

# Improved prediction of chronological age from DNA methylation limits it as a biomarker of ageing

Qian Zhang<sup>1</sup>, Costanza L. Vallergera<sup>1</sup>, Rosie M Walker<sup>2</sup>, Tian Lin<sup>1</sup>, Anjali K. Henders<sup>1</sup>, Grant W. Montgomery<sup>1</sup>, Ji He<sup>3</sup>, Dongsheng Fan<sup>3</sup>, Javed Fowdar<sup>4</sup>, Martin Kennedy<sup>5</sup>, Toni Pitcher<sup>6,7</sup>, John Pearson<sup>5</sup>, Glenda Halliday<sup>8</sup>, John B. Kwok<sup>8</sup>, Ian Hickie<sup>8</sup>, Simon Lewis<sup>8</sup>, Tim Anderson<sup>6,7</sup>, Peter A. Silburn<sup>9</sup>, George D. Mellick<sup>4</sup>, Sarah E. Harris<sup>2,10</sup>, Paul Redmond<sup>10</sup>, Alison D. Murray<sup>11</sup>, David J. Porteous<sup>2,10</sup>, Christopher S. Haley<sup>12</sup>, Kathryn L. Evans<sup>2</sup>, Andrew M. McIntosh<sup>2,10,13</sup>, Jian Yang<sup>1,9</sup>, Jacob Gratten<sup>1,9</sup>, Riccardo E. Marioni<sup>2,10</sup>, Naomi R. Wray<sup>1,9</sup>, Ian J. Deary<sup>10,14</sup>, Allan F. McRae<sup>1,15</sup>, Peter M. Visscher<sup>1,9,15</sup>

<sup>1</sup> Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland 4072, Australia

<sup>2</sup> Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU

<sup>3</sup> Department of Neurology, Peking University, Third Hospital, No. 49, North Garden Road, Haidian District, 100191, Beijing, China

<sup>4</sup> Griffith Institute for Drug Discovery (GRIDD), Griffith University, Brisbane, Australia

<sup>5</sup> Department of Pathology and Biomedical Science, University of Otago, Christchurch, New Zealand

<sup>6</sup> New Zealand Brain Research Institute, Christchurch, New Zealand

<sup>7</sup> Department of Medicine, University of Otago, Christchurch, New Zealand

<sup>8</sup> Brain and Mind Centre, Sydney Medical School, The University of Sydney, Sydney, Australia

<sup>9</sup> Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia

<sup>10</sup> Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ

<sup>11</sup> Aberdeen Biomedical Imaging Centre, Lilian Sutton Building, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK

<sup>12</sup> MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU

<sup>13</sup> Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, EH10 5HF, UK

<sup>14</sup> Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ

<sup>15</sup> Jointly supervised the study

1

## 2 **Abstract**

3

4 DNA methylation is associated with age. The deviation of age predicted from DNA  
5 methylation from actual age has been proposed as a biomarker for ageing. However, a  
6 better prediction of chronological age implies less opportunity for biological age. Here we  
7 used 13,661 samples (from blood and saliva) in the age range of 2 to 104 years from 14  
8 cohorts measured on Illumina HumanMethylation450/EPIC arrays to perform prediction  
9 analyses. We show that increasing the sample size achieves a smaller prediction error and  
10 higher correlations in test datasets. We demonstrate that smaller prediction errors provide  
11 a limit to how much variation in biological ageing can be captured by methylation and  
12 provide evidence that age predictors from small samples are prone to confounding by cell  
13 composition. Our predictor shows a similar or better performance in non-blood tissues  
14 including saliva, endometrium, breast, liver, adipose and muscle, compared with Horvath's  
15 across-tissue age predictor.

16

17 **Key words:** DNA methylation, age prediction, best linear unbiased prediction, elastic net

18

## 19 **Introduction**

20 Ageing as a complex biological phenomenon is related to diseases and mortality <sup>1,2</sup>, and  
21 chronological age has been widely used as a marker of ageing due to ease and accuracy of  
22 measurement <sup>1</sup>. However, chronological age is not necessarily a good predictor of biological  
23 ageing since individuals with the same chronological age can vary in health, especially in  
24 later life <sup>3</sup>. Therefore, biomarkers of ageing have become popular as they can indicate the  
25 presence or severity of some disease states <sup>4,5</sup>. In 2013, Hannum *et al.* and Horvath built age  
26 predictors based on DNA methylation and implemented them as biomarkers of ageing <sup>6,7</sup>.  
27 DNA methylation as a part of the epigenome plays an essential role in the regulation of gene

1 expression in the human body<sup>8,9</sup>. Unlike DNA which is (mostly) stable across the lifetime of  
 2 an individual, DNA methylation is dynamic, and previous studies have discovered a number  
 3 of CpG sites associated with chronological age<sup>10-12</sup>. The age predictor developed by Hannum  
 4 et al. was based on 482 blood samples with methylation measured on the Illumina 450K  
 5 methylation arrays, and they reported a correlation of 0.91 and a Root Mean Square Error  
 6 (RMSE) of 4.9 years in their test set<sup>6</sup>. Horvath's age predictor was based on 8,000 samples  
 7 from different tissues and cell types, and probes of these samples were from the Illumina  
 8 27K DNA methylation arrays. He reported a correlation of 0.96 and a Median Absolute  
 9 Deviation (MAD) of 3.6 years in the test set. Age Acceleration Residuals (AAR) is defined as  
 10 the residuals from regressing predicted age on chronological age. It has been reported to be  
 11 associated with mortality, obesity and other complex traits<sup>13-16</sup>.

12

13 According to its definition, AAR is the prediction error in a chronological age predictor.  
 14 Although previous survival analysis showed a significant association between AAR and  
 15 mortality<sup>13</sup>, AAR was found not to be a mitotic clock<sup>17</sup>. Therefore, whether the significance  
 16 of that association is inflated (e.g. by potential confounders in the error) needs to be  
 17 investigated. To study the use of predicted age from DNA methylation as a biomarker of  
 18 ageing, we calculated AAR based on age predictors with different prediction accuracy and  
 19 investigated the relationship between prediction accuracy and the significance of AAR in  
 20 survival analysis. We also investigated the effect of training sample size and statistical  
 21 method on age prediction.

22

23 In the present study, we built DNA methylation-based age predictors by integrating 13,661  
 24 samples (13,402 from blood and 259 from saliva) measured on 450K DNA methylation

arrays and Illumina EPIC (850K) arrays. Two approaches were evaluated: Elastic Net<sup>18</sup> and Best Linear Unbiased Prediction (BLUP)<sup>19</sup>. We discussed the implications of our results for the scope and utility of DNA methylation based age predictor as a biomarker for biological ageing. We also explored the factors that explain prediction accuracy. Finally, the performance of predictors on samples from tissues other than blood was investigated.

## Results

### Availability of DNA methylation in age prediction

We downloaded eight datasets from the public domain and used six datasets from our own studies (**Table 1**). All data underwent identical quality control criteria before statistical analyses (**Material and Methods**).

### Estimation of variation in age from using all probes

We used the unrelated individuals from our two largest datasets (GS, N = 2,586, SGPD, N = 1,299) to estimate the proportion of the observed variation in age that is explained when fitting all probes simultaneously, using a mixed linear model analogous to estimating heritability from SNP data<sup>20</sup>. The proportion of variance of age explained by DNA methylation was close to 1 in both cohorts (proportion explained = 1, SE = 0.0036, REML analysis using the software package OSCA<sup>21</sup> in GS, and 0.99 in SGPD, SE = 0.058), indicating a perfect age predictor can in principle be developed based on DNA methylation data if all probe associations are estimated without error. To demonstrate that this result is not caused by a violation of assumptions, we undertook a permutation test using the same cohorts. We shuffled the ages across individuals and found that DNA methylation did not

1 explain any significant amount of variation in GS (proportion explained = 0, SE = 0.0030) and  
2 SGPD (proportion explained = 0.0079, SE = 0.013).

3

#### 4 **Building multiple age predictors**

5 To build our age predictors, we collected 14 cohorts and used a common set of 319,607  
6 probes that passed quality control (**Material and Methods**) in all cohorts. We randomly  
7 combined 1 to 13 cohorts as a training set, and used the remaining cohorts as test sets. We  
8 repeated this step 65 times to generate different training sets with various sample sizes and  
9 age spectrum (**Material and Methods**). We implemented two estimates to evaluate the  
10 performance of our age predictors: (1) correlation between predicted age and chronological  
11 age in the test data set; (2) Root Mean Square Error (RMSE) of the predicted age in the test  
12 data set. Correlation indicates the strength of a linear relationship between the predicted  
13 age and chronological age and RMSE reveals the variation of the difference between  
14 predicted and chronological age. Two methods, namely Elastic Net<sup>18</sup> and BLUP<sup>19</sup> were used.  
15 Elastic Net was previously used by Horvath<sup>7</sup> and Hannum et al.<sup>6</sup> to build their age predictors  
16 and BLUP was used to predict age in Peters et al.<sup>22</sup>. These methods differ in how they select  
17 probes that are associated with age and how their effects are estimated. Results show that  
18 both methods have a decrease of RMSE (**Figure 1**) and an increase of correlation  
19 (**Supplementary Figure 1**) when the training sample size increased. The smallest RMSE  
20 based on Elastic Net was 2.04 years. This method gave better results with RMSE relative to  
21 BLUP for small training sample size, although the difference with BLUP became smaller with  
22 increased sample size (**Supplementary Figure 2**).

23

#### 24 **Prediction accuracy and biological ageing**

1 The difference between predicted age from the Hannum/Horvath predictors and  
 2 chronological age (AAR) was found to be associated with all-cause mortality in later life<sup>13</sup>. To  
 3 investigate the relationship between the significance of this association and the prediction  
 4 accuracy of the predictor, we examined the association between AAR and mortality using  
 5 the updated data in Marioni et al.<sup>13</sup>. These data were from two cohorts: LBC1921 (wave  
 6 one, N = 436, N<sub>deaths</sub> = 386) and LBC1936 (wave one, N = 906, N<sub>deaths</sub> = 214) (**Materials and**  
 7 **Methods**). Age predictors excluding LBC1921/LBC1936 as part of the training set (sample  
 8 size ranges from 335 to 12,710) were used. We observed a decrease of the test statistics for  
 9 the effect of AAR on mortality from the survival analysis (Cox regression) with increasing  
 10 sample size in training data set (**Figure 2**). No significant ( $P < 0.05$ ) associations between  
 11 AAR and mortality was found based on the largest training sample size in either LBC1921 or  
 12 LBC1936 using BLUP or Elastic Net (**Table 2 and Supplementary Table 1**). In contrast, results  
 13 based on the age predictors of Hannum and Horvath were significant ( $P < 0.05$ , **Table 2**).

14

# 15 **Enrichment analysis on AAR associated probes**

16 Variation in cellular compositions is known to be associated with both DNA methylation<sup>23</sup>  
 17 and mortality<sup>24</sup>, which suggests it could be a confounder in the survival analysis. To  
 18 investigate whether AAR is affected by cellular composition, we applied an epigenome-wide  
 19 association study (EWAS) on AAR from different predictors. For each predictor, AAR  
 20 associated ( $P < 0.05/319,607$ ) probes were selected. We found that the selected CpG sites  
 21 from age predictors of Hannum and Horvath were enriched in the probes that show  
 22 heterogeneity in DNA methylation across cell types (cellular heterogeneity probes)<sup>25</sup> (**Table**  
 23 **3**), indicating AAR in these predictors was associated with variation in cellular composition.  
 24 We also observed a decrease of the odds ratio of the enrichment test with the increase of

1 training sample size for both Elastic Net and BLUP based age predictors (**Figure 3**). No  
2 significant enrichment was found for the age predictors based on the largest training sample  
3 size (**Table 3**).

4

5 To examine the effect of variation in cellular content in the association between AAR and  
6 mortality, we re-ran the survival analysis based on AAR adjusting for white blood cell (WBC)  
7 counts (basophils, eosinophils, monocytes, lymphocytes, and neutrophils) (**Materials and**  
8 **Methods**). A decrease of the test statistics (from survival analysis) after correcting for the  
9 WBC counts was observed, especially when the training sample size is small (**Supplementary**  
10 **Figure 3**). The effect of AAR on survival is reduced the most (the changes of test statistics  
11 from survival analysis are largest) when adding WBC to the models using the Hannum and  
12 Horvath age predictors (**Supplementary Figure 3**). After adjustment for WBC, none of the  
13 associations remained significant ( $P < 0.05$ ) except for the association in LBC1936 based on  
14 the predictor of Horvath ( $P = 0.032$ ). Nevertheless, the significance of this association did  
15 not pass the Bonferroni corrected P value threshold ( $P < 0.05/4$ ) (**Table 2**). These results  
16 suggest that the significant associations between AAR and mortality is biased due to the  
17 existence of confounders like WBC counts, and that improved prediction accuracy of the age  
18 predictor reduces the effect of these confounders in the survival analysis.

19

## 20 **Other factors related to chronological age prediction**

21 Although the potential effect of confounders limit the “epigenetic clock” to be a biomarker  
22 of ageing, a chronological age predictor with good prediction accuracy would be a useful  
23 tool in forensics and/or other fields where chronological age is needed but not observed. To  
24 determine the factors that explain prediction accuracy, we examined the contribution of

1 age ranges (including absolute age difference between training and test set ( $Age_{diff}$ ) and  
 2 standard deviation of age ( $Age_{sd}$ ) of the training set) to the RMSE/correlation of the  
 3 prediction results in the test set by estimating the effect of  $Age_{diff}$ ,  $Age_{sd}$  and sample size in  
 4 the training set on the prediction accuracy jointly (**Material and Methods**). Results showed  
 5 that RMSE was significantly associated ( $P < 0.05$ ) with training sample size in 13 (out of 14)  
 6 cohorts based on BLUP predictors, confirming that increasing the sample size leads to  
 7 smaller prediction errors (**Figure 1**). In addition, eight out of 14 cohorts had a significant ( $P <$   
 8  $0.05$ ) and positive  $Age_{diff}$  effect, indicating similar ages between training and test set results  
 9 in better prediction accuracy (**Supplementary Table 2**). Five cohorts were found to have a  
 10 statistically significant ( $P < 0.05$ )  $Age_{sd}$  effect on RMSE, suggesting the prediction accuracy  
 11 benefits from a larger age range of the samples in the training set. Similar results were  
 12 found based on Elastic Net (**Supplementary Table 3**). In addition, we did not observe any  
 13 steady improvement using power-transformed Beta values (**Supplementary Figure 4**,  
 14 **Materials and Methods**), the M values of DNA methylation or the arcsine square root  
 15 transformed Beta values (**Supplementary Figure 5, Materials and Methods**).

16

17 There is a complex correlation structure in DNA methylation, and the effective number of  
 18 independent methylation probes was previously reported to be around 200<sup>26</sup>, indicating a  
 19 dense correlation structure. To compare the prediction performance between using the full  
 20 probe set (319,607 probes) and a pruned probe set (128,405 probes) (**Material and**  
 21 **methods**), we applied the same cross-validation steps to both probe sets using BLUP and  
 22 Elastic Net. We identified a higher RMSE and a lower correlation for the pruned set  
 23 (**Supplementary Figure 6**), indicating a loss of information when using fewer methylation  
 24 probes for prediction. In addition, we found that probes in the age predictors of Hannum



1 and Horvath were not necessary for age prediction. Compared with these two predictors,  
2 better prediction accuracy can still be observed based on the probe set without the probes  
3 from these predictors (**Supplementary Figure 7, Material and methods**), consistent with  
4 widespread correlation among probes.

5

## 6 **Age prediction in non-blood tissues**

7 The majority of our samples are from blood, and we observed a significant improvement in  
8 the prediction results for the samples from saliva when more blood samples were included  
9 in the training set (**Figure 1, Supplementary Figure 1**). This increase is expected since  
10 samples from saliva were reported to exhibit more than 80% contamination by immune  
11 cells<sup>27</sup>. To quantify whether our predictor has a good performance in non-blood tissues, we  
12 downloaded 13 data sets (**Supplementary Table 4**) that contain samples from other tissues.  
13 We compared the performance of our predictor (samples are from blood and saliva, age  
14 predictor generated by Elastic Net) with Horvath's age predictor (samples from multiple  
15 tissues, using Elastic Net) in these cohorts and found that our predictor has better  
16 performance in samples from endometrium and saliva. On the other hand, Horvath's age  
17 predictor outperformed our predictor in samples from brain (**Figure 4**). Their performance  
18 in other tissues (breast, liver, adipose and muscle) were similar, even though training  
19 samples in our predictor are not from these tissues. These results demonstrate that our  
20 predictor can also be used to predict the chronological age of samples from non-blood  
21 tissues.

22

## 23 **Discussion**

1 We investigated the relationship between the prediction accuracy of an age predictor and  
 2 its application as a biomarker of ageing. Age predictors with various prediction performance  
 3 were built based on datasets with different sample sizes (ranging from  $n=335$  to 13566). We  
 4 ran survival analysis (based on age acceleration residuals AAR) using samples from LBC1921  
 5 and LBC1936, with AAR calculated using different age predictors. We observed a decrease in  
 6 the significance of association between mortality and AAR with the improvement of the age  
 7 predictor. No significant ( $P < 0.05$ ) associations were found based on the age predictor with  
 8 the largest training sample size (**Table 2**), suggesting the improved prediction of  
 9 chronological age from DNA methylation limits it as a biomarker of ageing.

10

11 We found potential effects of confounders in the association between AAR and mortality in  
 12 the age predictors of Hannum and Horvath. The AAR associated probes from the age  
 13 predictors of Horvath and Hannum were enriched in CpG sites showing DNA methylation  
 14 heterogeneity across cell types, suggesting that AAR from these predictors is affected by  
 15 variation in cellular composition. The further sensitivity analysis confirmed that, although  
 16 the AAR from these two predictors are associated with mortality in LBC1921 and LBC1936,  
 17 no significant ( $P < 0.05/4$ ) association was observed after adjusting for white blood cell  
 18 counts. This demonstrates that although the Hannum and Horvath age predictors appear to  
 19 capture differences in biological ageing between people of the same age, these effects are  
 20 largely driven by differences in cellular makeup of the samples, and limits their usage as a  
 21 marker of biological ageing.

22

23 We also examined the factors that can affect the accuracy of chronological age prediction,  
 24 including the effect of the training sample size, the age range of the training samples, the

1 number of probes used and the statistical methods utilised. We found a positive association  
2 between the training sample size and the prediction accuracy in test sets. Our predictors  
3 showed substantially improved prediction accuracy compared to using the estimated  
4 coefficients previously reported by Hannum <sup>6</sup> and Horvath <sup>7</sup> in blood samples. Most of this  
5 improvement appears to come from simply increasing the experimental sample size in the  
6 training set. We also found that increased similarity of ages between samples in the training  
7 and test data set can improve the prediction accuracy in the test sets (**Supplementary**  
8 **Tables 2 and 3**). We provide estimated effect sizes on chronological age from the largest  
9 training set of 13,566 individuals for both Elastic Net and BLUP in **Supplementary Table 5**.  
10  
11 Notwithstanding the highly correlated pattern of DNA methylation across the genome, we  
12 observed a decline of prediction accuracy when using a correlation pruned probe set, so  
13 that including more probes in the training model is beneficial, especially when the training  
14 sample size is small (**Supplementary Figure 6**). The improvement of prediction accuracy  
15 could be explained by the decrease of noise effect (such as batch effects) of DNA  
16 methylation in age prediction since using more probes can reduce the unexpected effects of  
17 the noise. It could also be caused by the existence of many probes with a small correlation  
18 with age and the cumulative effect of these may be lost when using a pruned set of probes.  
19  
20 Although most of the samples in our age predictor are from blood, it showed good out-of-  
21 sample prediction performance in samples from non-blood tissues. Compared with  
22 Horvath's age predictor, we observed larger correlations (between predicted age and  
23 chronological age) in samples from saliva and endometrium, but smaller correlations in  
24 samples from brain. These smaller correlations are expected since a large proportion (23.4%)

1 of training samples in Horvath's age predictor are from brain. Moreover, these two  
2 predictors have similar performance in other tissues, which implies that our age predictor is  
3 also useful in samples from non-blood tissues.

4  
5 Our results have several implications for the utility of DNA methylation patterns of age as  
6 biomarkers of ageing. From the REML analysis on the SGPD and GS cohorts we estimated  
7 that almost 100% of variation in chronological age in those samples could be effectively  
8 captured by all the DNA methylation probes on the arrays. For prediction, this implies that  
9 for a very large training set a near-perfect predictor of chronological age can be built. Our  
10 results showing that larger sample sizes lead to more accurate prediction is consistent with  
11 this implication. It is clear that DNA methylation measured in blood is associated with  
12 environmental exposures such as smoking, sex and BMI<sup>28-30</sup>. In addition, "age acceleration",  
13 the difference between actual age and that predicted from methylation, has been reported  
14 to be associated with a number of outcomes, including mortality<sup>13,16</sup>. However, there is  
15 currently no good DNA-methylation-based estimator of an individual's "epigenetic clock"  
16 that is free from confounders (e.g., white blood cell counts) and from prediction error  
17 caused by other factors (e.g., measurement error). The difference between actual and  
18 predicted age contains both a prediction error term based on unknown factors and possible  
19 effects of confounders. These confounders could bias the results when using "epigenetic  
20 clock" as a biomarker of ageing, for example the association between "age acceleration"  
21 and mortality is confounded by the variation in cellular composition.

22

## 23 **Methods**

### 24 **Data**

1 We collected 14 data cohorts with samples measured on the DNA methylation 450K chips  
2 and Illumina EPIC (850K) arrays (**Table 1**), eight of which were from the public domain and  
3 six datasets from the investigators. Details of the BSGS and LBC cohorts can be found in  
4 Powell et al.<sup>31</sup> and Deary et al.<sup>32,33</sup>. GS is a population and family based cohort recruited  
5 through the NHS Scotland general practitioner research network<sup>34,35</sup>. The SGPD cohort is  
6 from a collaborative research project on systems genomics of Parkinson's Disease. Similarly,  
7 the MND cohort is from a systems genomics study of Motor Neuron Disease in Chinese  
8 subjects (see descriptions in Benyamin et al.<sup>36</sup>). For the purpose of this study, age at sample  
9 collection was the focus, disease status and ethnicity of individuals were not considered in  
10 any cohort. DNA methylation Beta value at each probe was used for analysis.

11  
12 A total of 319,607 probes (No Pruned Set) passed our quality control and 128,405 probes  
13 (Pruned Set) were retained after pruning based upon the pairwise correlation of probes (see  
14 next section). To test the performance of age predictors in non-blood tissues, we  
15 downloaded 13 cohorts from GEO database with accession ID GSE61431 (brain)<sup>37</sup>, GSE59685  
16 (brain)<sup>38</sup>, GSE80970 (brain), GSE101961 (breast)<sup>39</sup>, GSE108213 (breast), GSE48325 (liver)<sup>40</sup>,  
17 GSE61257 (adipose)<sup>41</sup>, GSE61258 (liver)<sup>41</sup>, GSE61259 (breast)<sup>41</sup>, GSE88883 (breast)<sup>42</sup>,  
18 GSE90060 (endometrium)<sup>43</sup>, GSE92767 (saliva)<sup>44</sup>, GSE99029 (saliva)<sup>45</sup>.

## 20 **Quality Control**

21 All the samples were measured on either the Illumina HumanMethylation450 arrays or  
22 Illumina EPIC arrays. Probes with call rate less than 0.95 were removed, and probes found to  
23 contain SNPs or potentially cross-hybridizing to different locations were excluded from

1 further analysis<sup>46</sup>. After combining all the samples from different cohorts, a set of 319,607  
 2 probes remained (called No Pruned set). Pruning was performed by removing one of two  
 3 probes on the same chromosome when their correlation ( $R^2$ ) was higher than 0.2; this  
 4 resulted in a set of 128,405 probes (called Pruned set). Both sets were used for further  
 5 analysis. DNA methylation Beta value was standardized by removing the mean value and  
 6 divided by the standard deviation for each sample.

7

## 8 **Selection of DNA methylation cohorts**

9 We collected 14 different cohorts in total, including a single cohort (GSE78874) measured in  
 10 saliva rather than blood tissue. Since DNA methylation is sensitive to batch effects, cell type  
 11 and tissue type<sup>23</sup>, we applied a PCA analysis (using probes from the No Pruned Set) on the  
 12 samples from these 14 cohorts to assess the presence of any “outlier” cohorts (i.e. cohorts  
 13 with a low prediction accuracy from the age predictor based on the other cohorts). All the  
 14 cohorts were closely matched with the exception of GSE78874 and GS (**Supplementary**  
 15 **Figure 8**). Samples in GSE78874 were from saliva instead of blood, and the samples in GS  
 16 were measured using Illumina EPIC arrays instead of 450K DNA methylation arrays. To  
 17 investigate if this difference could potentially adversely influence performance in age  
 18 prediction for these two cohorts, we used a “leave-one-cohort-out” strategy to leave these  
 19 two cohorts out as the test set separately and built the age predictor based on the  
 20 remaining cohorts. We found both of them to have good prediction accuracy (GS:  $R = 0.98$ ,  
 21  $RMSE = 3.52$ , GSE78874:  $R = 0.88$ ,  $RMSE = 5.39$ ), indicating a small difference between these  
 22 two cohorts and other cohorts in age prediction. We used all cohorts for subsequent  
 23 analyses.

1

## 2 **Generation of training set**

3 We generated training sets from the 14 cohorts. Each training set has a certain number of  
4 cohorts ranging between 1 to 13. For each number, we repeated random sampling five  
5 times. In total, 65 ( $13 \times 5$ ) training sets were generated.

6

## 7 **Proportion of variance of chronological age explained by DNA methylation**

8 The GS and SGPD samples were used in estimating the proportion of variance of  
9 chronological age explained by DNA methylation. Among the 5,101 samples in the GS cohort,  
10 a subset of 2,586 unrelated individuals, with a genetic relationship coefficient below 0.05  
11 and with no shared nuclear family environment were considered for the analysis. 1,299  
12 unrelated (genetic relationship coefficient  $< 0.05$ ) individuals with available age information  
13 in SGPD were selected. Variance of age was estimated by the REML method implemented in  
14 OSCA<sup>21</sup>.

15

## 16 **Prediction algorithm**

17 We compared the age prediction performance of two methods, namely Elastic Net and  
18 BLUP. Both methods are based on a linear regression:

$$Y = \sum \beta_i X_i + e$$

19 where  $Y$  is the chronological age,  $X_i$  is the DNA methylation of probe  $i$  and  $e$  is the Gaussian  
20 noise.

21 Elastic Net is a regularized regression method<sup>18</sup>, and its objective function is defined as:

22

$$L(\alpha, \beta) = ||Y - X\beta||^2 + \lambda \left( \frac{1-\alpha}{2} ||\beta||_2^2 + \alpha ||\beta||_1 \right)$$

1 where  $\alpha$  and lambda are regularisation parameters.  $||\beta||_1$  is defined as  $\sum_{i=1}^n |\beta_i|$  and  $||\beta||_2^2$   
2 equals  $\sum_{i=1}^n \beta_i^2$ , with  $n$  the number of probes.  $\alpha$  is set to 0.5 and  $\lambda$  is chosen based on cross-  
3 validation. We used the implementation of Elastic Net from the Python package glmnet<sup>47</sup>.

4

5 BLUP is special case of ridge regression with a fixed  $\lambda$ .

$$6 \quad \hat{\beta} = (X'X + \lambda I)^{-1} X'Y \text{ with } \lambda = \frac{\sigma_e^2}{\sigma_u^2},$$

7  $\sigma_u^2$  the variance of the effect size of the probe set, and  $\sigma_e^2$  the variance of the residuals. We  
8 used the R package rrBLUP<sup>48</sup> to build the age predictor, and  $\sigma_u^2$  and  $\sigma_e^2$  were estimated using  
9 the REML analysis implemented in this package.

10

## 11 **Survival analysis**

12 We followed the same analysis approach as previously described<sup>13</sup>. Briefly, Cox proportional  
13 hazards regression models were used to detect the association between the AAR and  
14 mortality with age at sample collection and sex as the covariates. AAR is defined as the  
15 difference between  $m_{age}$  and chronological age, where  $m_{age}$  is the predicted age correcting  
16 for plate, array, position on the array, and hybridisation date (all treated as fixed effect  
17 factors), all of which could be confounder in survival analysis. Additional adjustments of AAR  
18 were made for WBC counts measured on the same blood samples that were analysed for  
19 methylation. Hazard ratios for AAR were expressed per five years of methylation age  
20 acceleration (**Table 2**) and per standard deviation of methylation age acceleration  
21 (**Supplementary Table 1**), respectively. Cox models were performed utilizing the 'survival'  
22 library<sup>49</sup> in R. Samples from wave one of LBC1921 and LBC1936 were used in this analysis.

23



# 1 Transformation of DNA methylation

2 There are non-linear patterns in age-related DNA methylation<sup>50</sup>. To investigate if  
 3 transformed data can remove the nonlinearity and hence improve the prediction accuracy,  
 4 we selected eight DNA methylation cohorts with sample size larger than 600 to evaluate the  
 5 impact of data transformation: LBC1921, LBC1936, GS, BSGS, SGPD, MND, GSE40279 and  
 6 GSE42861. For each cohort, we randomly selected 70% of the samples as training set, and  
 7 the remaining 30% were used as test set. Only 50,000 randomly selected probes were used  
 8 for computational efficiency. Power parameter  $\lambda$  (ranges from 0.1 to 2 with 0.05 as the  
 9 interval) was used to transform the original Beta value of DNA methylation BV to  $BV^\lambda$ . Only  
 10 BLUP was used for age prediction because of its low bias. DNA methylation M value and  
 11 arcsine square root transformed methylation Beta value were also used to compare to raw  
 12 DNA methylation Beta value in prediction accuracy.

13

# 14 Age prediction without probes from age predictors of Horvath and Hannum

15 We compared the probes selected by Elastic Net (based on 13,566 training samples) with  
 16 those in Horvath's and Hannum's age predictors. 11 out of the 514 probes in our analysis  
 17 were identified in Horvath's age predictor and 30 in Hannum's age predictor. In addition, we  
 18 estimated the squared correlation ( $R^2$ ) of DNA methylation between probes selected by  
 19 Elastic Net and probes from the age predictor of Hannum/Horvath. We found 11 (Elastic  
 20 Net-Hannum) and 10 (Elastic Net-Horvath) pairs with an  $R^2$  larger than 0.5 (**Supplementary**  
 21 **Figure 9**), indicating that most of the probes selected by Elastic Net are not strongly  
 22 correlated with those in the other two predictors. To quantify whether the probes in the  
 23 Hannum and Horvath predictors were necessary for age prediction, we re-built our age  
 24 predictors by excluding these probes. No difference in prediction accuracy was found before

1 and after removing these probes for the BLUP based method (**Supplementary Figure 10**).  
 2 The prediction accuracy decreased for the Elastic Net based method; however, its  
 3 performance was still better than when using the Hannum and Horvath age predictors  
 4 (**Supplementary Figure 7**).

5

## 6 **Acknowledgements**

7 This research was supported by the Australian Research Council (DP160102400), the  
 8 Australian National Health and Medical Research Council (1078037, 1078901, 1103418,  
 9 1107258, 1127440 and 1113400), and the Sylvia & Charles Viertel Charitable Foundation.  
 10 Riccardo Marioni was supported by Alzheimer's Research UK Major Project Grant [ARUK-  
 11 PG2017B-10]. Generation Scotland received core support from the Chief Scientist Office of  
 12 the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council  
 13 [HR03006]. Genotyping and DNA methylation profiling of the GS:SFHS samples was carried  
 14 out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility,  
 15 Edinburgh, Scotland and was funded by the Medical Research Council UK and the Wellcome  
 16 Trust (Wellcome Trust Strategic Award "STratifying Resilience and Depression Longitudinally"  
 17 ((STRADL) Reference 104036/Z/14/Z).

18

## 19 **Author contributions**

20 A.F.M and P.M.V conceived and designed the experiments. Q.Z performed all statistical  
 21 analyses. Q.Z, A.F.M and P.M.V wrote the paper. R.E.M, I.J.D, J.Y and N.W.R advised on  
 22 statistical methodology, C.L.V, R.M.W, T.L, A.K.H, G.W. M, J.H, D.F, J.F, M.K, T.P, J.P, G.H, J.B.  
 23 K, I.H, S.L, T.A, P.A.S, G.D.M, S.E.H, P.R, A.D.M, D.J.P, C.S.H, K.L.E, A.M.M, J.G contributed  
 24 data. All authors read and approved the final manuscript.

25

## 26 **Competing interests**

27 The authors declare no competing financial interests.

# Reference

- 1 Fries, J. F. Aging, natural death, and the compression of morbidity. *B World Health Organ* **80**, 245-250 (2002).
- 2 Gavrilov, L. A. & Gavrilova, N. S. The reliability theory of aging and longevity. *J Theor Biol* **213**, 527-545, doi:10.1006/jtbi.2001.2430 (2001).
- 3 Baker, G. T. & Sprott, R. L. Biomarkers of aging. *Experimental gerontology* **23**, 223-239 (1988).
- 4 McClearn, G. E. Biomarkers of age and aging. *Experimental gerontology* **32**, 87-94 (1997).
- 5 Johnson, T. E. Recent results: Biomarkers of aging. *Experimental Gerontology* **41**, 1243-1246, doi:10.1016/j.exger.2006.09.006 (2006).
- 6 Hannum, G. *et al.* Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular cell* **49**, 359-367 (2013).
- 7 Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**, 3156 (2013).
- 8 Razin, A. & Riggs, A. D. DNA methylation and gene function. *Science* **210**, 604-610 (1980).
- 9 Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213 (1986).
- 10 Hernandez, D. G. *et al.* Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Human molecular genetics* **20**, 1164-1172 (2011).
- 11 Florath, I., Butterbach, K., Müller, H., Bewerunge-Hudler, M. & Brenner, H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Human molecular genetics* **23**, 1186-1201 (2013).
- 12 Bell, J. T. *et al.* Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS genetics* **8**, e1002629 (2012).
- 13 Marioni, R. E. *et al.* DNA methylation age of blood predicts all-cause mortality in later life. *Genome biology* **16**, 25 (2015).
- 14 Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *P Natl Acad Sci USA* **111**, 15538-15543, doi:10.1073/pnas.1412759111 (2014).
- 15 Wolf, E. J. *et al.* Accelerated DNA methylation age: Associations with PTSD and neural integrity. *Psychoneuroendocrino* **63**, 155-162, doi:10.1016/j.psyneuen.2015.09.020 (2016).
- 16 Chen, B. H. *et al.* DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)* **8**, 1844-1865, doi:10.18632/aging.101020 (2016).
- 17 Yang, Z. *et al.* Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biol* **17**, 205, doi:10.1186/s13059-016-1064-3 (2016).
- 18 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).
- 19 Robinson, G. K. That BLUP is a good thing: the estimation of random effects. *Statistical science*, 15-32 (1991).

- 1 20 Yang, J. *et al.* Common SNPs explain a large proportion of the heritability for human  
2 height. *Nat Genet* **42**, 565-569, doi:10.1038/ng.608 (2010).
- 3 21 Zhang, F. *et al.* OSCA: a tool for omic-data-based complex trait analysis. *bioRxiv*,  
4 doi:10.1101/445163 (2018).
- 5 22 Peters, M. J. *et al.* The transcriptional landscape of age in human peripheral blood.  
6 *Nature communications* **6** (2015).
- 7 23 Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture  
8 distribution. *Bmc Bioinformatics* **13**, doi: 10.1186/1471-2105-13-86 (2012).
- 9 24 Ruggiero, C. *et al.* White blood cell count and mortality in the Baltimore Longitudinal  
10 Study of Aging. *J Am Coll Cardiol* **49**, 1841-1850, doi:10.1016/j.jacc.2007.01.076  
11 (2007).
- 12 25 Jaffe, A. E. & Irizarry, R. A. Accounting for cellular heterogeneity is critical in  
13 epigenome-wide association studies. *Genome Biology* **15**, doi:10.1186/gb-2014-15-2-  
14 r31 (2014).
- 15 26 Shah, S. *et al.* Improving phenotypic prediction by combining genetic and epigenetic  
16 associations. *The American Journal of Human Genetics* **97**, 75-85 (2015).
- 17 27 Zheng, S. C. *et al.* A novel cell-type deconvolution algorithm reveals substantial  
18 contamination by immune cells in saliva, buccal and cervix. *Epigenomics* **10**, 925-940,  
19 doi:10.2217/epi-2018-0037 (2018).
- 20 28 Joehanes, R. *et al.* Epigenetic signatures of cigarette smoking. *Circulation:*  
21 *Cardiovascular Genetics* **9**, 436-447 (2016).
- 22 29 Mendelson, M. M. *et al.* Association of body mass index with DNA methylation and  
23 gene expression in blood cells and relations to cardiometabolic disease: a mendelian  
24 randomization approach. *PLoS medicine* **14**, e1002215 (2017).
- 25 30 Yousefi, P. *et al.* Sex differences in DNA methylation assessed by 450 K BeadChip in  
26 newborns. *BMC Genomics* **16**, 911, doi:10.1186/s12864-015-2034-y (2015).
- 27 31 Powell, J. E. *et al.* The Brisbane Systems Genetics Study: genetical genomics meets  
28 complex trait genetics. *PLoS One* **7**, e35430, doi:10.1371/journal.pone.0035430  
29 (2012).
- 30 32 Deary, I. J., Gow, A. J., Pattie, A. & Starr, J. M. Cohort profile: the Lothian Birth  
31 Cohorts of 1921 and 1936. *International journal of epidemiology* **41**, 1576-1584  
32 (2011).
- 33 33 Taylor, A. M., Pattie, A. & Deary, I. J. Cohort Profile Update: The Lothian Birth  
34 Cohorts of 1921 and 1936. *Int J Epidemiol*, doi:10.1093/ije/dyy022 (2018).
- 35 34 Smith, B. H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study  
36 (GS:SFHS). The study, its participants and their potential for genetic research on  
37 health and illness. *Int J Epidemiol* **42**, 689-700, doi:10.1093/ije/dys084 (2013).
- 38 35 Navrady, L. B. *et al.* Cohort Profile: Stratifying Resilience and Depression  
39 Longitudinally (STRADL): a questionnaire follow-up of Generation Scotland: Scottish  
40 Family Health Study (GS:SFHS). *Int J Epidemiol* **47**, 13-14g, doi:10.1093/ije/dyx115  
41 (2018).
- 42 36 Benyamin, B. *et al.* Cross-ethnic meta-analysis identifies association of the GPX3-  
43 TNIP1 locus with amyotrophic lateral sclerosis. *Nat Commun* **8**, 611,  
44 doi:10.1038/s41467-017-00471-1 (2017).
- 45 37 Pidsley, R. *et al.* Methylomic profiling of human brain tissue supports a  
46 neurodevelopmental origin for schizophrenia. *Genome Biol* **15**, 483,  
47 doi:10.1186/s13059-014-0483-2 (2014).

1 38 Lunnon, K. *et al.* Methylomic profiling implicates cortical deregulation of ANK1 in  
2 Alzheimer's disease. *Nat Neurosci* **17**, 1164-1170, doi:10.1038/nn.3782 (2014).  
3 39 Song, M. A. *et al.* Landscape of genome-wide age-related DNA methylation in breast  
4 tissue. *Oncotarget* **8**, 114648-114662, doi:10.18632/oncotarget.22754 (2017).  
5 40 Ahrens, M. *et al.* DNA methylation analysis in nonalcoholic fatty liver disease  
6 suggests distinct disease-specific and remodeling signatures after bariatric surgery.  
7 *Cell Metab* **18**, 296-302, doi:10.1016/j.cmet.2013.07.004 (2013).  
8 41 Horvath, S. *et al.* Obesity accelerates epigenetic aging of human liver. *Proc Natl Acad*  
9 *Sci U S A* **111**, 15538-15543, doi:10.1073/pnas.1412759111 (2014).  
10 42 Johnson, K. C., Houseman, E. A., King, J. E. & Christensen, B. C. Normal breast tissue  
11 DNA methylation differences at regulatory elements are associated with the cancer  
12 risk factor age. *Breast Cancer Res* **19**, doi:10.1186/s13058-017-0873-y (2017).  
13 43 Kukushkina, V. *et al.* DNA methylation changes in endometrium and correlation with  
14 gene expression during the transition from pre-receptive to receptive phase. *Sci Rep*  
15 **7**, 3916, doi:10.1038/s41598-017-03682-0 (2017).  
16 44 Hong, S. R. *et al.* DNA methylation-based age prediction from saliva: High age  
17 predictability by combination of 7 CpG markers. *Forensic Sci Int Genet* **29**, 118-125,  
18 doi:10.1016/j.fsigen.2017.04.006 (2017).  
19 45 Gopalan, S. *et al.* Trends in DNA Methylation with Age Replicate Across Diverse  
20 Human Populations. *Genetics* **206**, 1659-1674, doi:10.1534/genetics.116.195594  
21 (2017).  
22 46 Price, E. M. *et al.* Additional annotation enhances potential for biologically-relevant  
23 analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenet*  
24 *Chromatin* **6**, doi: 10.1186/1756-8935-6-4 (2013).  
25 47 Balakumar, B. J., Hastie, T., Friedman, J., Tibshirani, Simon, N. R. *Glmnet for Python*,  
26 <[http://www.stanford.edu/~hastie/glmnet\\_python/](http://www.stanford.edu/~hastie/glmnet_python/)> (2016).  
27 48 Endelman, J. B. Ridge regression and other kernels for genomic selection with R  
28 package rrBLUP. *The Plant Genome* **4**, 250-255 (2011).  
29 49 A Package for Survival Analysis in S v. 2.38 (2015).  
30 50 Johnson, N. D. *et al.* Non-linear patterns in age-related DNA methylation may reflect  
31 CD4(+) T cell differentiation. *Epigenetics-U S* **12**, 492-503,  
32 doi:10.1080/15592294.2017.1314419 (2017).  
33 51 Horvath, S. *et al.* An epigenetic clock analysis of race/ethnicity, sex, and coronary  
34 heart disease. *Genome Biology* **17**, doi:10.1186/s13059-016-1030-0 (2016).  
35 52 Horvath, S. *et al.* Aging effects on DNA methylation modules in human brain and  
36 blood tissue. *Genome Biol* **13**, R97, doi:10.1186/gb-2012-13-10-r97 (2012).  
37 53 Liu, Y. *et al.* Epigenome-wide association data implicate DNA methylation as an  
38 intermediary of genetic risk in rheumatoid arthritis. *Nat Biotechnol* **31**, 142-147,  
39 doi:10.1038/nbt.2487 (2013).  
40 54 Li, Y. *et al.* An Epigenetic Signature in Peripheral Blood Associated with the  
41 Haplotype on 17q21.31, a Risk Factor for Neurodegenerative Tauopathy. *Plos Genet*  
42 **10**, doi: 10.1371/journal.pgen.1004211 (2014).  
43  
44

1 Table 1: Description of DNA methylation cohorts

| COHORT <sup>1</sup>             | SAMPLE SIZE <sup>2</sup> | NUMBER OF SAMPLES WITH VALID AGE | MEAN AGE(SD) | AGE RANGE    | SOURCE | DISEASE                                                                                                                                                      |
|---------------------------------|--------------------------|----------------------------------|--------------|--------------|--------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>LBC1921</b> <sup>32,33</sup> | 692                      | 692                              | 82.3 (4.3)   | [77.8, 90.6] | blood  | Not Available                                                                                                                                                |
| <b>LBC1936</b> <sup>32,33</sup> | 2326                     | 2326                             | 72.4 (2.8)   | [67.7, 77.7] | blood  | Not Available                                                                                                                                                |
| <b>BSGS</b> <sup>31</sup>       | 614                      | 614                              | 21.4 (14.1)  | [9.9, 74.9]  | blood  | Not Available                                                                                                                                                |
| <b>SGPD</b>                     | 1962                     | 1556                             | 67.2 (9.5)   | [23.0,104.0] | blood  | Parkinson's Disease: 988, Control: 974                                                                                                                       |
| <b>MND</b> <sup>36</sup>        | 695                      | 600                              | 45.2 (15.0)  | [17.0,76.0]  | blood  | Motor Neuron Disease (MND): 497, Control: 198                                                                                                                |
| <b>GS</b> <sup>34,35</sup>      | 5101                     | 5100                             | 48.5 (14.0)  | [18.0,94.5]  | blood  | Not Available                                                                                                                                                |
| <b>GSE72775</b> <sup>51</sup>   | 335                      | 335                              | 70.2 (10.3)  | [36.5,90.5]  | blood  | Not Available                                                                                                                                                |
| <b>GSE78874</b> <sup>51</sup>   | 259                      | 259                              | 68.8(9.7)    | [36.0,88.0]  | saliva | Not Available                                                                                                                                                |
| <b>GSE72773</b> <sup>51</sup>   | 310                      | 310                              | 65.6 (13.9)  | [35.1,91.9]  | blood  | Not Available                                                                                                                                                |
| <b>GSE72777</b> <sup>51</sup>   | 46                       | 46                               | 14.7 (10.4)  | [2.2,35.0]   | blood  | Not Available                                                                                                                                                |
| <b>GSE41169</b> <sup>52</sup>   | 95                       | 95                               | 31.6 (10.3)  | [18.0,65.0]  | blood  | Schizophrenia:62, Control:33                                                                                                                                 |
| <b>GSE40279</b> <sup>6</sup>    | 656                      | 656                              | 64.0 (14.7)  | [19.0,101.0] | blood  | Not Available                                                                                                                                                |
| <b>GSE42861</b> <sup>53</sup>   | 689                      | 689                              | 51.9 (11.8)  | [18.0,70.0]  | blood  | Rheumatoid Arthritis:354, Control:335                                                                                                                        |
| <b>GSE53740</b> <sup>54</sup>   | 384                      | 383                              | 67.8(9.6)    | [34.0,93.0]  | blood  | Alzheimer's Disease:15, Corticobasal Degeneration:1, Frontotemporal Dementia (FTD):121, FTD/MND:7, Progressive Supranuclear Palsy:43, Control:193, Unknown:4 |

2  
3  
4 <sup>1</sup> LBC = Lothian Birth Cohort; BSGS = Brisbane Systems Genomics Study; SGPD = Systems Genomic of Parkinson's Disease  
5 consortium; MND = Motor Neuron Disease cohort; GS = Generation Scotland. Cohorts with prefix GSE are from the GEO  
6 database.

7 <sup>2</sup> The number of samples in each cohort. Some samples in LBC were measured from the same individual but at different  
chronological age.

1 Table 2: Summary details of two LBC cohorts and the relationship between all-cause  
2 mortality and predicted age from different methods (before and after cell counts correction)

|                                                          | LBC1921 wave one        | LBC1936 wave one         |
|----------------------------------------------------------|-------------------------|--------------------------|
| N                                                        | 436                     | 906                      |
| N <sub>deaths</sub>                                      | 386                     | 214                      |
| Chronological Age: mean (SD) <sup>1</sup>                | 79.1 (0.6)              | 69.5 (0.8)               |
| Before cell counts correction                            |                         |                          |
| Hannum: mean (SD)                                        | 80.3 (6.2)              | 71.3 (5.7)               |
| Hannum: Hazard Ratio (P-value, 95% CI) <sup>2</sup>      | 1.12 (0.016, 1.02-1.23) | 1.18 (0.020, 1.02-1.37)  |
| Horvath: mean (SD)                                       | 73.8 (6.9)              | 66.1 (6.4)               |
| Horvath: Hazard Ratio (P-value, 95% CI)                  | 1.09 (0.038, 1.00-1.20) | 1.19 (0.0022, 1.06-1.32) |
| Elastic Net: mean (SD) <sup>3</sup>                      | 77.4 (3.6)              | 72.5 (3.2)               |
| Elastic Net: Hazard Ratio (P-value, 95% CI)              | 1.08 (0.38, 0.91-1.27)  | 1.00 (0.96, 0.79-1.28)   |
| BLUP: mean (SD) <sup>3</sup>                             | 77.3 (3.3)              | 72.5 (2.8)               |
| BLUP: Hazard Ratio (P-value, 95% CI)                     | 1.20 (0.066, 0.99-1.46) | 1.25 (0.12, 0.95-1.64)   |
| After cell counts correction                             |                         |                          |
| Hannum: Hazard Ratio (P-value, 95% CI)                   | 1.10 (0.057, 1.00-1.21) | 1.11 (0.15, 0.96-1.29)   |
| Horvath: Hazard Ratio (P-value, 95% CI)                  | 1.07 (0.13, 0.98-1.17)  | 1.14 (0.032, 1.01-1.28)  |
| Elastic Net: Hazard Ratio (P-value, 95% CI) <sup>3</sup> | 1.07 (0.39, 0.91-1.27)  | 1.03 (0.79, 0.82-1.31)   |
| BLUP: Hazard Ratio (P-value, 95% CI) <sup>3</sup>        | 1.21 (0.05, 1.00-1.48)  | 1.21 (0.17, 0.92-1.60)   |

<sup>1</sup> Mean (predicted) age and its standard deviation.

<sup>2</sup> Hazard Ratio, P-value and 95% confidence interval from the survival analysis based on the predicted age. Hazard Ratios were expressed per 5 years of methylation age acceleration.

<sup>3</sup> Both results of Elastic Net and BLUP were based on the age predictor with largest training sample size (sample size = 10,411 for LBC1936 and sample size = 12,710 for LBC1921).

1 Table 3: Enrichment test on the AAR associated CpG sites from different methods

2

|                          | Number of significant associations (P < 0.05/319,607) | $\lambda_{\text{median}}$ | Number of CpG sites showing significant cellular heterogeneity | Odds ratio (P-value) <sup>1</sup> |
|--------------------------|-------------------------------------------------------|---------------------------|----------------------------------------------------------------|-----------------------------------|
| Hannum                   | 12,015                                                | 3.6                       | 4,958                                                          | 3.85 (P < 2.2×10 <sup>-16</sup> ) |
| Horvath                  | 18,847                                                | 5.4                       | 5,955                                                          | 2.53 (P < 2.2×10 <sup>-16</sup> ) |
| Elastic Net <sup>2</sup> | 159                                                   | 2.1                       | 21                                                             | 0.78 (P = 0.33)                   |
| BLUP <sup>2</sup>        | 793                                                   | 2.6                       | 130                                                            | 1.00 (P = 1.0)                    |

3

4

<sup>1</sup> The odd ratio for the enrichment of EWAS significant CpG sites in the probe set showing significant cellular heterogeneity.

5

<sup>2</sup> Both results of Elastic Net and BLUP were based on the age predictor with largest training sample size (sample size = 10,411 for LBC1936 and sample size = 12,710 for LBC1921).

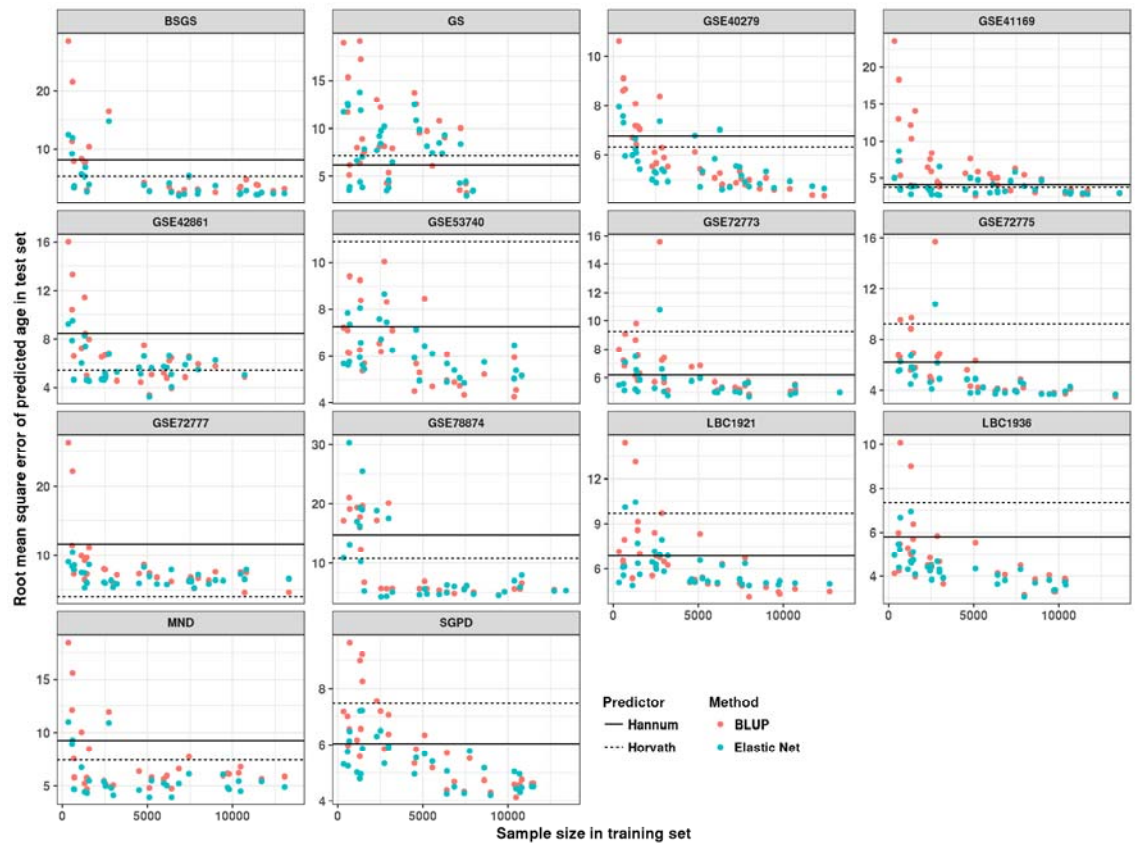
6

7

8

9





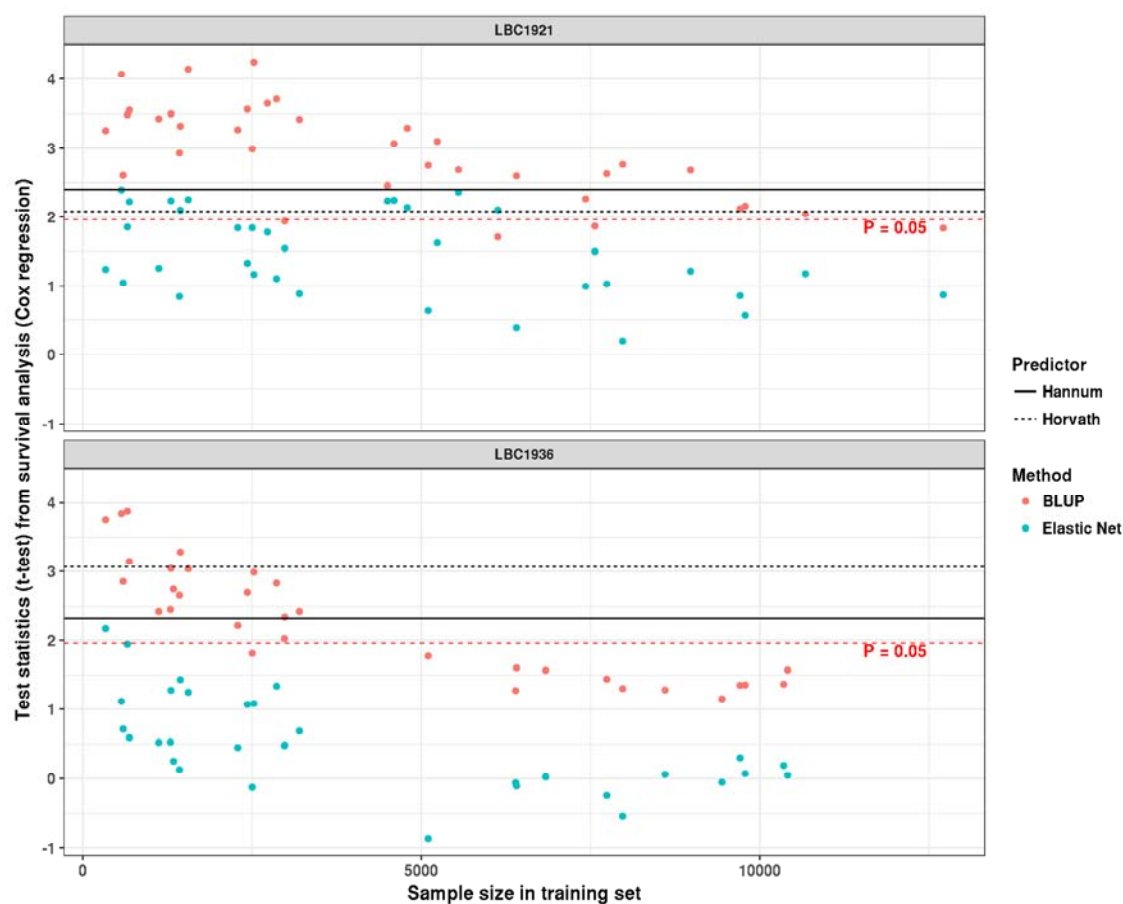


Figure 2: Relationship between the training sample size and the test statistics (t-test) from the association between age acceleration residual (AAR) and mortality. Each point represents the test statistic from the survival analysis based on the predicted ages from predictors with different training sample sizes.

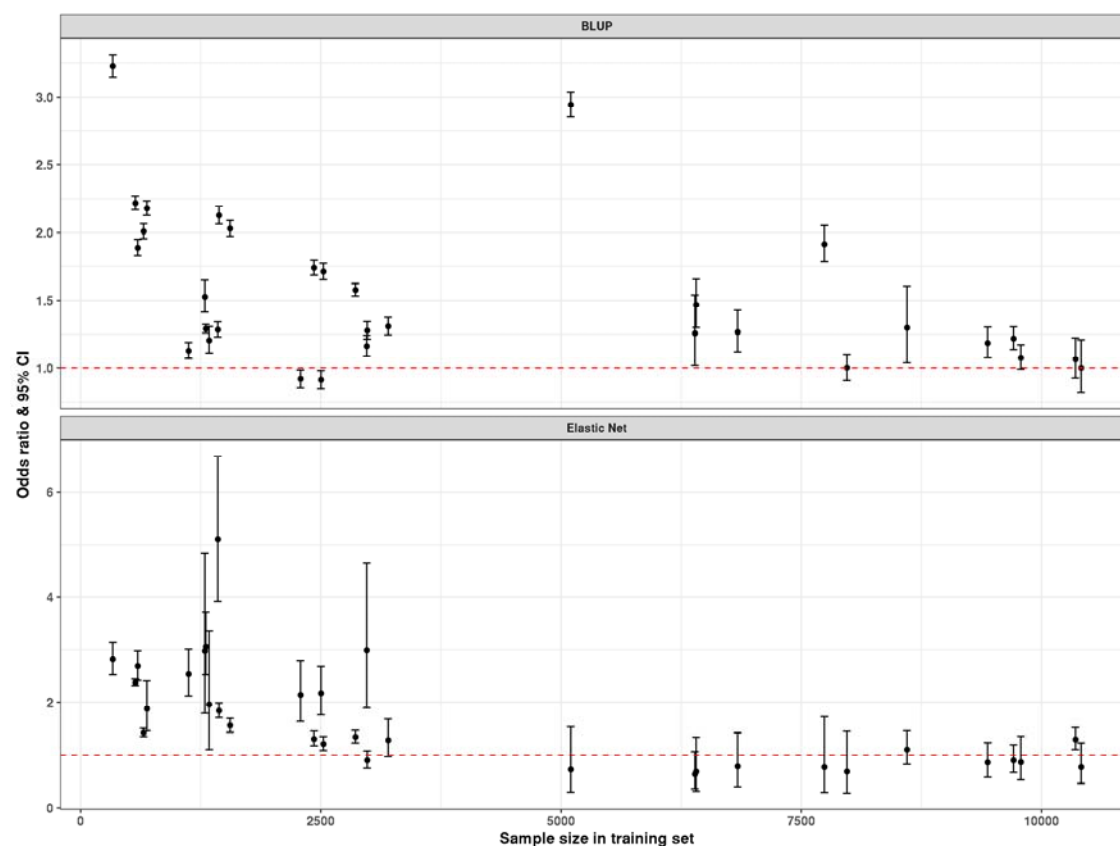


Figure 3: The change of odds ratio from the enrichment test with the increase of training sample size. The enrichment test examines whether AAR associated CpG sites are enriched in probes with cellular heterogeneity.

