DNA methylation age acceleration and risk factors for Alzheimer's disease[1]

Daniel L McCartney[a], Anna J Stevenson[a], Rosie M Walker[a], Jude Gibson[b], Stewart W Morris[a], Archie Campbell[a], Alison D Murray[c], Heather C Whalley[b], David J Porteous[a,d], Andrew M McIntosh[a,b,d], Kathryn L Evans[a,d], Ian J Deary[d,e], Riccardo E Marioni[a,d,*]

a. Medical Genetics Section, Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU

b. Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, EH10 5HF

c. Aberdeen Biomedical Imaging Centre, University of Aberdeen, Aberdeen, AB25 2ZD

d. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ

e. Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ

* Corresponding Author: riccardo.marioni@ed.ac.uk

Word Count (abstract): 150

Word Count (body): 3,067

Running Title: DNA Methylation Age and Alzheimer's risk factors

Keywords: DNA methylation, epigenetic clock, Alzheimer's disease

Declarations of interest: None

[1] Abbreviations used: AD = Alzheimer's disease, BMI = Body mass index, EEAA = Extrinsic epigenetic age acceleration, GS:SHFS = Generation Scotland: Scottish family health study, HBP = High blood pressure, HDL = high-density lipoprotein, IEAA = Intrinsic epigenetic age acceleration, PGRS = Polygenic risk score, SIMD = Scottish index of multiple deprivation, T2D = Type-2 diabetes

Abstract

INTRODUCTION: The 'epigenetic clock' is a DNA methylation-based estimate of biological age and is correlated with chronological age – the greatest risk factor for Alzheimer's disease (AD). Genetic and environmental risk factors exist for AD, several of which are potentially modifiable. Here, we assess the relationship associations between the epigenetic clock and AD risk factors.

METHODS: Linear mixed modelling was used to assess the relationship between age acceleration (the residual of biological age regressed onto chronological age) and AD risk factors relating to cognitive reserve, lifestyle, disease, and genetics in the Generation Scotland study (n=5,100).

RESULTS: We report significant associations between the epigenetic clock and BMI, total:HDL cholesterol ratios, socioeconomic status, and smoking behaviour (Bonferroni-adjusted $P<0.05$).

DISCUSSION: Associations are present between environmental risk factors for AD and age acceleration. Measures to modify such risk factors might improve the risk profile for AD and the rate of biological ageing. Future longitudinal analyses are therefore warranted.

1. Introduction

DNA methylation is an epigenetic modification typically characterised by the addition of a methyl group to a ctosine-guanine dinucleotide (CpG). Both genetic and environmental factors influence DNA methylation, which in turn can regulate gene expression [1]. The "epigenetic clock" is an estimation of biological age derived from DNA methylation data, and is strongly correlated with chronological age [2]. From biological age, a measure of age acceleration can be obtained based on the difference between an individual's biological (estimated) and chronological (actual) age. Age acceleration has been linked to a range of age-related health outcomes, including increased Alzheimer's disease (AD) pathology [3], reduced cognitive and physical fitness [4], and an increase in all-cause mortality [5]. The epigenetic clock has therefore been proposed as a biomarker of ageing and may be predictive of age-related disorders, such as dementia [6].

Dementia is one of the leading global health concerns of the 21st century. The most common form of dementia is AD. Lifestyle factors such as smoking have been linked to an increased risk of AD [7], as have disease-related factors including type 2 diabetes (T2D) and high blood pressure (HBP) [8, 9]. Moreover, resilience to age-related brain changes (e.g. cognitive reserve), has been linked to AD risk [10]. Factors such as educational attainment and socioeconomic status have been proposed as proxy measures of cognitive reserve, and lower levels of these are established AD risk factors [11, 12]. Genetic studies of AD have revealed several risk factors [13], with the *APOE* locus (encoding apolipoprotein E) being among the strongest [14].

A recent review [15] suggested that up to a third of cases of all-cause dementia might be delayed by actively addressing its modifiable risk factors. In the current study, we investigate the relationship between epigenetic age acceleration and both genetic and environmental AD risk factors in over 5,000 individuals from the Generation Scotland cohort. Two measures of age acceleration were assessed: intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA). These measures are described in greater detail in the methods section. Briefly, IEAA is a measure of age acceleration that is independent of age-related changes in the cellular composition of blood [16], whereas EEAA captures the age-related functional decline of the immune system. Age is the strongest risk factor for AD [17] and epigenetic age is a robust predictor of chronological age. We therefore hypothesise that individuals with poorer profiles for AD risk factors display accelerated ageing in comparison to those with more favourable profiles.

2. Methods

2.1 The Generation Scotland Cohort

Details of the Generation Scotland Scottish Family Health Study (GS:SFHS) have been described previously [18, 19]. Briefly, the cohort comprises 23,960 individuals, each with at least one family member participating in the study. DNA samples were collected for genotype- and DNA methylation-profiling along with detailed clinical, lifestyle, and sociodemographic data. The current study comprised 5,101 individuals from the cohort for whom DNA methylation data were available. A summary of all variables assessed in this analysis is presented in Table 1.

2.2 Ethics

All components of GS:SFHS received ethical approval from the NHS Tayside Committee on Medical Research Ethics (REC Reference Number: 05/S1401/89). GS:SFHS has also been granted Research Tissue Bank status by the Tayside Committee on Medical Research Ethics (REC Reference Number: 10/S1402/20), providing generic ethical approval for a wide range of uses within medical research.

2.3 GS:SHFS DNA methylation

Genome-wide DNA methylation was profiled in blood samples from 5,200 individuals using the Illumina HumanMethylationEPIC BeadChips. Quality control was conducted in R [20]. ShinyMethyl [21] was used to plot the log median intensity of methylated versus unmethylated signal per array with outliers being excluded upon visual inspection. WateRmelon [22] was used to remove (1) samples where ≥1% of CpGs had a detection p-value in excess of 0.05 (2) probes with a beadcount of less than 3 in more than 5 samples, and (3) probes where ≥0.5% of samples had a detection p-value in excess of 0.05. ShinyMethyl was used to exclude samples where predicted sex did not match recorded sex. This left a sample of 5,101 available for analysis.

2.4 Calculation of age acceleration

Methylation-based estimates of age were calculated using the online age calculator (https://dnamage.genetics.ucla.edu/) developed by Horvath [23]. Normalised GS:SHFS DNA methylation data were used as input for the algorithm and data were underwent a further round of normalisation by the age calculator. Two measures of age acceleration were calculated: intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA). IEAA is defined as the residual term of a multivariate model regressing estimated Horvath

methylation age [23] on chronological age, fitting counts of naive CD8+ T cells, exhausted CD8+ T cells, plasmablasts, CD4+ T cells, natural killer cells, monocytes, and granulocytes estimated from the methylation data. IEAA therefore does not consider age-related changes in the cellular composition of blood. Conversely, the estimate of EEAA tracks age-related changes in blood cell composition as well as intrinsic epigenetic changes. EEAA is calculated first by calculating a weighted average of Hannum's DNA methylation age [24], and three cell types whose abundance is known to change with age (naïve cytotoxic T-cells, exhausted cytotoxic T-cells, and plasmablasts) using the approach described by Klemera and Doubal [25]. EEAA is defined as the residual term of a univariate model regressing the weighted estimated age on chronological age.

2.5 Definition of AD risk factors

AD risk factors were divided into four categories: cognitive reserve, disease, lifestyle, and genetics. Cognitive reserve factors comprised education years and socioeconomic status as measured by the Scottish index of multiple deprivation (SIMD). Disease-related factors comprised self-reported type 2 diabetes (T2D) status and high blood pressure (HBP) status. Lifestyle factors comprised smoking pack years (defined as packs smoked per day times years as a smoker), body mass index (BMI), high-density lipoprotein (HDL), total cholesterol and total:HDL cholesterol ratio. Genetic factors comprised family history (defined as having a parent or grandparent with AD), AD polygenic risk score (PGRS), and *APOE* genotype.

2.6 Calculation of AD PGRS

PGRS for AD were created for all individuals with genotype data in the GS:SHFS cohort. All autosomal SNPs which passed quality control were included in the calculation of the PGRS for

AD (see Supplementary Information for quality control parameters). PGRS for AD were estimated using summary statistics from an independent GWAS of AD (17,008 cases, 37,154 controls), conducted by the International Genomics of Alzheimer's Project (IGAP) [13]. PGRS were estimated using the PRSice software package, according to previously described protocols [26], with LD threshold and distance threshold for clumping of $R^2 > 0.25$ and 250 kb, respectively. After excluding SNPs within a 500kb region of *APOE*, a score was created for each individual, using all possible remaining SNPs, in accordance with previous GS:SFHS analyses [27].

2.7 Statistical analysis

Mixed-effects models were performed in R [20], assessing the relationship between epigenetic age acceleration (IEAA and EEAA) and factors related to cognitive reserve, disease, lifestyle, and genetics. Sex and risk factors were fitted as fixed effects and a kinship matrix was fitted as a random effect to control for pedigree structure. Models were built using the *lmekin()* function from the coxme R package [28]. Correction for multiple testing was applied separately to IEAA and EEAA-based analyses using the Bonferroni method.

3. Results

3.1 Estimation of epigenetic age

Methylation data from 5,101 individuals were submitted to the online age calculator. One individual was flagged for an ambiguous gender prediction and was omitted from downstream analysis, leaving 5,100 individuals. A summary of chronological and estimated ages in the GS:SHFS cohort is provided in Table 1. Both Horvath's and Hannum's estimates of biological age were strongly correlated with chronological age (r = 0.94 and 0.93, respectively). As

reported previously [29], there was a strong effect of biological sex on age acceleration with men showing greater acceleration than women (Mean EEAA: males = 0.47, females = -0.3, P = 3.58 x 10-12; Mean IEAA: males = 1.13, females = -0.71, P = 8.68 x 10^{-53}).

3.2 Cognitive reserve and epigenetic age acceleration

Two cognitive reserve factors were evaluated for association with age acceleration: socioeconomic status based on the SIMD, and education years (Table 2; Figure 1). No significant associations were present between these factors and IEAA. Nominally significant negative associations (at P<0.05) were observed between education and SIMD and EEAA (-0.08 years per education category, P = 0.048; -0.26 years per SD increase in SIMD, P = 1.9 x 10^{-5}).

3.3 Disease-related risk factors and epigenetic age acceleration

We assessed the relationship between age acceleration and two disease-related risk factors: type 2 diabetes and high blood pressure (Table 2; Figure 1). No significant associations were observed between either measure of epigenetic age acceleration and type 2 diabetes. Individuals with high blood pressure displayed an average of 0.37 years of extrinsic age acceleration (P = 0.02).

3.4 Lifestyle-related risk factors and epigenetic age acceleration

Four factors related to lifestyle were considered: BMI, smoking habits (pack years), HDL, and total cholesterol (Table 2; Figure 1). Higher values of both measures of epigenetic age acceleration were observed with higher BMI (IEAA: 0.06 years per kg/m^2 of BMI, P = 1.6 x

178 10^{-10}; EEAA: 0.05 years per kg/m^2 of BMI, P = 1.2 x 10^{-5}), and more pack years (IEAA: 0.007
179 years per smoking pack year, P = 0.025; EEAA: 0.01 years per smoking pack year, P = 4.5 x
180 10^{-5}). Greater IEAA was associated with lower levels of HDL cholesterol (-0.34 years per
181 mmol/L of HDL. P = 0.015), and higher levels of total cholesterol (0.12 years per mmol/L of
182 total cholesterol, P = 0.015). A significant positive association was present between IEAA and
183 total:HDL cholesterol ratio (beta = 0.168, P = 2.7 x 10^{-4}). There were no significant associations
184 observed between EEAA and any of the three cholesterol-related metrics assessed.

185

186 **3.5 Genetic risk factors and epigenetic age acceleration**

187 Three genetic risk factors for AD were assessed for association with age acceleration: family
188 history, AD PGRS, and *APOE* genotype (Table 2; Figure 1). No significant associations were
189 present between any of the genetic risk factors assessed and either measure of epigenetic age
190 acceleration.

191

192 **3.6 Correction for multiple testing**

193 Applying a Bonferroni correction separately for the IEAA and EEAA regressions
194 (P<(0.05/12)=0.0042) identified significant IEAA associations with BMI and total:HDL
195 cholesterol ratio (BMI adjusted P = 2.6 x 10^{-9}; total:HDL cholesterol ratio adjusted P = 4.3 x
196 10^{-3}); and significant EEAA associations with SIMD, BMI, and smoking (SIMD adjusted P =
197 3.0 x 10^{-4}; BMI adjusted P = 1.9 x 10^{-4}; smoking adjusted P = 7.2 x 10^{-4}). Of these, higher age
198 acceleration was associated with higher total:HDL cholesterol ratios, BMI levels, smoking
199 levels and social deprivation.

200

4. Discussion

In the current study, we hypothesised that age acceleration might be associated with dementia risk factors in the Generation Scotland cohort. Using both intrinsic (cell-adjusted) and extrinsic (immune system-associated) estimates of epigenetic age acceleration in a cohort of 5,100 individuals, we identified significant associations between multiple dementia risk factors and age acceleration Several of the AD risk factors associated with age acceleration are potentially modifiable lifestyle factors, suggesting the rate of epigenetic ageing can be altered through behavioural changes.

Biological age has been linked to an increased risk of all-cause mortality and is strongly correlated with chronological age [5]. The epigenetic clock has been proposed as a biomarker of ageing as well as a predictor of an individual's health and susceptibility to age-related health outcomes [3,5]. As chronological age increases, so too does the risk of dementia. Individuals with greater age acceleration (i.e. with greater epigenetic age relative to chronological age) have slightly poorer cognitive ability [4] and a modest increase in burden of pathological hallmarks of dementia [3].

Of the risk factors assessed, BMI and smoking levels were associated (at a nominal significance threshold) with both estimates of age acceleration. BMI has previously been associated with an increased risk of dementia and AD when high in middle-age and low in old-age [30, 31]. Consistent with our findings, others have observed an association between higher BMI and increased IEAA and EEAA [32]. Previous studies have failed to find associations between smoking levels and epigenetic age acceleration [16, 33]. However, the effect sizes observed in the present study were roughly comparable with those reported by Gao et al. [33] (IEAA: Gao

et al. Beta = 0.002-0.0039, current study Beta = 0.007; EEAA: Gao et al. Beta = 0.0073-0.0099, current study Beta = 0.014). Our findings of a significant positive association between self-reported smoking and both measures of age acceleration may be attributable to our larger sample size (N = 4,997 individuals compared to maximum N = 978 individuals with smoking data available [33]) although only EEAA was significantly associated with smoking after correction for multiple testing.

In the present study, factors relating to cholesterol were associated with age acceleration based on the intrinsic (cell-adjusted) estimate of epigenetic age acceleration. HDL levels were negatively correlated with epigenetic age acceleration whereas both total cholesterol levels and total:HDL cholesterol ratio were positively correlated with age acceleration. To our knowledge, significant associations between methylation-based estimates of age acceleration and total:HDL cholesterol ratios have not been reported to date. Consistent with our findings, others have observed an association between lower HDL cholesterol and increased age acceleration [32]. A relationship between increased age acceleration and both total and HDL cholesterol levels using a transcriptomic estimate of biological age has also been reported [34]. HDL cholesterol, colloquially known as “good cholesterol”, primarily functions in lipid transport. Higher levels of HDL cholesterol have been linked to a reduction in cardiovascular disease [35], as well as a decreased risk of AD and dementia [36, 37]. Conflicting evidence exists for the association between mid-life levels of total cholesterol and dementia risk [38, 39]. However, studies have consistently reported an inverse association between total cholesterol levels and AD risk in elderly individuals [40, 41, 42]. Longitudinal analyses have revealed different trajectories of BMI in dementia cases compared to controls [31]. Similarly, longitudinal analyses have also indicated that mid-to-late-life trajectories of cholesterol levels are related to both *APOE* genotype [43] and dementia status [44]. *APOE*, a strong genetic risk

250 factor for AD also functions in lipid transport. The association between cholesterol levels and
251 AD risk, coupled with the functions of *APOE* and other genetic risk factors (e.g. *SORL1*) [13]
252 supports a role for lipid metabolism and transport in dementia [45, 46].

253

254 Of the risk factors related to cognitive reserve, both educational attainment and socioeconomic
255 status were associated with EEAA. However, of the two, only socioeconomic status remained
256 significant following Bonferroni correction. Individuals with fewer education years showed
257 increased age acceleration, as did individuals from more deprived socioeconomic backgrounds.
258 Individuals with increased levels of education have displayed delays in the age of onset of
259 dementia [47]. Consistent with our findings, others have reported a similar pattern between
260 EEAA and educational attainment [32, 48]. The biological differences linked to social
261 deprivation are possibly due to the association between socioeconomic status and other, more
262 biologically direct, risk factors for dementia. For example, several lifestyle-related AD risk
263 factors have been shown to be associated with socioeconomic status, including smoking and
264 BMI [49, 50].

265

266 No significant associations were observed between either measure of age acceleration and any
267 of the genetic risk factors assessed. Epigenetic age acceleration effects of environmental factors
268 such as smoking and cholesterol may be more visible in blood due to direct contact with the
269 tissue. Whereas genetic risk factors should be consistent across all tissues, it is possible that
270 they only influence epigenetic age acceleration in cell types where AD pathology is primarily
271 observed (i.e. brain tissue).

272

With a sample size in excess of 5,000 individuals, this is the largest study of DNA methylation-based ageing to date. The use of a comprehensively phenotyped cohort has permitted the assessment of both genetic and environmental AD risk factors and their relationship with epigenetic ageing. This resource is further strengthened by the potential for data linkage to medical records and re-contact of participants, making future longitudinal analyses possible. The cross-sectional design of the current study poses a limitation as it does not permit the assessment of longitudinal changes in age acceleration in response to altered lifestyle habits. However, such a study might be informative in determining whether the trajectory of biological age can be modified through efforts to reduce the risk of AD and other forms of dementia. With the exception of BMI and smoking, significant associations were specific to either IEAA or EEAA. This discordance is possibly due to differences in the two estimates of age acceleration. As described in the methods section, IEAA does not reflect differences in blood cell composition that may be due to age whilst these differences are incorporated into the estimate of EEAA. High blood pressure and both cognitive reserve factors were associated with EEAA, but not IEAA. This may reflect a relationship between these risk factors and immunosenescence. Conversely, the cholesterol-related factors were associated with IEAA but not EEAA, possibly reflecting a relationship between these factors and "pure" epigenetic ageing.

In conclusion, we reported associations between both intrinsic and extrinsic measures of epigenetic age acceleration and environmental AD risk factors. However, no associations were present for the genetic risk factors assessed. At a nominal ($P<0.05$) significance threshold, IEAA was associated with all of the lifestyle-related factors assessed, whereas EEAA was associated with high blood pressure, BMI, smoking, and both cognitive reserve factors assessed. Following Bonferroni correction, BMI, cholesterol ratios, smoking and

298 socioeconomic status remained significantly associated with epigenetic age acceleration. Risk
299 factors such as cholesterol levels, smoking and BMI can be modulated by behavioural changes
300 with regard to exercise, dietary intake and smoking behaviour. The epigenetic clock is a robust
301 predictor of chronological age, and the greatest risk factor for AD is advanced age [17].
302 Individuals displaying accelerated ageing have demonstrated increased AD neuropathology
303 and lower cognitive test scores [3, 4]. In the current study, we observed a relationship between
304 age acceleration and AD risk factors. It is reasonable to suggest that, by improving one's AD
305 risk profile where possible, the process of biological ageing process could be "slowed".

306

307 **Acknowledgements**

308 This work was supported by Alzheimer's Research UK Major Project Grant [ARUK-
309 PG2017B-10]. Generation Scotland received core funding from the Chief Scientist Office of
310 the Scottish Government Health Directorates [CZD/16/6] and the Scottish Funding Council
311 [HR03006]. We are grateful to all the families who took part, the general practitioners and the
312 Scottish School of Primary Care for their help in recruiting them, and the whole Generation
313 Scotland team, which includes interviewers, computer and laboratory technicians, clerical
314 workers, research scientists, volunteers, managers, receptionists, healthcare assistants and
315 nurses. Genotyping of the GS:SFHS samples was carried out by the Genetics Core Laboratory
316 at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the
317 Medical Research Council UK and the Wellcome Trust (Wellcome Trust Strategic Award
318 "STratifying Resilience and Depression Longitudinally" (STRADL) [104036/Z/14/Z]. DNA
319 methylation data collection was funded by the Wellcome Trust Strategic Award
320 [10436/Z/14/Z]. The research was conducted in The University of Edinburgh Centre for
321 Cognitive Ageing and Cognitive Epidemiology (CCACE), part of the cross-council Lifelong

322 Health and Wellbeing Initiative [MR/K026992/1]; funding from the Biotechnology and
323 Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is
324 gratefully acknowledged. CCACE supports Ian Deary, with some additional support from
325 Dementias Platform UK [MR/L015382/1]. AMM and HCW have received support from the
326 Sackler Institute.

327

References

[1] Jaenisch R, Bird A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. Nat Genet 2003;33:245–54. doi:10.1038/ng1089.

[2] Marioni RE, Harris SE, Shah S, McRae AF, von Zglinicki T, Martin-Ruiz C, et al. The epigenetic clock and telomere length are independently associated with chronological age and mortality. Int J Epidemiol 2016;45:424–32. doi:10.1093/ije/dyw041.

[3] Levine ME, Lu AT, Bennett DA, Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging (Albany NY) 2015;7:1198–211. doi:10.18632/aging.100864.

[4] 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 2015a;44:1388–96. doi:10.1093/ije/dyu277.

[5] Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol 2015b;16. doi:10.1186/s13059-015-0584-6.

[6] Degerman S, Josefsson M, Nordin Adolfsson A, Wennstedt S, Landfors M, Haider Z, et al. Maintained memory in aging is associated with young epigenetic age. Neurobiol Aging 2017;55:167–71. doi:10.1016/j.neurobiolaging.2017.02.009.

[7] Durazzo TC, Mattsson N, Weiner MW. Smoking and increased Alzheimer's disease risk: A review of potential mechanisms. Alzheimers Dement 2014;10:S122-45. doi:10.1016/j.jalz.2014.04.009.

[8] Kivipelto M, Helkala E, Laakso MP, Hänninen T, Hallikainen M, Alhainen K, et al. Midlife vascular risk factors and Alzheimer's Disease in later life: Longitudinal, population based study. Bmj 2001;322:1447–51. doi:10.1136/bmj.322.7300.1447.

353 [9] Arvanitakis Z, Wilson RS, Bienias JL, Evans DA, Bennett DA. Diabetes Mellitus and Risk
354 of Alzheimer Disease and Decline in Cognitive Function. Arch Neurol 2004;61:661–6.
355 doi:10.1001/archneur.61.5.661.
356 [10] Stern Y. Cognitive reserve in ageing and Alzheimer's disease. Lancet Neurol
357 2012;11:1006–12. doi:10.1016/S1474-4422(12)70191-6.
358 [11] Jefferson AL, Gibbons LE, Rentz DM, Carvalho JO, Manly J, Bennett DA, et al. A life
359 course model of cognitive activities, socioeconomic status, education, reading ability, and
360 cognition. J Am Geriatr Soc 2011;59:1403–11. doi:10.1111/j.1532-5415.2011.03499.x.
361 [12] Mortimer JA, Graves AB. Education and other socioeconomic determinants of
362 dementia and Alzheimer's disease. Neurology 1993;43:S39–44.
363 [13] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al.
364 Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's
365 disease. Nat Genet 2013;45:1452–8. doi:10.1038/ng.2802.
366 [14] Liu CC, Kanekiyo T, Xu H, Bu G. Apolipoprotein e and Alzheimer disease: Risk,
367 mechanisms and therapy. Nat Rev Neurol 2013;9:106–18. doi:10.1038/nrneurol.2012.263.
368 [15] Livingston G, Frankish H. A global perspective on dementia care: A Lancet
369 Commission. Lancet 2015;386:933–4. doi:10.1016/S0140-6736(15)00078-1.
370 [16] Horvath S, Ritz BR. Increased epigenetic age and granulocyte counts in the blood of
371 Parkinson's disease patients. Aging (Albany NY) 2015;7:1130–42.
372 doi:10.18632/aging.100859.
373 [17] Alzheimer's Association. 2015 Alzheimer's disease facts and figures. Alzheimers
374 Dement 2015;11:332–84. doi:10.1016/j.jalz.2015.02.003.
375 [18] Smith BH, Campbell H, Blackwood D, Connell J, Connor M, Deary IJ, et al. Generation
376 Scotland: The Scottish Family Health Study; a new resource for researching genes and
377 heritability. BMC Med Genet 2006;7. doi:10.1186/1471-2350-7-74.

[19] Smith BH, Campbell A, Linksted P, Fitzpatrick B, Jackson C, Kerr SM, et al. Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. Int J Epidemiol 2013;42:689–700. doi:10.1093/ije/dys084.

[20] R Core Team. R Development Core Team. R A Lang Environ Stat Comput 2017;55:275–86.

[21] Fortin J-P, Fertig E, Hansen K. shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R. F1000Research 2014. doi:10.12688/f1000research.4680.2.

[22] Schalkwyk LC, Pidsley R, Wong CCY. wateRmelon: Illumina 450 methylation array normalization and metrics. R Packag Version 122 2013.

[23] Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14. doi:10.1186/gb-2013-14-10-r115.

[24] Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda SV, et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. Mol Cell 2013;49:359–67. doi:10.1016/j.molcel.2012.10.016.

[25] Klemera P, Doubal S. A new approach to the concept and computation of biological age. Mech Ageing Dev 2006;127:240–8. doi:S0047-6374(05)00265-4 pii]\r10.1016/j.mad.2005.10.004.

[26] Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2015;31:1466–8. doi:10.1093/bioinformatics/btu848.

[27] Marioni RE, Campbell A, Hagenaars SP, Nagy R, Amador C, Hayward C, et al. Genetic Stratification to Identify Risk Groups for Alzheimer's Disease. J Alzheimer's Dis 2017;57:275–83. doi:10.3233/JAD-161070.

[28] Therneau (2018). coxme: Mixed Effects Cox Models. R package version 2.2-7. https://CRAN.R-project.org/package=coxme

[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. doi:10.1186/s13059-016-1030-0.

[30] Tolppanen AM, Ngandu T, Kåreholt I, Laatikainen T, Rusanen M, Soininen H, et al. Midlife and late-life body mass index and late-life dementia: Results from a prospective population-based cohort. J Alzheimer's Dis 2014;38:201–9. doi:10.3233/JAD-130698.

[31] Singh-Manoux A, Dugravot A, Shipley M, Brunner EJ, Elbaz A, Sabia S, et al. Obesity trajectories and risk of dementia: 28 years of follow-up in the Whitehall II Study. Alzheimer's Dement 2017. doi:10.1016/j.jalz.2017.06.2637.

[32] 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 (Albany NY) 2017;9:419–46. doi:10.18632/aging.101168.

[33] 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. doi:10.18632/oncotarget.9795.

[34] Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nat Commun 2015;6. doi:10.1038/ncomms9570.

[35] Ali KM, Wonnerth A, Huber K, Wojta J. Cardiovascular disease risk reduction by raising HDL cholesterol - Current therapies and future opportunities. Br J Pharmacol 2012;167:1177–94. doi:10.1111/j.1476-5381.2012.02081.x.

[36] Reitz C, Tang M-X, Schupf N, Manly JJ, Mayeux R, Luchsinger JA. Association of higher levels of high-density lipoprotein cholesterol in elderly individuals and lower risk

427 of late-onset Alzheimer disease. Arch Neurol 2010;67:1491–7.
428 doi:10.1001/archneurol.2010.297.
429 [37]Zuliani G, Cavalieri M, Galvani M, Volpato S, Cherubini a, Bandinelli S, et al. Relationship
430 between low levels of high-density lipoprotein cholesterol and dementia in the elderly. The
431 InChianti study. J Gerontol A Biol Sci Med Sci 2010;65:559–64.
432 doi:10.1093/gerona/glq026.
433 [38] Solomon A, Kivipelto M, Wolozin B, Zhou J, Whitmer RA. Midlife serum cholesterol
434 and increased risk of Alzheimer's and vascular dementia three decades later. Dement
435 Geriatr Cogn Disord 2009;28:75–80. doi:10.1159/000231980.
436 [39] Mielke MM, Zandi PP, Shao H, Waern M, Östling S, Guo X, et al. The 32-year
437 relationship between cholesterol and dementia from midlife to late life. Neurology
438 2010;75:1888–95. doi:10.1212/WNL.0b013e3181feb2bf.
439 [40] Romas SN, Tang MX, Berglund L, Mayeux R. APOE genotype, plasma lipids
440 lipoproteins, and AD in community elderly. Neurology 1999;53:517–21.
441 doi:10.1212/WNL.53.3.517.
442 [41] Reitz C, Tang M-X, Luchsinger J, Mayeux R. Relation of Plasma Lipids to Alzheimer
443 Disease and Vascular Dementia. Arch Neurol 2004;61:705.
444 doi:10.1001/archneur.61.5.705.
445 [42] Mielke MM, Zandi PP, Sjögren M, Gustafson D, Ostling S, Steen B, et al. High total
446 cholesterol levels in late life associated with a reduced risk of dementia. Neurology
447 2005;64:1689–95. doi:10.1212/01.WNL.0000161870.78572.A5.
448 [43] Downer B, Estus S, Katsumata Y, Fardo DW. Longitudinal trajectories of cholesterol
449 from midlife through late life according to apolipoprotein E allele status. Int J Environ Res
450 Public Health 2014;11:10663–93. doi:10.3390/ijerph111010663.

451 [44] Stewart R, White LR, Xue QL, Launer LJ. Twenty-six-year change in total cholesterol
452 levels and incident dementia: The Honolulu-Asia aging study. Arch Neurol 2007;64:103–
453 7. doi:10.1001/archneur.64.1.103.
454 [45] Reitz C. Dyslipidemia and the risk of Alzheimer's disease. Curr Atheroscler Rep
455 2013;15:307. doi:10.1007/s11883-012-0307-3.
456 [46] Wong MW, Braidy N, Poljak A, Pickford R, Thambisetty M, Sachdev PS.
457 Dysregulation of lipids in Alzheimer's disease and their role as potential biomarkers.
458 Alzheimer's Dement 2017;13:810–27. doi:10.1016/j.jalz.2017.01.008.
459 [47] Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of
460 Dementia over Three Decades in the Framingham Heart Study. N Engl J Med
461 2016;374:523–32. doi:10.1056/NEJMoa1504327.
462 [48] Karlsson Linner R, Marioni RE, Rietveld CA, Simpkin AJ, Davies NM, Watanabe K,
463 et al. An epigenome-wide association study meta-analysis of educational attainment. Mol
464 Psychiatry 2017. doi:10.1038/mp.2017.210.
465 [49] Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and
466 smoking: A review. Ann N Y Acad Sci 2012;1248:107–23. doi:10.1111/j.1749-
467 6632.2011.06202.x.
468 [50] Tyrrell J, Jones SE, Beaumont R, Astley CM, Lovell R, Yaghootkar H, et al. Height,
469 body mass index, and socioeconomic status: mendelian randomisation study in UK
470 Biobank. BMJ 2016:i582. doi:10.1136/bmj.i582.

McCartney et al. Figure 1: Effects of AD risk factors on age acceleration.

Bar plots are separated into four groups of AD risk factors: cognitive reserve, disease, lifestyle, and genetic. Model Beta coefficients (i.e. effect sizes) are presented along the Y-axes while risk factors are presented along the X-axes. Bars are coloured by extrinsic epigenetic age acceleration (EEAA; red) and intrinsic epigenetic age acceleration (IEAA; blue). Error bars show the standard error (SE). Bars accompanied by an asterisk (*) represent measures significantly associated with age acceleration at a Bonferroni $P < 0.05$.

SIMD: Scottish Index of Multiple Deprivation, HBP: High Blood Pressure, T2D: Type 2 Diabetes, BMI: Body Mass Index, HDL: High-Density Lipoprotein cholesterol, AD: Alzheimer's Disease, PGRS: Polygenic Risk Score. The effect sizes for smoking have been scaled to represent per 10 pack years of exposure. All other effect sizes are per unit increase (disease positive for HBP and T2D and positive family history of AD) with the exception of SIMD and AD polygenic risk score (per SD), and *APOE* status, where the effects are relative to the ε3ε3 reference category.

498 **McCartney et al. Table 1:** Summary of variables assessed in the Generation Scotland cohort

499 * The following smoking categories were available: current smoker (N = 939); former smoker,
500 stopped within past 12 months (N = 158); former smoker, stopped more than 12 months ago
501 (N = 1,309); never smoker (N = 2,533). Data were unavailable for 62 participants.

502 † Scottish Index of Multiple Deprivation

503 ‡ Education was measured as an ordinal variable. 0: 0 years, 1: 1-4 years, 2: 5-9 years, 3: 10-
504 11 years, 4: 12-13 years, 5: 14-15 years, 6: 16-17 years, 7: 18-19 years, 8: 20-21 years, 9: 22-
505 23 years, 10: ≥24 years.

506

507

508

509

510

511

512

513

514

515

516

517

518 **McCartney et al. Table 2:** Age acceleration and AD risk factors. Significant associations after
519 accounting for multiple comparisons are highlighted in bold.

520 * Scottish Index of Multiple Deprivation

521 † Education was measured as an ordinal variable. 0: 0 years, 1: 1-4 years, 2: 5-9 years, 3: 10-
522 11 years, 4: 12-13 years, 5: 14-15 years, 6: 16-17 years, 7: 18-19 years, 8: 20-21 years, 9: 22-
523 23 years, 10: ≥24 years.

524 ‡ The following smoking categories were available: current smoker (N = 939); former smoker,
525 stopped within past 12 months (N = 158); former smoker, stopped more than 12 months ago
526 (N = 1,309); never smoker (N = 2,533), and not coded (N = 62).

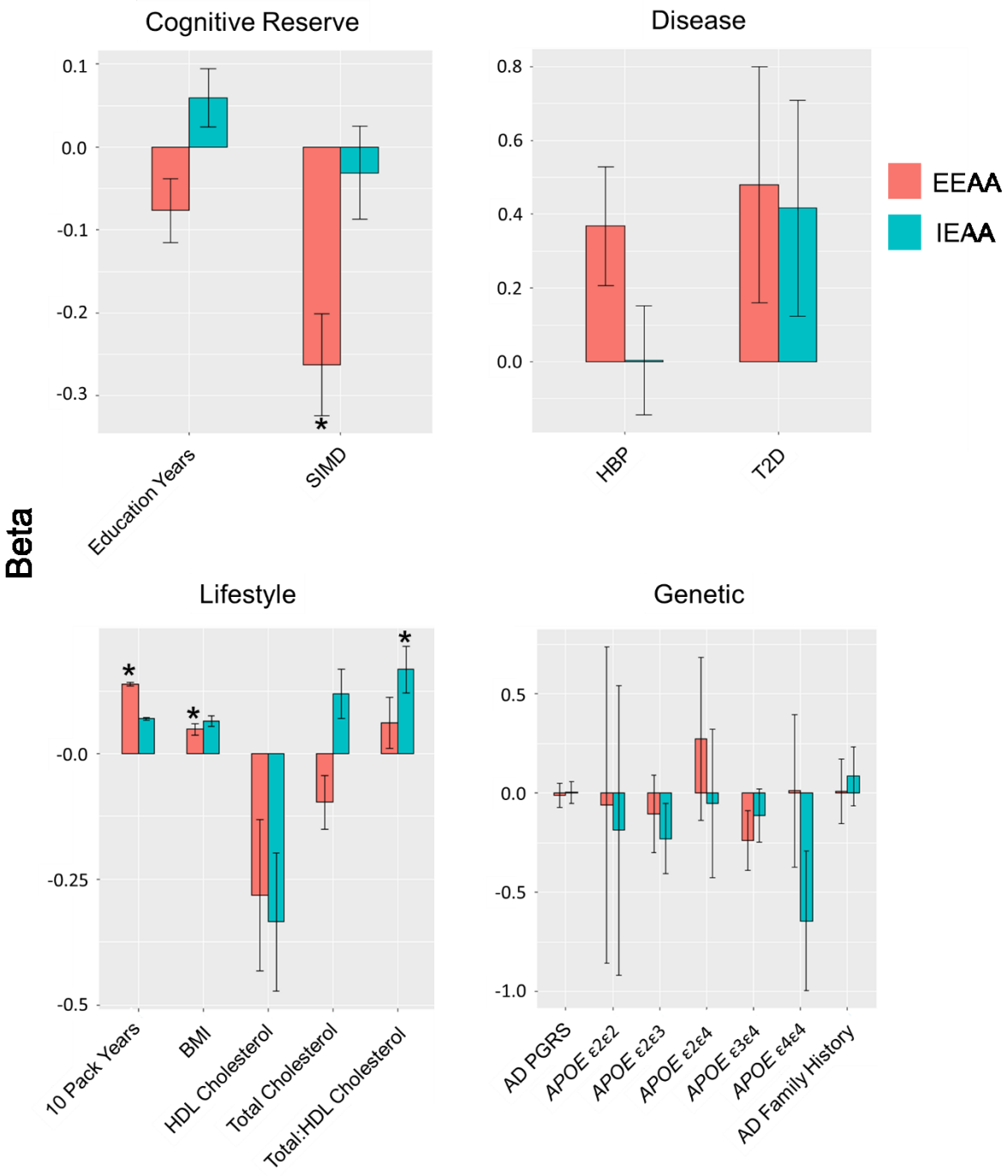

Figure 1: Effects of AD risk factors on age acceleration.

Bar plots are separated into four groups of AD risk factors: cognitive reserve, disease, lifestyle, and genetic. Model Beta coefficients (i.e. effect sizes) are presented along the Y-axes while risk factors are presented along the X-axes. Bars are coloured by extrinsic epigenetic age acceleration (EEAA; red) and intrinsic epigenetic age acceleration (IEAA; blue). Error bars

show the standard error (SE). Bars accompanied by an asterisk (*) represent measures significantly associated with age acceleration at a Bonferroni $P < 0.05$.

SIMD: Scottish Index of Multiple Deprivation, HBP: High Blood Pressure, T2D: Type 2 Diabetes, BMI: Body Mass Index, HDL: High-Density Lipoprotein cholesterol, AD: Alzheimer's Disease, PGRS: Polygenic Risk Score. The effect sizes for smoking have been scaled to represent per 10 pack years of exposure. All other effect sizes are per unit increase (disease positive for HBP and T2D and positive family history of AD) with the exception of SIMD and AD polygenic risk score (per SD), and *APOE* status, where the effects are relative to the ε3ε3 reference category.

McCartney et al. Table 1: Summary of variables assessed in the Generation Scotland cohort

Variable	**N**	**Mean**	**SD**
Chronological age (years)	5,100	48.51	13.99
Horvath's estimated age (years)	5,100	52.60	11.59
Hannum's estimated age (years)	5,100	39.42	11.68
Body mass index (BMI; kg/m^2)	4,977	27.03	5.37
Smoking (pack years)[*]	4,997	9.13	17.28
High-density lipoprotein (HDL) cholesterol (mmol/L)	4,948	1.49	0.42
Total cholesterol (mmol/L)	4,960	5.13	1.09
Total:HDL cholesterol (ratio)	4,948	3.67	1.22
	N	**Median**	**IQR**
Socioeconomic status (SIMD[†]; rank)	4,728	4,230	2,148.5 – 5,423
Education[‡]	4,816	4	3 - 6
AD Polygenic risk score	4,994	1.7×10^{-4}	$1.6 \times 10^{-4} – 1.9 \times 10^{-4}$
	N		
Sex (male/female)	1,918/3,083	-	-
Type 2 diabetes (yes/no)	171/4,830	-	-
High blood pressure (yes/no)	768/4,830	-	-
AD Family history (yes/no)	834/4,167	-	-
APOE (ε2ε2)	27	-	-
APOE (ε2ε3)	572	-	-
APOE (ε2ε4)	108	-	-
APOE (ε3ε3)	2952	-	-

APOE (ε3ε4)	1126	-	-
APOE (ε4ε4)	124	-	-

* The following smoking categories were available: current smoker (N = 939); former smoker, stopped within past 12 months (N = 158); former smoker, stopped more than 12 months ago (N = 1,309); never smoker (N = 2,533). Data were unavailable for 62 participants.

† Scottish Index of Multiple Deprivation

‡ Education was measured as an ordinal variable. 0: 0 years, 1: 1-4 years, 2: 5-9 years, 3: 10-11 years, 4: 12-13 years, 5: 14-15 years, 6: 16-17 years, 7: 18-19 years, 8: 20-21 years, 9: 22-23 years, 10: ≥24 years.

McCartney et al. Table 2: Age acceleration and AD risk factors. Significant associations after accounting for multiple comparisons are highlighted in bold.

	IEAA			EEAA		
Cognitive Reserve	**Beta**	**SE**	**P**	**Beta**	**SE**	**P**
Socioeconomic status (SIMD*; SD)	-0.031	0.056	0.580	-0.262	0.061	**1.90×10^{-5}**
Education† (per unit)	0.059	0.035	0.092	-0.076	0.039	0.048
Disease	**Beta**	**SE**	**P**	**Beta**	**SE**	**P**
Type 2 diabetes (yes/no)	0.417	0.293	0.150	0.480	0.320	0.130
High blood pressure (yes/no)	0.004	0.148	0.980	0.368	0.161	0.022
Lifestyle	**Beta**	**SE**	**P**	**Beta**	**SE**	**P**
Body mass index (BMI; kg/m^2)	0.064	0.010	**1.6×10^{-10}**	0.048	0.011	**1.2×10^{-5}**
Smoking‡ (Pack years)	0.007	0.003	0.025	0.014	0.003	**4.5×10^{-5}**
High-density lipoprotein (HDL) cholesterol (mmol/L)	-0.335	0.138	0.015	-0.282	0.151	0.062
Total cholesterol (mmol/L)	0.120	0.049	0.015	-0.097	0.054	0.072
Total:HDL cholesterol (ratio)	0.168	0.046	**2.70×10^{-4}**	0.062	0.05	0.22
Genetic	**Beta**	**SE**	**P**	**Beta**	**SE**	**P**
AD Polygenic risk score (SD)	0.004	0.056	0.940	-0.010	0.061	0.870
AD Family history (y/n)	0.085	0.149	0.570	0.009	0.163	0.960
APOE (ε2ε2)	-0.187	0.730	0.800	-0.061	0.798	0.940
APOE (ε2ε3)	-0.229	0.178	0.200	-0.106	0.195	0.590
APOE (ε2ε4)	-0.050	0.374	0.890	0.274	0.409	0.500
APOE (ε3ε4)	-0.113	0.136	0.410	-0.239	0.149	0.110
APOE (ε4ε4)	-0.644	0.352	0.067	0.012	0.385	0.970

[*] Scottish Index of Multiple Deprivation

[†] Education was measured as an ordinal variable. 0: 0 years, 1: 1-4 years, 2: 5-9 years, 3: 10-11 years, 4: 12-13 years, 5: 14-15 years, 6: 16-17 years, 7: 18-19 years, 8: 20-21 years, 9: 22-23 years, 10: ≥24 years.

[‡] The following smoking categories were available: current smoker (N = 939); former smoker, stopped within past 12 months (N = 158); former smoker, stopped more than 12 months ago (N = 1,309); never smoker (N = 2,533), and not coded (N = 62).