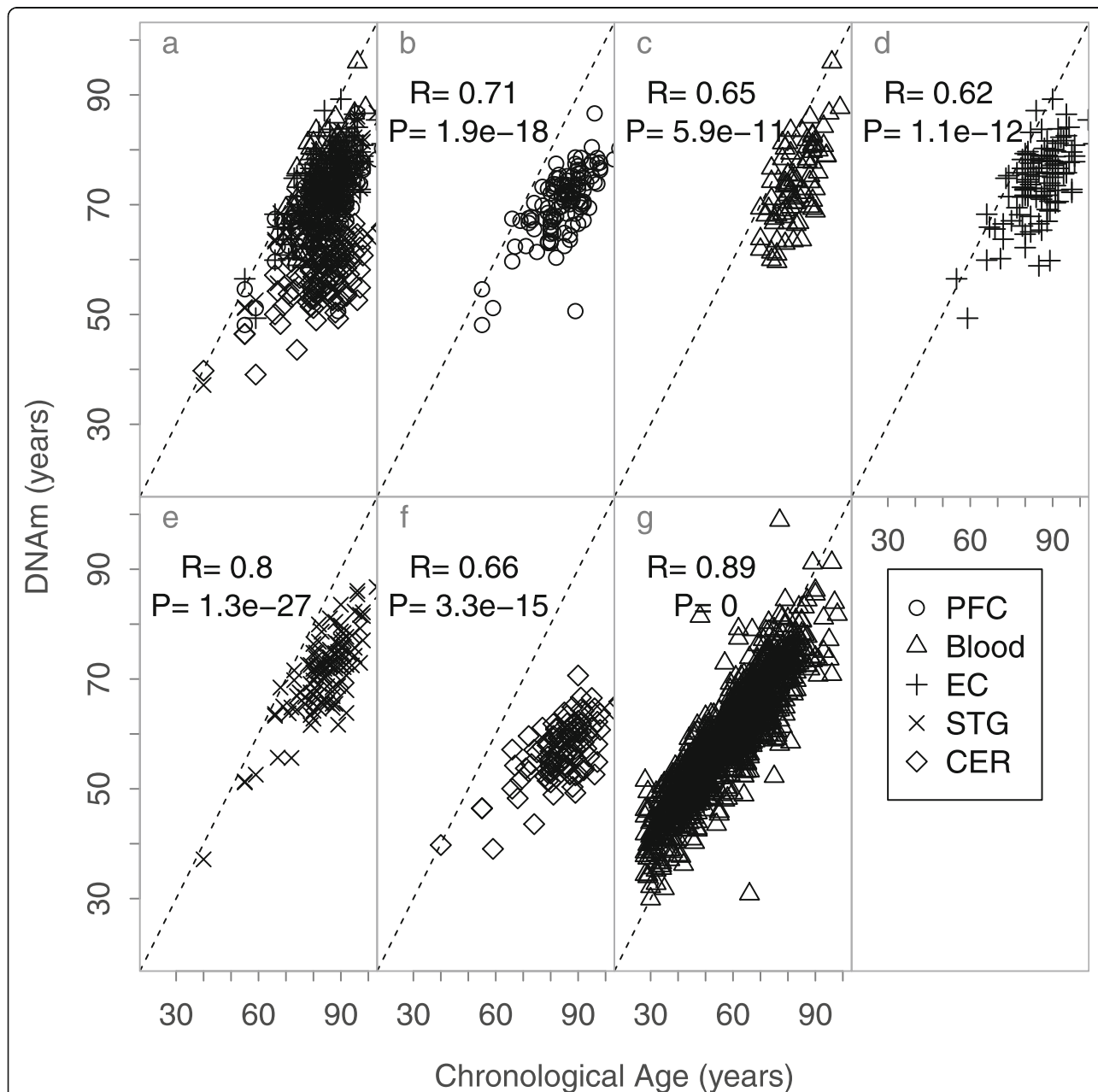
**Results**

**Age estimation**

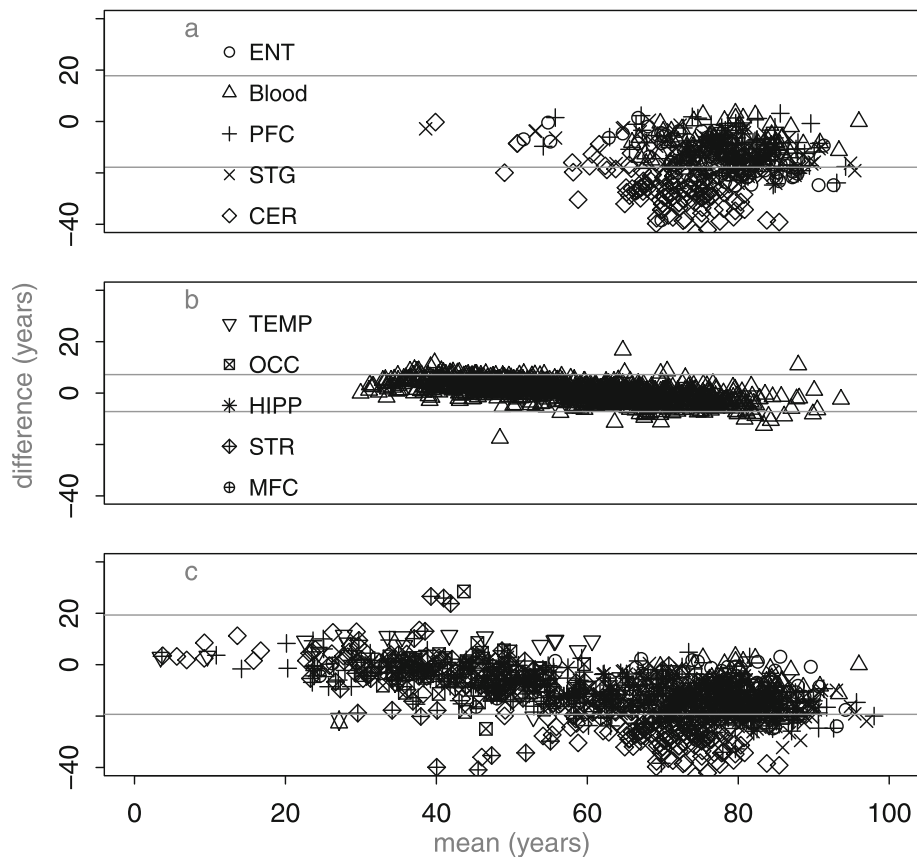
Initially, we observed in our AD dataset [9] that ages were strikingly underestimated using the Horvath [8] clock. Indeed, in this elderly data set, across multiple brain regions and in blood, the model did predict age but with a slope of predicted against actual age clearly less than 1 (Fig. 1a–f). This was borne out in the much larger UKHLS set of blood DNA samples measured with the Illumina EPIC array [10] (Fig. 1g).

The discrepancy can be more clearly demonstrated with a mean-difference (Bland-Altman) plot (thanks to a reviewer for this suggestion). There is a trend to larger discrepancy with age in the the AD data set, UKHLS, and a collection of additional datasets listed in Additional file 1: Table S1 (Fig. 2).

We focus in this study on the characteristics of the Horvath model [8] because it remains widely used and because it is designed to be applicable across tissues. For the UKHLS dataset, we also looked at the



**Fig. 1** Scatterplots of chronological vs DNAm ages of brain and blood samples. Each point corresponds to an independent sample. The dotted line is the  $y = x$  bisector line, and the solid lines correspond to the regression line of each tissue. PFC, prefrontal cortex; EC, entorhinal cortex; STG, the superior temporal gyrus; CER, cerebellum (data from [9] for panels a–f and [10] for panel g)



**Fig. 2** Mean-difference (Bland-Altman) plots showing the difference between Horvath model age and chronological age. **a** Elderly brain: AD data set [9], **b** population blood sample [10], and **c** additional publicly available datasets (see Additional file 1: Table S1). The horizontal lines in each case are at  $\pm 1.96 * sd$ ; for normally distributed difference due to error 5% of points would lie outside these and in each case many more do

blood-specific Hannum [6] model. This model is simpler than the Horvath one, with 71 coefficients. On the UKHLS dataset, it shows a very similar deviation, with increasing underestimation in advanced age (Fig. 3). There is an overlap of 6 loci between the two models.

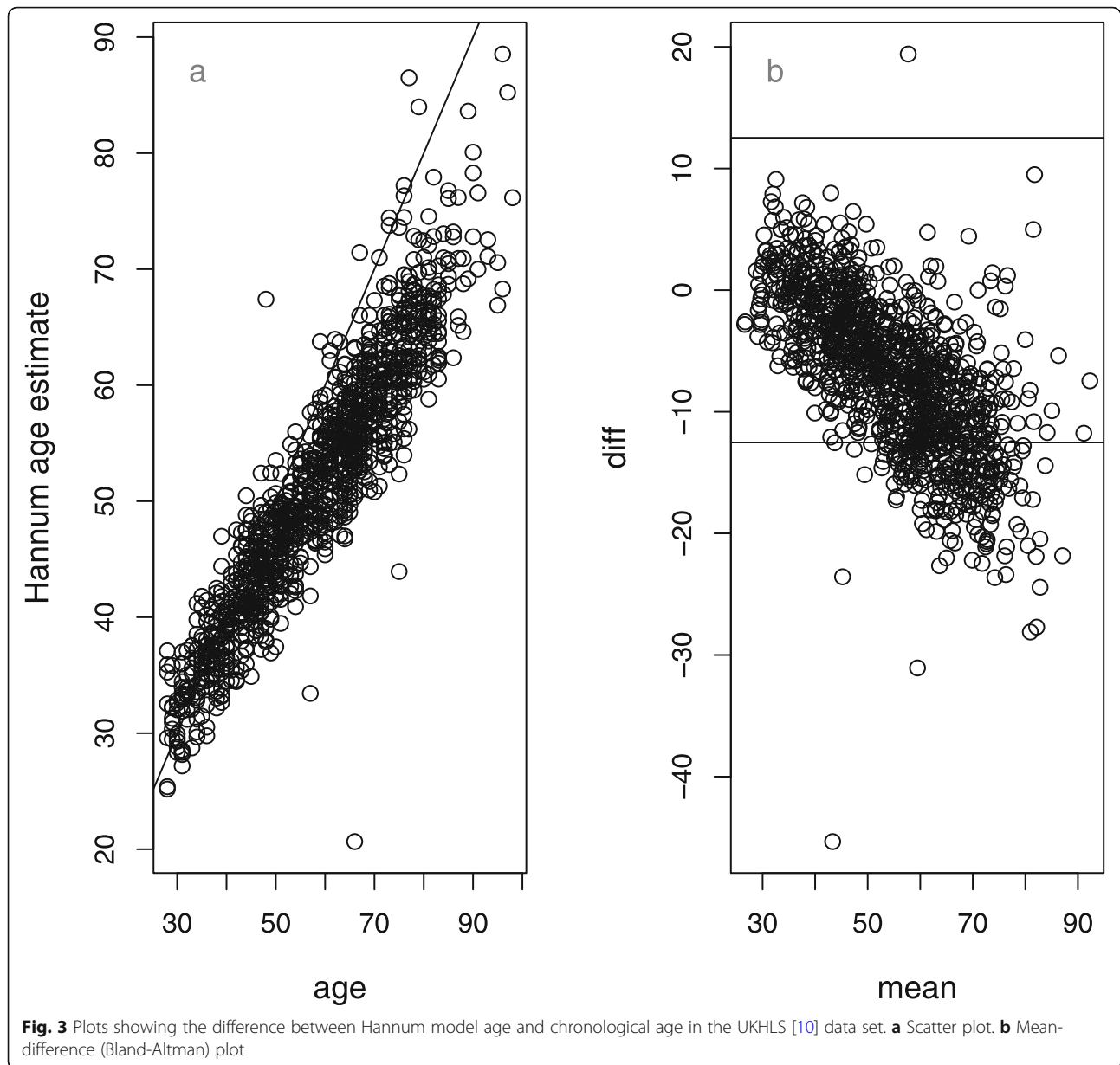
#### Clock properties

Next, to get some insight into the behavior of the clock, we dissected possible reasons for its nonlinear response. The absolute values of the coefficients in the Horvath model range from  $5.9e-6$  to 3.07, so some of the CpG sites are much more important than others. To investigate this, we make an index of the influence of each locus on the age estimate by dividing the absolute value of the coefficient from the Horvath clock by an index of dispersion from our data, the coefficient of variation. The ten highest ranked probes by this measure (Fig. 4a, black circles at the top and bottom of the plot) include examples of both small variance (and large coefficient) and large variance (and possibly smaller coefficient),

although clearly the smallest coefficients are not going to make an appreciable difference to the age estimate no matter what the variance. Two of the ten most influential probes, cg22736354 and cg06493994, are also used in the Hannum model.

One simple way the clock rate could be reduced in later life is by saturation, i.e., the CpG sites used in the clock reaching either full methylation or complete demethylation. To investigate this, we dissect the ten most influential further (Fig. 4b). Of these ten probes, three (cg12830694, cg24580001, and cg02580606) might be candidates for saturation because they are highly methylated and expected from the model to be increasing with age.

To test this further, we fitted a regression line between chronological age and the beta values of each of the 353 loci and plotted the slopes against the Horvath coefficients. Of the ten most influential loci, four have a slope opposite in sign to the Horvath coefficient (cg08090772, cg03019000, cg04268405, and cg24580001) (Fig. 4c). Loci with the largest coefficients demonstrate the same direction of effect in our data as in the Horvath model.



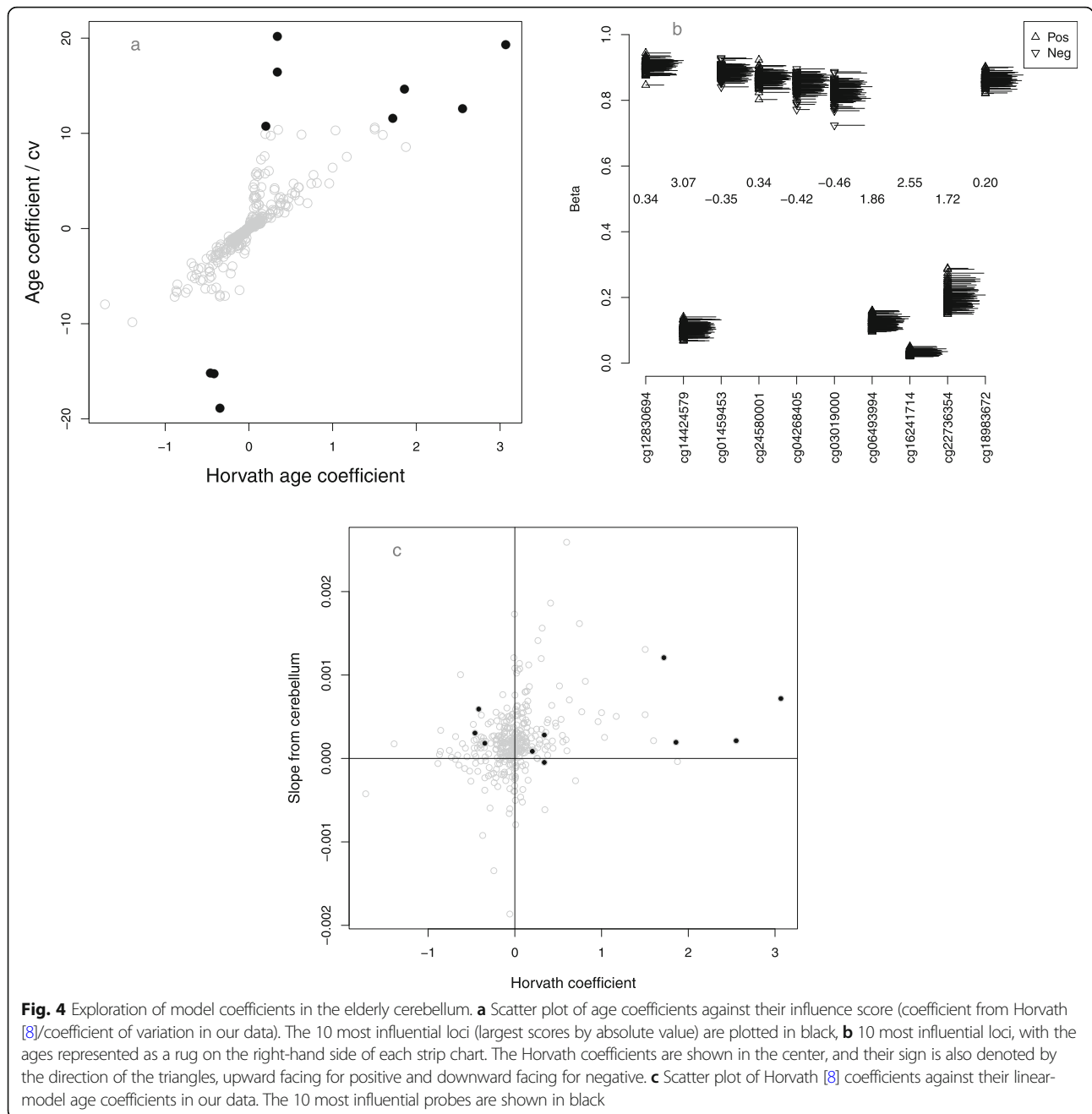
For this tissue and age range, many of the smaller coefficients are effectively random, but they may be influential in the model's performance in other tissues. Slieker et al. [11] have shown that the majority of age effects are tissue-specific.

#### Age acceleration associations

As an example of an association test using age acceleration, we examined whether age acceleration (calculated as the difference between DNAm age and chronological age) associates with AD neuropathology (measured using Braak score) in the London cohort [9]. We observed a weak association in some brain tissues. However, when age is included as a covariate, the association between age

acceleration and AD pathology disappears. We also see this in the Mount Sinai cohort [12, 13] where no correlation was found between age acceleration and amyloid plaque levels when age is included as a covariate (Table 1).

In a broad but non-comprehensive survey of the literature (Table 2), we observe a variety of methods of calculating age acceleration, and many studies do not correct for chronological age. Initially,  $\Delta$ -age (the difference between chronological age and the DNAm predicted age) was reported, but alternative methods have since arisen: (1) the residual of regressing DNAm predicted age on chronological age (possibly in a model including covariates), (2) AgeAccel (difference between DNAm age value and the value predicted by a regression model in the



control group), and (3) intrinsic (IEAA) and (4) extrinsic epigenetic age acceleration (EEAA) methods. Both IEAA and EEAA are methods applicable only on blood since they subtract out the effect of blood cell count [16, 18].

**Discussion**

The Horvath epigenetic clock [8] has been of practical use in predicting the age of unknown samples and as a quality check in epigenetic research. Additional widely used age predictors specific for blood were published by Hannum [6] and Levine [42] (phenotype-based). Here

we analyze the Horvath model, but the methods and many of the conclusions may be more widely applicable, in particular the Hannum clock model shows a similar underestimation of ages in elderly subjects.

The mechanism or mechanisms of the apparent change of gears in a person’s sixties are not clear. At least part of the effect with these models seems to be saturation, i.e., loci approaching the limits of 0 or 100% methylation. Another intriguing part of the picture, at least for the brain, could be 5-hydroxymethyl cytosine, which is present at appreciable levels in brain tissues,

**Table 1** Regression analysis of epigenetic age acceleration of four brain tissues and blood from the London Brain Bank cohort [9] versus brain Braak stage and of two brain tissues from the Mount Sinai cohort [12] versus amyloid plaque levels

	London cohort		London cohort (age in model)		Mount Sinai cohort (age in model)	
	Coefficient	<i>P</i> value	Coefficient	<i>P</i> value	Coefficient	<i>P</i> value
Prefrontal cortex	0.434	0.164	−0.345	0.107	−0.029	0.629
Entorhinal cortex	0.969	0.004	0.424	0.125	–	–
Sup. Temporal Gyrus	0.653	0.019	−0.062	0.786	0.004	0.944
Cerebellum	1.127	0.002	−0.059	0.788	–	–
Blood	−0.278	0.632	−0.090	0.870	–	–

Shown for each cohort the coefficient and *P* value for the regression analysis between age acceleration and Braak stages (London cohort) or amyloid plaque levels (Mount Sinai cohort) with chronological age as a covariate

especially cerebellum which is characterized by elevated levels of 5-hydroxymethylcytosine (5hmC) [15]. We found that 31 out of the 353 Horvath clock sites were amongst the 65,663 elevated 5hmC probes found in the cerebellum by Lunnon et al. [15]. Of these, two sites (cg04268405, and cg24580001) are amongst the most influential sites (Fig. 4). Given that 5hmC is not distinguished from 5mC following bisulfite conversion, it is possible that age-associated changes to the 31 5hmC sites of the Horvath algorithm are offsetting the age predictions.

These two models both use a small fraction of the available age indexing GpGs, especially since much more comprehensive arrays are now in use, and in fact their site contents overlap. Although improved age prediction can no doubt be achieved by making use of additional informative loci, especially in tissue-specific models, we believe it is more important to use existing models with an awareness of their properties and limitations and not as a black box.

In addition to age prediction, the Horvath [8] paper also featured the idea of “age acceleration” in which discrepancies between DNA methylation (DNAm) age and chronological age might tell us something about the biological aging status of the organism. A number of positive association findings with age association, particularly mortality [43], make it compelling to think of the epigenetic clock as an index of an underlying aging program that adapts to health and environment. In light of the methodological variety though, we are concerned that the different epigenetic clocks, and the variety of age acceleration methods to choose from, lay a trap of potentially hidden multiple testing, as the temptation will be to survey the available methods for interesting results.

When comparing DNA methylation profiles across tissues, individuals, and other variables such as health, the dominant source of variation is the tissue, or more precisely the cell type. It is reasonable to suppose that this developmental blueprint can change over time in response to the environment, or simply drift or decay. This

point of view corresponds roughly with the “epigenetic maintenance” model posited by Horvath [8], and developed further by Horvath and Raj [44].

The “decay clock” or epigenetic maintenance models are perhaps more likely to be accurate than a biochemical aging clock, but they are somewhat at odds with the age acceleration concept. Association tests with age acceleration are very common but should be treated with caution, especially if the effect is small. As shown in this study, in the latter third of the human age range, where such associations are most likely observed, negative age acceleration increases with age. This means that any phenotype associated with age will appear to be associated with age acceleration as well, and a correct analysis should include chronological age as a covariate, as in the Alzheimer disease example that we report.

A preprint which appeared while this paper was under review [45] suggests that age acceleration may result from confounding of age with other phenomena such as blood cell composition, and indicates that adding further age-predicting loci to a clock model reduces association of mortality with age acceleration.

The clock model has interesting and useful characteristics, but it is an extremely narrow summary of the DNA methylation profile based on only 353 CpG sites representing  $1.15 \times 10^{-5}\%$  of the methylome. EWAS, association tests of the full DNA methylation profile, using appropriate genome-wide confidence limits are much more likely to lead to biological insights.

## Conclusion

The age prediction properties of both Horvath [8] and Hannum et al. [6] DNA methylation clock models begin to degrade as subjects enter old age. This is at least partly due to saturation, i.e., DNA methylation proportion at some loci approaching 0 or 1, and confounding with the effects of other age-related processes will also play a role. It is likely that this could be ameliorated with additional loci and/or further refined modeling of the

**Table 2** Literature survey of age acceleration

Reference	Phenotype	AIM	SIG	Age acceleration method
Horvath [14]	Obesity	No	Yes	Residuals of DNAm age regressed on chronological age
Marioni et al. [15]	All-cause mortality	Yes	Yes	$\Delta$ -age
Levine et al. [16]	Lung cancer incidence	Yes	Yes	IEAA
Levine et al. [17]	Neuritic plaque	Yes	Yes	Residuals of DNAm age regressed on chronological age and sex
Horvath et al. [18]	Parkinson's disease	Yes	No	AgeAccel
		Yes	No	IEAA
		Yes	Yes	EEAA
Horvath et al. [19]	Down syndrome	No	Yes	Residuals DNAm age regressed on chronological age in controls
Marioni et al. [20]	Fitness variables	Yes	Yes	Residuals DNAm age regressed on chronological age
Horvath et al. [21]	HIV	No	Yes	AgeAccel
Horvath et al. [22]	Alzheimer in the cerebellum	Yes	No	Residual DNAm age against chronological age in non-cerebellar brain sample
Horvath et al. [23]	Longevity	No	Yes	AgeAccel
		No	Yes	IEAA
		No	Yes	EEAA
Walker [24]	Multifocal developmental disorders	No	No	Residuals DNAm age regressed on chronological age
Lin et al. [25]	Life expectancy	Yes	Yes	$\Delta$ -age
Perna [26]	All-cause mortality	Yes	Yes	$\Delta$ -age
	Cancer mortality	Yes	Yes	$\Delta$ -age
	Cardiovascular mortality	Yes	Yes	$\Delta$ -age
Horvath [27]	Huntington's disease	No	Yes	Residuals DNAm age regressed on chronological age
Levine et al. [28]	Age at menopause	Yes	Yes	AgeAccel
Horvath [29]	Race/ethnicity	Yes	Yes	EEAA
Chen et al. [30]	Time of death	Yes	Yes	EEAA
		Yes	Yes	IEAA
		Yes	Yes	AgeAccel
Simpkin et al. [31]	Birth weight	No	Yes	$\Delta$ -age
Gao et al. [32]	Smoking status	Yes	No	Residuals DNAm age regressed on chronological age
	Cumulative exposure	Yes	No	Residuals DNAm age regressed on chronological age
	Cessation time	Yes	No	Residuals DNAm age regressed on chronological age
Breitling et al. [33]	Frailty	Yes	Yes	$\Delta$ -age
Ward-Caviness et al. [34]	Air pollution	No	Yes	Residuals DNAm age regressed on chronological age
		Yes	Yes	EEAA
		Yes	Yes	IEAA
Levine et al. [35]	HIV-associated neurocognitive disorders	Yes	Yes	Residuals DNAm age regressed on chronological age
Armstrong et al. [36]	Longevity	No	Yes	AgeAccel - Hannum
McKinney et al. [37]	Schizophrenia duration	No	No	Residuals DNAm age regressed on chronological age
Wolf et al. [38]	PTSD hyperarousal	No	Yes	Residuals DNAm age regressed on chronological age
	PTSD severity	No	No	Residuals DNAm age regressed on chronological age
Quach et al. [39]	Diet and lifestyle	Yes	Yes	EEAA
	Diet and lifestyle	Yes	Yes	IEAA
Binder et al. [40]	Time to menarche	No	Yes	AgeAccel
	Pubertal tempo	No	Yes	AgeAccel
	Breast fibro-glandular volume	No	Yes	AgeAccel

**Table 2** Literature survey of age acceleration (Continued)

Reference	Phenotype	AIM	SIG	Age acceleration method
Dugué [41]	Mortality	Yes	Yes	Residuals DNAm age regressed on chronological age
		Yes	Yes	IEAA

currently used set. Association tests using age acceleration should incorporate age as a covariate (as should those using DNA methylation values for individual loci) to avoid spurious associations.

## Methods

This study was designed to investigate age prediction from DNA methylation profiles across multiple brain regions and blood, especially in older subjects. It uses a number of existing data sets.

## Samples

### Tissue samples

Brain tissue samples (London cohort) were obtained from individuals diagnosed with Alzheimer's disease (AD,  $n = 61$ ) and from non-demented elderly control individuals (CON,  $n = 31$ ) through the MRC London Neurodegenerative Disease Brain Bank as described in Lunnon et al. [9, 46]. In total, four brain regions were analyzed (prefrontal cortex (PFC), the entorhinal cortex (EC), the superior temporal gyrus (STG), and the cerebellum (CER)) and pre-mortem blood from a subset of individuals, collected as part of the Biomarkers of AD Neurodegeneration study. A second independent cohort (Mount Sinai cohort) was obtained from the Mount Sinai Alzheimer's disease and Schizophrenia Brain Bank. This cohort consisted of two brain regions (PFC and STG) for 75 AD and 72 CON donors [12, 13].

### Population sample: the UK Household Longitudinal Study (UKHLS)

UKHLS is an annual household-based panel study which started collecting information about the social, economic, and health status of its participants in 2009. Our analysis data set is drawn from one of the arms of UKHLS, namely, the British Household Panel Survey (BHPS), which merged with UKHLS in 2010 at the start of wave two. UKHLS collected additional health information, including blood samples for genetic and epigenetic analysis, at wave 3 for BHPS ([www.understandingsociety.ac.uk](http://www.understandingsociety.ac.uk)). DNA methylation profiling and initial analysis are described in [10, 47].

### Methylomic profiling

DNA from the London cohort tissue samples were bisulfite-treated using Zymo EZ 96 DNA methylation kit (Zymo Research) according to the manufacturer's protocol. DNA methylation levels were assessed on an Illumina

HiScan System using the Illumina Infinium HumanMethylation450 BeadChip as previously described by Lunnon et al. [9]. Raw signal intensities and probes for the London cohort were extracted using Illumina Genome Studio software and were transformed into beta values using the Bioconductor `wateRmelon` package [48]. These were later normalized using the method implemented in the Horvath [8] script. Data is available from both cohorts under GEO accession numbers GSE59685 (London cohort) and GSE80970 (Mount Sinai cohort).

One thousand one hundred ninety-three DNA samples from UKHLS were bisulfite-treated using Zymo EZ 96 DNA methylation kit (Zymo Research) according to the manufacturer's protocol. DNA methylation levels were assessed on an Illumina HiScan System (Illumina) using the Illumina Infinium Epic Methylation BeadChip, and samples were randomly assigned to chips and plates to minimize batch effects. Furthermore, in order to resolve any experimental inconsistencies, and to approve data quality, a fully methylated control (CpG Methylated HeLa Genomic DNA; New England BioLabs, MA, USA) was included in a random position on each plate. Raw signal intensities and probes for UKHLS were extracted using Illumina Genome Studio software and were transformed into beta values using the Bioconductor `bigmelon` package (<https://doi.org/10.18129/B9.bioc.bigmelon>) [49]. These were later normalized using `dasen` function from the `wateRmelon` package [48]. After QC, a final  $n$  of 1175 was reached.

### DNA methylation age prediction

DNA methylation (DNAm) age was assessed for all samples of the London and Mt Sinai datasets on the R statistical environment (R Development Core Team, 2015) using the script provided by Horvath [8] as well as through the online DNAm Age Calculator (<https://dna-age.genetics.ucla.edu/>). These methods predicted age based on the DNAm coefficients of 353 CpG sites. The model (although not the custom normalization method) is also implemented in the `agep()` function of the `wateRmelon` package (version 1.17.0). This is expected to perform very similarly to the original Horvath protocol as long as reasonable preprocessing steps are used. A BA plot demonstrates that this is the case for the AD study samples (Additional file 1: Figure S1). The differences are small (sd of difference 1.8 years) and approximately normally distributed. The `agep()` function was used to predict the ages of the UKHLS samples for this study.



To maximize the number of brain samples included in our assessment of age prediction, publicly available 450K Methylation brain tissue datasets obtained from GEO (GSE40360, GSE53162, GSE59457, GSE61380, GSE61431, GSE67748, GSE67749, and GSE89702 [50–60]) along with the London and Mount Sinai cohorts were analyzed (Additional file 1: Table S1).

## Supplementary information

The online version of this article (<https://doi.org/10.1186/s13059-019-1810-4>) contains supplementary material, which is available to authorized users.

**Additional file 1:** Fig. S1. Describing the differences between DNA methylation ages estimated with the Horvath 2013 calculator and the `agep()` function. As well as Table S1. Detailing additional data sets used in this study.

## Acknowledgements

*Understanding Society*: Analysis was facilitated by access to the High Performance Computing facilities at the University of Essex.

## Authors' contributions

LS designed the study; LEK, TGS, and LS did the analyses; MS, AH, YB, AA, JB, EH, and TGS generated and curated the UKHLS DNA methylation data; MK and JM advised and oversaw the work; and all authors were involved in drafting and approving the final manuscript.

## Funding

This work was supported by grants from the UK Medical Research Council (MRC) (grant numbers MR/K013807/1 and MR/R005176/1) to JM and LS. The AD data sets were produced under NIH grant R01 AG036039 to JM and LS. The UK Household Longitudinal Study is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council (Grant Number: ES/M008592/1).

## Availability of data and materials

The data used in this publication is all previously published. UKHLS DNA methylation data is available from the European Genome-phenome Archive under accession EGAS00001002836 (<https://www.ebi.ac.uk/ega/home>) [61]. Specific details can be found here (<https://www.understandingsociety.ac.uk/about/health/data>). Phenotypes linked to DNA methylation data are available through application to the METADAC ([www.metadac.ac.uk](http://www.metadac.ac.uk)). Publicly available 450K Methylation brain tissue datasets were obtained from GEO (GSE40360, GSE53162, GSE59457, GSE61380, GSE61431, GSE67748, GSE67749, and GSE89702 [50–60]). Detailed description is available from GEO and listed in Additional file 1: Table S1.

## Ethics approval and consent to participate

London cohort—ethical approval NHS was provided by South East London REC 3; Mount Sinai Cohort—ethical approval for the project was provided by the University of Exeter Medical School Research Ethics Committee under application number 14/02/041; UKHLS—participants gave informed written consent for their blood to be taken and stored for future scientific analysis. The UKHLS has been approved by the University of Essex Ethics Committee and the nurse data collection by the National Research Ethics Service (10/H0604/2). All experimental methods performed comply with the Helsinki Declaration.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>School of Life Sciences, University of Essex, Colchester, UK. <sup>2</sup>Present Address: Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA. <sup>3</sup>Institute for Social and Economic Research, University of Essex, Colchester, UK. <sup>4</sup>MRC Integrative Epidemiology Unit -

University of Bristol, Bristol, UK. <sup>5</sup>Medical School, University of Exeter, Exeter, UK.

Received: 20 August 2018 Accepted: 3 September 2019

Published online: 17 December 2019

## References

- Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Bürkle A, Caiafa P. Reconfiguration of DNA methylation in aging. *Mech Ageing Dev.* 2015;151:1–11.
- Bjornsson HT, Sigurdsson MI, Fallin MD, Irizarry RA, Aspelund T, Cui H, et al. Intra-individual change over time in DNA methylation with familial clustering. *JAMA.* 2008;299:2877–83.
- Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochim Biophys Acta.* 2007;1775:138–62.
- Bell JT, Tsai P-C, Yang T-P, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 2012;8:e1002629.
- Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CCY, O'Donovan MC, et al. Methylation trajectories across human fetal brain development. *Genome Res.* 2015;25:338–52.
- Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* 2013;49:359–67.
- Zbieć-Piekarska R, Spólnicka M, Kupiec T, Parys-Proszek A, Makowska Ż, Pałeczka A, et al. Development of a forensically useful age prediction method based on DNA methylation analysis. *Forensic Sci Int Genet.* 2015;17:173–9.
- Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14:R115.
- Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylation profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. *Nat Neurosci.* 2014;17:1164–70.
- Hughes A, Smart M, Gorrie-Stone T, Hannon E, Mill J, Bao Y, et al. Socioeconomic position and DNA methylation age acceleration across the life course. *Am J Epidemiol.* 2018;187:2346–54.
- Slieker RC, Relton CL, Gaunt TR, Slagboom PE, Heijmans BT. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. *Epigenetics Chromatin.* 2018;11:25.
- Smith RG, Hannon E, De Jager PL, Chibnik L, Lott SJ, Condliffe D, et al. Elevated DNA methylation across a 48-kb region spanning the HOXA gene cluster is associated with Alzheimer's disease neuropathology. *Alzheimer's Dement.* 2018; <https://doi.org/10.1016/j.jalz.2018.01.017>.
- Smith R, Hannon E, De Jager P, Chibnik L, Lott S, Condliffe D, Smith A, Haroutunian V, Troakes C, Bennett D, Powell J, Lovestone S, Schalkwyk L, Mill J, Lunnon K GSE80970 Gene Expression Omnibus. 2016 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE80970>.
- Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad Sci U S A.* 2014;111:15538–43.
- Lunnon K, Hannon E, Smith RG, Dempster E, Wong C, Burrage J, et al. Variation in 5-hydroxymethylcytosine across human cortex and cerebellum. *Genome Biol.* 2016;17:27.
- Levine ME, Hosgood HD, Chen B, Absher D, Assimes T, Horvath S. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. *Aging.* 2015;7:690–700.
- 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.* 2015;7:1198–211.
- Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging.* 2015;7:1130–42.
- Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, et al. Accelerated epigenetic aging in Down syndrome. *Aging Cell.* 2015;14:491–5.
- Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. *Int J Epidemiol.* 2015;44:1388–96.
- Horvath S, Levine AJ. HIV-1 infection accelerates age according to the epigenetic clock. *J Infect Dis.* 2015;212:1563–73.
- Horvath S, Mah V, Lu AT, Woo JS, Choi O, Jasinska AJ, et al. The cerebellum ages slowly according to the epigenetic clock. *Aging.* 2015;7:294–306.

23. Horvath S, Pirazzini C, Bacalini MG, Gentilini D, Di Blasio AM, Delledonne M, et al. Decreased epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. *Aging*. 2015;7:1159–70.
24. Walker RF, Liu JS, Peters BA, Ritz BR, Wu T, Ophoff RA, et al. Epigenetic age analysis of children who seem to evade aging. *Aging*. 2015;7:334–9.
25. Lin Q, Weidner CI, Costa IG, Marioni RE, Ferreira MRP, Deary IJ, et al. DNA methylation levels at individual age-associated CpG sites can be indicative for life expectancy. *Aging*. 2016;8:394–401.
26. Perna L, Zhang Y, Mons U, Holleczeck B, Saum K, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. *Clin Epigenetics*. 2016;8:64.
27. Horvath S, Langfelder P, Kwak S, Aaronson J, Rosinski J, Vogt TF, et al. Huntington's disease accelerates epigenetic aging of human brain and disrupts DNA methylation levels. *Aging*. 2016;8:1485–512.
28. Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, et al. Menopause accelerates biological aging. *Proc Natl Acad Sci U S A*. 2016;113:9327–32.
29. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biol*. 2016;17:171.
30. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai P, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging*. 2016;8:1844–65.
31. Simpkin AJ, Hemani G, Suderman M, Gaunt TR, Lyttleton O, Mcardle WL, et al. Prenatal and early life influences on epigenetic age in children: a study of mother-offspring pairs from two cohort studies. *Hum Mol Genet*. 2016;25:191–201.
32. Gao X, Zhang Y, Breitling LP, Brenner H. Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. *Oncotarget*. 2016;7:46878–89.
33. Breitling LP, Saum K, Perna L, Schöttker B, Holleczeck B, Brenner H. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. *Clin Epigenetics*. 2016;8:21.
34. Ward-Caviness CK, Nwanaji-Enwerem JC, Wolf K, Wahl S, Colicino E, Trevisi L, et al. Long-term exposure to air pollution is associated with biological aging. *Oncotarget*. 2016;7:74510–25.
35. Levine AJ, Quach A, Moore DJ, Achim CL, Sontorniyomkij V, Masliah E, et al. Accelerated epigenetic aging in brain is associated with pre-mortem HIV-associated neurocognitive disorders. *J Neuro-Oncol*. 2016;22:366–75.
36. Armstrong NJ, Mather KA, Thalamuthu A, Wright MJ, Trollor JN, Ames D, et al. Aging, exceptional longevity and comparisons of the Hannum and Horvath epigenetic clocks. *Epigenomics*. 2017;9:689–700.
37. McKinney BC, Lin H, Ding Y, Lewis DA, Sweet RA. DNA methylation age is not accelerated in brain or blood of subjects with schizophrenia. *Schizophr Res*. 2018;196:39–44.
38. Wolf EJ, Logue MW, Stoop TB, Schichman SA, Stone A, Sadeh N, et al. Accelerated DNA methylation age: associations with PTSD and mortality. *Psychosom Med*. 2017. <https://doi.org/10.1097/PSY.0000000000000506>.
39. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging*. 2017;9:419–46.
40. Binder AM, Corvalan C, Mericq V, Pereira A, Santos JL, Horvath S, et al. Faster ticking rate of the epigenetic clock is associated with faster pubertal development in girls. *Epigenetics*. 2018;13:85–94.
41. Dugué P, Bassett JK, Joo JE, Baglietto L, Jung C, Wong EM, et al. Association of DNA methylation-based biological age with health risk factors and overall and cause-specific mortality. *Am J Epidemiol*. 2018;187:529–38.
42. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*. 2018;10:573–91.
43. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, et al. DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biology*. 2015;16:25.
44. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*. 2018;19:371–84.
45. Zhang Q, Vallerga CL, Walker RM, Lin T, Henders AK, et al. Improved prediction of chronological age from DNA methylation limits it as a biomarker of ageing. *bioRxiv* 327890; doi: <https://doi.org/10.1101/327890>.
46. 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, Haroutunian V, Kaminsky Z, Joachim C, Powell J, Lovestone S, Bennett DA, Schalkwyk LC, Mill J. Cross-tissue methylomic profiling implicates cortical deregulation of ANKI in Alzheimer's disease neuropathology. *GSE59685 Gene Expression Omnibus*. 2014 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59685>.
47. Lynn, Peter & Knies, Gundi. Understanding Society Quality Profile (Waves 1–5). 2016 <https://doi.org/10.13140/RG.2.1.2389.2887>.
48. Pidsley R, Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*. 2013;14:293.
49. Gorrie-Stone TJ, Smart MC, Saffari A, Malki K, Hannon E, Burrage J, et al. Bigmelon: tools for analysing large DNA methylation datasets. Kelso J, editor. *Bioinformatics*. 2018; <https://doi.org/10.1093/bioinformatics/bty713>.
50. Huynh JL, Garg P, Sharp AJ, Casaccia P. DNA methylation differences between multiple sclerosis and controls in frontal lobe white matter. *GSE40360 Gene Expression Omnibus*. 2013 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE40360>.
51. Ladd-Acosta C, Feinberg AP. DNA methylation in postmortem brain samples from autism cases and controls. *GSE53162 Gene Expression Omnibus*. 2013 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53162>.
52. Ladd-Acosta C, Hansen KD, Briem E, Fallin MD, Kaufmann WE, Feinberg AP. Common DNA methylation alterations in multiple brain regions in autism. *Mol Psychiatry*. 2014;19:862–71.
53. Horvath S, Levine AJ. DNA methylation levels of various brain regions from HIV+ and HIV- subjects. *GSE59457 Gene Expression Omnibus*. 2015 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59457>.
54. Pidsley R, Viana J, Hannon E, Spiers H, Troakes C, Al-Sarraj S, Mechawar N, Turecki G, Schalkwyk L, Bray N, Mill J. Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *GSE61380 Gene Expression Omnibus*. 2014 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61380>.
55. Huynh JL, Garg P, Thin TH, Yoo S, Dutta R, Trapp BD, et al. Epigenome-wide differences in pathology-free regions of multiple sclerosis-affected brains. *Nat Neurosci*. 2014;17:121–30.
56. Horvath S, Levine AJ. Cerebellum Illumina Inf 450 DNA methylation samples from HIV+ and HIV- human subjects. *GSE67748 Gene Expression Omnibus*. 2015 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67748>.
57. Horvath S, Levine AJ. Cerebellum Illumina Inf 450 DNA methylation samples from HIV+ and HIV- human subjects. *GSE67749 Gene Expression Omnibus*. 2015 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67749>.
58. Schizophrenia-associated methylomic variation: molecular signatures of disease and polygenic risk burden across multiple brain regions (cerebellum DBCBB). *GSE89702 Gene Expression Omnibus*. 2016 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE89702>.
59. Pidsley R, Viana J, Hannon E, Spiers H, Troakes C, Al-Sarraj S, Mechawar N, Turecki G, Schalkwyk L, Bray N, Mill J. Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *GSE61431 Gene Expression Omnibus*. 2014 <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE61431>.
60. Viana J, Hannon E, Dempster E, Pidsley R, Macdonald R, Knox O, et al. Schizophrenia-associated methylomic variation: molecular signatures of disease and polygenic risk burden across multiple brain regions. *Hum Mol Genet*. 2016;26:210–25.
61. Hughes A, Smart M, Gorrie-Stone T, Hannon E, Mill J, Bao Y, Burrage J, Schalkwyk L, Kumari M, Benzeval M. EGAS00001002836 European Genome-phenome Archive. 2018 <https://www.ebi.ac.uk/ega/studies/EGAS00001002836>.

## Publisher's Note

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

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

