Peripheral DNA methylation, cognitive decline and brain ageing: Findings from the Whitehall II Imaging Study

Leonidas Chouliaras¹,²,³,⁶, Ehsan Pishva³,⁴,⁶, Rita Haapakoski¹,⁵, Eniko Zsoldos¹, Abda Mahmood¹, Nicola Filippini¹, Joe Burrage³, Jonathan Mill³, Mika Kivimäki⁵,⁶, Katie Lunnon³,², Klaus P. Ebmeier¹,²

¹Department of Psychiatry, University of Oxford, Warneford Hospital, OX3 7JX, United Kingdom
present: Department of Psychiatry
²Current: Department of Psychiatry, University of Cambridge, Cambridgeshire, UK
³University of Exeter Medical School, RILD, University of Exeter, Barrack Road, Devon, UK
⁴Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience (MHeNS), Maastricht University, Maastricht, The Netherlands
⁵Department of Epidemiology and Public Health, University College London, United Kingdom;
⁶Clinicum, Faculty of Medicine, University of Helsinki, Finland.

#These authors contributed equally to this work

*Corresponding author: Leonidas Chouliaras (MD, PhD, MRCPsych), Department of Psychiatry, University of Cambridge School of Clinical Medicine, Box 189, Level E4 Cambridge Biomedical Campus, Cambridge CB2 0SP, UK. Tel.: +44 1223 760682; fax: +44 1223 336968
DISCLOSURE STATEMENT
The authors declare that they have no conflicts of interest in regard to this work.

ACKNOWLEDGMENTS
We thank the Whitehall II cohort participants for their time. The study was funded by the National Institute of Health Research (NIHR) Academic Clinical Fellowship programme (ACF; LC) the Oxfordshire Health Services Research Committee (OHSRC; LC), the Oxford University Clinical Academic Graduate School (OUCAGS; LC), the UK Medical Research Council (G1001354; KPE; K013351; MK), the Gordon Edward Small’s Charitable Trust (SC008962; KPE), the HDH Wills 1965 charitable trust (charity No: 1117747; KPE), the Medical Research Council (MRC) (Grant number: MR/N027973/1; KL, EP) as part of the Joint Programme—Neurodegenerative Disease Research (JPND) grant for the EPI-AD consortium.

ABSTRACT

Background: DNA methylation (DNAm) has been linked with the pathophysiology of brain aging, cognitive impairment and dementia. Methods: The present study investigated the association between blood genome-wide DNAm profiles, cognitive dysfunction and brain magnetic resonance imaging (MRI) measures in 48 participants of the Whitehall II imaging sub-study. Results: We identified eight differentially methylated regions (DMRs) associated with cognitive impairment. Accelerated aging based on the Hannum epigenetic clock was associated with mean diffusivity and global fractional anisotropy. We also identified modules of co-methylated loci associated with white matter hyperintensities. These co-methylation modules were enriched among pathways relevant to beta-amyloid processing and glutamatergic signalling. Conclusion: Our data support the notion that blood DNAm changes may have utility as a biomarker for cognitive dysfunction and brain aging.

Keywords: Mild Cognitive impairment, DNA methylation, Magnetic resonance imaging.
Summary Points

- Differentially methylated regions in 8 genes are associated with cognitive dysfunction
- Epigenetic clock age is associated with fractional anisotropy and mean diffusivity
- White matter hyperintensities co-methylated modules are enriched in pathways involved in beta amyloid processing and glutamatergic signalling
1. **INTRODUCTION**

DNA methylation (DNAm) changes have been linked with the pathophysiology of brain aging, Alzheimer’s disease (AD) and other types of dementia [1,2]. Considering the relative stability of DNAm and the fact that it is directly modulated by both underlying genetic sequence and environmental exposures it appears as a promising peripheral biomarker for brain related changes [3]. Peripheral changes in mRNA and proteins that are downstream of DNAm regulation further support this concept [4,5]. Studies have shown differences in blood DNAm when comparing AD subjects with controls [6], but interestingly when looking at specific genes there is little overlap between brain and blood DNAm changes in AD [7,8]. This could be due to several factors, such as the effect of cell differentiation, developmental programming and tissue specificity of methylation profiles [9]. Moreover, the rapid turnover of peripheral cells may not mirror the cumulative environmental exposures and hence possible epi-mutations that a post mitotic neuron can accumulate over time. Alternatively, technical limitations, such as the limited coverage of total CpGs by the 450K array may fail to detect cross tissue overlap that could be driven by genetic or early life factors [10]. However, a good peripheral biomarker does not have to mirror disease-associated changes in the brain; it could also simply represent a peripheral response to central pathology. At present, there is limited data available on the potential utility of peripheral DNAm as a marker for cognitive decline before the onset of dementia. One study of patients with type 2 diabetes mellitus who later developed presymptomatic dementia highlighted leukocyte DNAm changes that were comparable to changes identified in blood in AD patients and could thus potentially represent early markers of dementia [11]. Recently DNAm epigenetic clocks have been developed, which provide a measure of epigenetic age, based on DNAm levels at 353 CpG sites [12] and 71 CpG sites [13], respectively. A number of studies have utilized the epigenetic age in blood, based on the DNAm epigenetic clock, to show correlations with cognitive function, white matter hyperintensities (WMH), Parkinson’s disease, Down’s syndrome and all-cause mortality [14–18].

Considering the limited access to brain tissue in living humans, multimodal magnetic resonance imaging (MRI) has been widely used to study brain aging *in vivo*. MRI measures of reduced brain
volume, reduced white matter integrity and increased white matter hyperintensities have been linked with the risk for developing dementia [19,20]. As such, brain imaging changes are useful surrogate markers of brain health and aging. Here, we investigated the potential utility of DNAm profiling as a peripheral marker for cognitive decline and brain health in the Whitehall II Imaging Sub-Study [21], which provides a valuable source of longitudinal data to explore factors hypothesized to affect brain health and cognitive aging.

2. MATERIALS AND METHODS

2.1. Participants

Forty-eight participants were selected from subjects recruited to the Whitehall II Imaging Sub-Study between May 2012 and December 2014 [21]. All participants were members of the Whitehall II study, a prospective occupation cohort study established in 1985 [22]. For the Imaging Sub-Study based at the University of Oxford, a sample of participants in the 2012-2013 clinical examination were selected at random from the whole Whitehall cohort. Ethical approval was obtained from the University of Oxford Central University Research Ethics Committee, and the University College London (UCL) Medical School Committee on the Ethics of Human Research. Informed written consent was obtained from all participants [21].

In 2012-2016, participants of the Whitehall II Imaging Sub-Study invited to Oxford had a clinical interview and cognitive assessment, followed by a multimodal 3T MRI scan (full details are provided in Supplementary Methods) and subsequently blood samples were taken. The detailed study protocol has been previously published elsewhere [21]. We assessed cognitive impairment using the Montreal Cognitive Assessment (MoCA) [23], and studied 24 cognitively impaired participants (MoCA<26), who would qualify for a clinical diagnosis of mild cognitive impairment (MCI), alongside 24 age and gender matched cognitively normal participants (MoCA>26), who were free form other comorbidities such as depression (Supplementary Table 1). A range of MRI measures were obtained, including total brain volume (TBV), left hippocampal volume (LHV), right hippocampal volume (RHV), cerebrospinal fluid volume (CSF), white matter volume (WM), grey matter volume (GM), fractional
anisotropy (FA), mean diffusivity (MD) and white matter hyperintensities (WMH), with full details provided in the Supplementary Methods.

### 2.2 DNA methylation profiling

Samples (peripheral blood mononuclear cells (PBMCs)) were drawn using the Vacutainer CPT tubes (Becton Dickinson) and the tube was immediately inverted eight times to mix anticoagulant and processed within two hours of collection. CPT tubes were centrifuged in 1600g for 30 minutes to separate serum from peripheral white blood cells and red blood cells. PBMCs were washed with phosphate buffered saline (PBS), counted, cryopreserved using Cell Freezing Medium (5% DMSO/11% HAS in PBS) and stored in -80°C until DNA isolation took place. DNA was isolated from all samples simultaneously using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer’s protocol and tested for degradation and purity before analysis. For each sample, 500ng of genomic DNA was bisulfite treated using the EZ-96 DNA Methylation-Gold™ Kit (Zymo Research) according the manufacturer’s protocol. A total of 48 samples were processed on Illumina Infinium HumanMethylation450K BeadChip (HM450K) (Illumina) according to the protocol supplied by the manufacturer. Intensity values generated by the Illumina iScanner were imported into R statistical software v3.3.2 using `readEPIC` function in the Bioconductor package wateRmelon v1.18 [24]. Data quality control (QC) and normalization were conducted using the methylumi [25], wateRmelon [24] and minfi [26] R packages according to the pipeline previously described [7]. After the QC steps, one sample was dropped from the analysis due to possessing an XXY genotype.

### 2.3 Statistical analyses

A linear regression model was conducted to identify differentially methylated positions (DMPs) associated with cognition, whilst controlling for co-variates. We used the MoCA [23] to determine cognitive impairment (full details provided in Supplementary Methods). Cellular composition of peripheral blood cells was measured using an online calculator (https://dnamage.genetics.ucla.edu), and included in the model as covariates together with gender, age, processing batch (chip and position on array). We have previously established the multiple testing threshold (experiment-wide significance) for epigenome wide association data (EWAS) data generated on the HM450K array as
To identify differentially methylated regions (DMRs) we used the \textit{comb-p} software tool [28], to identify ≥3 adjacent significant DMPs in a 1000bp sliding window. The \textit{comb-p} algorithm corrects the DMP P values for auto-correlation between probes and then scans the genome for peaks of association around a seed signal (set to P< 5xE05). For each region it calculates the Stouffer–Liptak corrected P value, which is then adjusted for multiple testing using Šidák’s correction [27,28]. Next, we separately tested the association between blood cell composition and the methylation status of the top 20 MoCA associated CpGs. We also performed targeted gene analysis, testing the associations between MoCA and neuroimaging variables (LHV, RHV, TBV, CSF, WM, GM, FA, MD, WMH) and the methylation of 45 CpGs that are included in the 450K array that are related to genes previously associated with AD in blood (20 CpGs in ABCA7, 7 CpGs in Trem2, 18 CpGs in SNCA; [29–31]). For the targeted gene analysis we used Bonferroni correction for multiple testing (\(\alpha=1.1\text{-E03}\)) and used the same covariates in those models as described above.

Our secondary analysis focussed on identifying the relationship between accelerated aging in blood (based on the Horvath and Hannum epigenetic clock age measures) and neuroimaging variables. The Horvath predicted age based on 353 CpG probes and the Hannum predicted age based on 71 CpG probes were generated using online calculators [12,13] (https://dnamage.genetics.ucla.edu). The variables “AAHOAdjCellCounts” and “AAHAAdjCellCounts” derived from the online calculator were used in linear regression models as independent variables and MoCA and MRI measures (LHV, RHV, GM, WM, CSF, FA, MD, WMH) were used as dependent variables. Cell-type composition, gender, age, smoking, alcohol consumption, premorbid IQ (Test of premorbid function; TOPF) were used as covariates in the model. Bonferroni correction was used for the DNAm clock analyses to adjust for multiple testing with level of significance set at \(\alpha=0.0055\).

Our final analysis used weighted gene co-methylation network analysis (WGCNA) [32] to group together highly correlated probes. First, we calculated the variance of methylation values across all 47 samples for each probe. Next, we excluded probes showing variance lower than the median of calculated variance in the entire dataset (variance <5.86E-04). A signed network was constructed using the automatic one-step blockwiseModules function within the WGCNA package, based on a
block size of 5000 and using a soft threshold parameter of six. Each module was then labelled with a
unique colour identifier according to size. Module-trait associations were examined using Pearson’s
correlation coefficients to examine associations between the module eigengene (ME) values and
cognitive and MRI outcomes (full details provided in Supplementary Methods) including MoCA total
scores, and various MRI measures as described in section 2.1. The association between modules and
the covariates that had been used for our initial linear regression analyses, including gender, age,
batch effects and cell type compositions were also examined. Subsequently, the module membership
(MM) of each probe was calculated as the correlation between DNAm values and ME values. To
incorporate cognitive performance (MoCA scores) in the WGCNA modules we calculated the gene
significance (GS) of each probe. Thus GS describes the strength of the correlation between the
DNAm of each probe and the MoCA scores while MM quantifies the extent to which a gene conforms
to the characteristics of a module (i.e. correlates with ME) [33]. Subsequently, the correlation between
MM and GS for each probe was computed. The correlation between MM and GS for each probe
indicates how central a given probe is to the determined module and how associated it is with the
MoCA scores. Gene ontology (GO) and KEGG pathway enrichment analysis was conducted for the
CpG sites that were grouped in the modules showing significant associations with the outcomes. The
missMethyl package in R (v 1.9) [34] was used for this analysis, which adjusts for the number of
CpGs associated with each gene in the HM450K array.

3. RESULTS AND DISCUSSION

3.1. Blood DNA methylation signatures are associated with cognitive impairment

We identified a number of MoCA-associated DMPs (Table 1A), but none reached the multiple testing
threshold for epigenome wide association data (EWAS) (P<2.2E-07) [27]. Our most significant CpG
site residing in BNC1 (cg26429925; standardized regression coefficient= -0.75, P=4.71E-06). Differential methylation of BNC1, which encodes a zinc finger protein basonuclein, has been
previously found in frontal cortex of AD subjects compared to controls [35] and upregulation of
BNC1 has been observed in a novel chimeric model of AD [36]. Our subsequent regional analysis
identified eight DMRs (each containing ≥3 DMPs), which were significantly associated with MoCA scores, with a Sidak-corrected P<0.05 (Table 1B; Figure 1A). The eight identified DMRs were annotated to the HLA-DPA1 / HLA-DPB1, DRC1, PRKAA2, CALCB, CDH2, RTBDN, ZNF256 and SHANK2 genes. The top DMR we identified resided in genes encoding subunits of the Human Leukocyte Antigen (HLA) DP receptor, which is involved in the immune response. Alterations in the immune profile of leukocytes have been previously described in MCI patients [33]. Moreover increased RNA expression of HLA DPA1 and DPB1 in frontal cortex and hippocampus have been associated with the transition from MCI to AD [37]. It is unclear whether in our study the observed methylation changes in this DMR could affect gene expression. However, considering that they span the whole gene, including the TSS, it is likely that these methylation changes alter transcription. Interestingly, methylation changes in a probe of the SHANK2 gene, which is involved in the structural and functional organization of the dendritic spine and synaptic junction, has been previously reported in a study comparing blood DNAm profiles of AD subjects to non-demented controls [7]. We did not find any associations between cell composition and MoCA scores. With regards to targeted gene analysis the most remarkable was a negative association between methylation at cg25748868 in TREM2 and WMH (standardized regression coefficient=-11.3, P=3.9E-03) however this did not survive Bonferroni correction for multiple testing (a=1.1E-03) and neither did any of the other tested comparisons (full data not shown). WMH load has been consistently associated with AD, TREM2 hypomethylation and overexpression have been previously observed when comparing blood samples of AD and controls, while increased TREM2 methylation and hydroxymethylation have been observed in AD brain [29,38–41]. Overall it is likely that previously identified associations between cell composition and Parkinson’s disease as well as methylation and expression changes in ABCA7, TREM2 and SNCA with Alzheimer’s disease and SNCA methylation with Dementia with Lewy Bodies are specific for those conditions [29–31,42,43].

3.2. Accelerated blood DNAm age is associated with MRI alterations

Our linear regression analyses showed a significant association between chronological age and various imaging measures, as to be expected (Supplementary Table 2). We also observed a significant
association between accelerated DNAm age (based on the Hannum epigenetic age clock) and two MRI measures (Supplementary Table 2). Increased DNAm age (based on the Hannum epigenetic clock) was associated with global measures of increased FA (P= 8.99E-04) and reduced MD (P= 8.33E-04), with these associations being in the opposite direction than the associations between global FA and MD with chronological age. Such findings seem paradoxical as reduced FA and increased MD have been associated with aging and symptomatic neurodegeneration and could be chance findings. Alternatively, epigenetic aging associated with increased FA and decreased MD could be linked with comparable findings in presymptomatic carriers of familial AD mutations, where it has been suggested that such findings could reflect earlier axonal insults in the pathway towards subsequent neurodegeneration such as neuroinflammatory changes leading to neuronal and glial swelling [44,45]. Nevertheless most studies indicated no change or decreased FA and increased MD in preclinical familial AD [44]. No associations were observed between the Horvath epigenetic clock and any of the imaging variables tested and neither of the epigenetic clocks were associated with cognitive function or presence of cognitive impairment. In comparison to the study by Raina and colleagues, we found no associations between accelerated DNAm age and WMH [17].

3.3. WGCNA

WGCNA was used to construct clusters of highly correlated probes. For the current HM450K dataset, we identified 41 modules, with each assigned a unique colour (Supplementary Table 3). Of those 41 modules, 16 were significantly associated with at least one of the covariates such as gender, age, cell composition estimates (based on Horvath epigenetic calculator; [12]) or premorbid IQ. The grey module is considered to be 'background' CpG sites that are not clustered into any module, and were thus disregarded from further analyses. From the remaining 24 modules (Figure 1B), the “yellowgreen” module, which contained 278 probes, showed the greatest association with WMH, after correcting for multiple tests (R=0.74, P=2.28E-09). A positive correlation between module membership (MM) and gene significance (GS) for each of the 278 probes in the “yellowgreen” module was found with respect to MoCA scores (R=0.27, P=5.35E-06; Figure 1C). The GO and KEGG enrichment analysis for the probes in the “yellowgreen” module highlighted many pathways
associated with the pathophysiology of cognitive dysfunction and dementia (Supplementary Tables 4 and 5), for example the GO terms “main axon” (GO:0044304: P=4.37E-03) and “beta-amyloid binding” (GO:0001540: P =0.0109), and the KEGG pathways “Glutamatergic synapse” (hsa04727: P=4.79E-04) and “Alzheimer’s disease” (hsa05010: P=0.0491). DNA methylation and hydroxymethylation changes in genes involved in these pathways, such as axon and glutamatergic synapse have been previously linked to brain ageing and cognition in experimental models [46,47].

4. CONCLUSIONS

The current study has identified DNAm changes in PBMCs that correlate with age-related cognitive dysfunction, measures of white matter integrity and white matter hyperintensities. We have identified eight DMRs that were associated with age-related dysfunction, with the most significant being related to immune response. Another identified DMR, SHANK2, has also been found to be differentially methylated when comparing blood DNAm profiles of AD cases and controls and could represent an early dementia biomarker. We have also identified links between accelerated aging based on the epigenetic clock and MRI measures of white matter integrity and have found associations between white matter hyperintensities and methylation of genes involved in neurodegenerative processes.

Advantages of the study are related to the unique combination of multimodal MRI imaging, cognitive assessments and genome-wide DNAm analysis in a well characterised cohort with available extensive sociodemographic data giving the ability to control for confounders such as premorbid functioning. One of the main limitations of this pilot study is the relatively small sample size for EWAS. Therefore, any results will need to be replicated in larger independent cohorts. Future research will also need to assess the timing and specificity of such changes and expand research in more robust sequencing methods as the current arrays only cover a small proportion of DNAm sites. Moreover, the current approach does not distinguish between DNAm and hydroxymethylation and other DNA modifications. Additional limitations are related to use of blood for methylomic profiling, the limited overlap with brain profiles and hence limited utility in providing novel mechanistic insights with regards to improving our understanding of the pathophysiology of cognitive dysfunction. This is
further supported by the fact that we have detected correlations between brain white matter hyperintensitives and methylation of genes involved in some but not all aspects of neurodegenerative processes [48–50].

Overall, these findings suggest that DNA methylation changes in blood could have potential utility as a peripheral biomarker for age-related cognitive dysfunction and associated structural brain changes. Considering the need for robust biomarkers for early detection of dementia to test effective treatments in a timely manner, more research in the field of epigenetics and mild cognitive impairment is warranted.

REFERENCES


