

# **Polygenic coronary artery disease association with brain atrophy in the cognitively impaired.**

Eric de Silva, PhD<sup>1,2</sup>, Carole H Sudre, PhD<sup>1,3,4</sup>, Josephine Barnes, PhD<sup>5</sup>, Marzia A Scelsi, PhD<sup>1</sup> and André Altmann, PhD<sup>1</sup> for the Alzheimer's Disease Neuroimaging Initiative\*

<sup>1</sup>*Centre for Medical Image Computing, University College London, London, UK*

<sup>2</sup>*NIHR University College London Hospitals Biomedical Research Centre, London, UK*

<sup>3</sup>*MRC Unit for lifelong health and ageing, University College London, London, United Kingdom*

<sup>4</sup>*School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK*

<sup>5</sup>*Dementia Research Centre, UCL Queen Square Institute of Neurology, London, United Kingdom*

\*Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at [http://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

**Corresponding author: [eric.desilva@ucl.ac.uk](mailto:eric.desilva@ucl.ac.uk)**

## **ABSTRACT**

The role cardiac function plays in the predilection for, and progression of, Alzheimer's disease (AD), is complex and unclear. While a number of low-frequency genetic variants of large effect-size have been shown to underlie both cardiovascular disease and dementia, recent studies have highlighted the importance of common genetic variants of small-effect size, which, in aggregate, are embodied by a polygenic risk score (PRS). In this study we aim to investigate the effect of polygenic risk for coronary artery disease (CAD) on brain atrophy in AD using whole brain volume (WBV) and put our findings in context with the polygenic risk for AD and presumed small vessel disease as quantified by white matter hyperintensities (WMH). We used 730 subjects from the ADNI database to investigate PRS effects (beyond

APOE) on whole brain volumes, total and regional WMH and amyloid beta across diagnostic groups. In a subset of these subjects (N=602) we utilise longitudinal changes in whole brain volume over a maximum of 24 months using the boundary shift integral approach. Linear regression and linear mixed effects models were used to investigate the effect of WMH at baseline as well as AD-PRS and CAD-PRS on whole brain atrophy and whole brain atrophy acceleration, respectively. All genetic associations were examined under oligogenic ( $p=1e-5$ ) and the more variant-inclusive polygenic ( $p=0.5$ ) scenarios. Our results suggest no evidence for a link between PRS score and markers of AD pathology at baseline (when stratified by diagnostic group). However, both AD-PRS and CAD-PRS were associated with longitudinal decline in WBV (AD PRS  $t=3.3$ ,  $P_{FDR}=0.007$  over 24 months in healthy controls) and surprisingly, under certain conditions WBV atrophy is statistically more correlated with cardiac PRS than AD PRS (CAD PRS  $t=2.1$ ,  $P_{FDR}=0.04$  over 24 months in the MCI group). Further, in our regional analysis of WMH, AD PRS beyond APOE is predictive of white matter volume in the occipital lobe in AD subjects in the polygenic regime. Finally, the *rate of change* of brain volume (or *atrophy acceleration*) may be sensitive to AD polygenic risk beyond APOE in healthy individuals ( $t=2$ ,  $p=0.04$ ). For subjects with mild cognitive impairment (MCI), beyond APOE, a more inclusive polygenic risk score including more variants, shows CAD PRS to be more predictive of WBV atrophy, than an oligogenic approach including fewer larger effect size variants.

**Keywords: Alzheimer’s disease, coronary artery disease, GWAS, brain atrophy, polygenic risk score, white matter hyperintensities**

## 1. Introduction

It is estimated that more than one million people in the UK, and over 44 million individuals globally<sup>1</sup> are living with dementia. Alzheimer's disease (AD) is the most common form of dementia and is usually diagnosed in the elderly (over the age of 65 years). There are many symptoms associated with AD and these include changes in memory, language and personality<sup>2</sup>.

Beyond cognitive testing there has been a large focus on the use of brain imaging to help diagnose and track the disease. In particular, a decrease in brain volume and the build-up of protein in the form of amyloid plaques (between neurons) and misfolded neurofibrillary tangles of hyperphosphorylated tau (within neurons) are seen in brain imaging studies and observed in post-mortem examinations of AD patients<sup>3</sup>. It is thought that the amyloid and tau aggregations contribute to the death of neuronal cells resulting in a reduction in regional grey-matter volume and neuronal connectivity<sup>4</sup>. Surprisingly, years of clinical trials involving pharmacological interventions targeting these protein deposits have largely been unsuccessful<sup>5</sup>, possibly because they are formed as a result of earlier pathological changes. Another explanation may be that drug interventions, thus far, have all been administered too late in the disease life-course to be effective, and that intervention earlier in life may be required. Current symptomatic treatments of AD symptoms come from acetylcholinesterase inhibitors and glutamate blockers which serve, in some cases, to reduce disease severity, but are by no means curative<sup>6</sup>. At the time of writing the monoclonal antibody aducanumab, which targets beta-amyloid plaques, had just been approved for clinical use in the United States (see [www.fda.gov](http://www.fda.gov)).

AD is often diagnosed alongside vascular dementia in what is called mixed dementia<sup>7</sup>. Vascular dementia (itself the second most common form of dementia following AD) is the result of reduced blood flow to brain cells resulting in cell-death. This can be caused by cerebral small vessel disease, resulting in subcortical vascular dementia, which affects vessels deep in the brain. One imaging marker of small vessel disease is MRI-visible white matter changes, or, white matter hyperintensities (WMHs). WMHs are known to increase with age<sup>8,9</sup>, but are also strongly associated with AD<sup>10</sup>. WMHs have also been associated with a range of other pathologies, including amyloid angiopathy, arteriosclerosis, axonal loss, blood-brain barrier leakage, demyelination, gliosis, hypoperfusion, hypoxia and inflammation<sup>11</sup>.

In coronary artery disease (CAD) plaques aggregate in blood vessels that feed the heart oxygen and nutrients. In the extreme, this can lead to angina and heart attack, but a smaller prolonged reduction in cardiac function may be responsible for cerebral hypoperfusion. CAD has a strong genetic basis, being ~50% heritable with ~60 genetic loci identified<sup>12</sup>. However, the relationship between the genetic contribution to heart health and AD remains largely unexplored.

While age is the most salient factor in AD risk, there is, alongside environmental and lifestyle factors, a genetic component underlying both AD and CAD. In AD, a small number of cases (<5%) are due to autosomal dominant early-onset AD for which there are a number of rare, large effect-size genetic variants that contribute to the pathology<sup>13</sup>. These include mutations that result in abnormal protein products of amyloid precursor protein (APP), or in the genes that code for the enzymes that alter the breakdown of APP, both of which may result in an increase in amyloid plaques.

Most AD cases are sporadic and late-onset (typically found in those aged 65 and older) where heritability is estimated to be between 60% and 80%<sup>14</sup>. Here, a number of identified common variants, most notably the e4 allele of the *APOE* gene which accounts for ~5% of AD heritability, plus about 20 additional loci, account for ~30% AD heritability<sup>15</sup>. It is likely that the remaining heritability is the result of the combination of a great many (1000s to 100,000s) of common variants, each contributing a very small effect.

As a result of these findings, much research has been performed to establish how to best capture a composite measure of these many common variants, that individually have such a small effect. Polygenic Risk Scores (PRSs) offer a way of doing this and have become increasingly used following the many large Genome Wide Association Studies (GWAS) which show associations between common genetic variants and diseases. The PRS sums up the effect size<sup>16</sup> across a selected set of genetic variants shared between a discovery sample (some CAD GWAS for example) and target sample (some other genotyped cohort such as the Alzheimer's Disease Neuroimaging Initiative, ADNI), resulting in an aggregate score that reflects the genetic contribution to the disease phenotype in the target cohort. AD-PRS have effectively discriminated between AD cases and controls<sup>17</sup>, been used as a predictor of conversion from MCI to AD<sup>18</sup>, been linked to inflammatory biomarkers<sup>19</sup>, CSF amyloid beta

levels<sup>20</sup>, CSF tau levels<sup>21</sup>, hippocampal volume<sup>22</sup>, cortical thickness<sup>23</sup> and age of onset of AD<sup>24</sup>. AD-PRS related work has been recently reviewed by<sup>25</sup>

Just as research has been devoted to investigating how AD-PRS affects AD phenotypes, the same concept can be used to investigate genetic risk for other diseases on AD pathology. Here, we specifically investigate the role of common, small-effect, cardiac-related genetic variants to Alzheimer's disease. This enables us to elucidate the role that cardiovascular health plays in relation to dementia. While we focus here on the influence of underlying genetics upon WMHs, whole brain atrophy and the changes therein, the cardiac-cerebrovascular axis is no doubt complex, encompassing many biological pathways. We believe, however, that a combination of genetics and imaging, in particular longitudinal images that capture changes over a disease trajectory, will provide important insights into this system.

## **2. Material and methods**

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date information, see [www.adni-info.org](http://www.adni-info.org). The R-data-package ADNIMERGE (dated 2020-05-19) was used to access ADNI data.

### **2.1 Computing white matter hyperintensities**

Regional and total WMH values were determined using the Bayesian Model Selection (BaMoS) software<sup>26</sup>, a white matter lesion segmentation algorithm. BaMoS was applied to 932 ADNIGO and ADNI2 participants following<sup>27</sup> and described therein. In short, the label fusion algorithm GIF (Geodesic Information Flows)<sup>28</sup> was used to parcellate the T1-weighted cortical grey matter into various cortical and subcortical brain structures in an automated fashion. It additionally carries out skull stripping and generates probabilistic atlases for each

individual. These atlases are then input to BaMoS alongside co-registered FLAIR images consisting of log-transformed normalised intensities. To determine white matter lesions, BaMoS computes the most suitable model description of the data accounting for prevailing outliers. It manages this by first partitioning data into inlier and outlier portions and then modelling input data in a hierarchical fashion, with elements then separated into one of four tissue types – grey matter, white matter, cerebrospinal fluid and non-brain. Each of these in turn are modelled via a Gaussian mixture model with the number of constituent parts determined using a split-and-merge strategy. An Expectation-Maximization (EM) algorithm is used for optimisation with model selection implemented using Bayesian Information Criterion to produce probabilistic lesion maps from which measurements of lesion volumes are inferred. Regional volumes in cubic mm were also computed across five lobes and four radial layers.

## 2.2 Brain atrophy

To compute changes in brain volume we utilise the Boundary Shift Integral (BSI) method<sup>29,30</sup>. Briefly, the BSI is defined as the difference in brain volume (either in total brain or in a brain region via displacement of the boundaries), automatically computed (brain mask creation is semi-automated) between a baseline scan and a repeat scan at a later time. The KN-BSI<sup>30</sup> is an extension of the classic-BSI and carries out tissue-specific intensity normalisation which deals with tissue-contrast differences, producing a smaller standard deviation in atrophy changes than the classic-BSI. BSI values were obtained from the *foxlabbsi* table in ADNI comprising 2348 subjects (across ADNI1, ADNI2, ADNI3 and ADNIGO), where all T1-weighted scans included in the core datasets pertaining to BSI were obtained using 3T scanners. Some months had very few available scans so these were removed to maintain consistency between scan interval times. MRI scans were originally made using accelerated and non-accelerated acquisitions and we have chosen to focus on the accelerated scans<sup>31</sup> as they have been shown to result in fewer motion artifacts in pairs of scans resulting from patient movement. Finally, some scans had BSI-determined brain volume increases which may be a result of better subject hydration or due to overall noise. Scans with  $BSI < 0$  (corresponding to an increase in whole brain volume over time) were removed from the analysis.

### **2.3 ADNI genetic target data pre-processing**

The genetic data used in this work is a combination ADNI1, ADNIGO and ADNI2 participant genotypes, comprising a total of 1674 subjects<sup>32</sup>. Details on quality control and imputation (use of the HRC reference panel, the Sanger server, EAGLE2 for phasing, and PBWT for imputation) are described in<sup>33</sup>. Briefly, following imputation, multi-allelic SNPs and SNPs with INFO score less than 0.3 were removed; then calls with less than 90% posterior probability of the imputed genotype were set to missing. SNPs missing in more than 10% subjects, deviating from Hardy-Weinberg equilibrium ( $p < 5e-7$ ) and with minor allele frequency less than 5% were all removed. These processing steps were carried out using PLINK v1.9<sup>34</sup>[[www.cog-genomics.org/plink/1.9/](http://www.cog-genomics.org/plink/1.9/)] and resulted in a final set of 5,082,879 autosomal SNPs. For ancestry determination and relatedness analysis, following<sup>21,33</sup> a HapMap 3 reference panel was utilised where individuals with greater than 80% Central European ancestry were held. PLINK v1.9 was then used to retain common SNPs with MAF  $\geq 5\%$  and carry out LD-pruning and construct a genetic relatedness matrix (threshold=0.1) and filter to remove related subjects. This resulted in the exclusion of 116 subjects most likely due to them being genetically non-central European (there are known to be a number of erroneous self-identifications of European descent in this cohort) or being related to other subjects.

For the CSF amyloid beta measurements we used *ADNIMERGE ABETA.bl* values and removed subjects with missing data and set values recorded as <1700 to 1700 (PET CSF Ab1-42; 192 pg/mL cut-off value; Luminex assay; data range 203 to 1700 with mean=1024.7).

### **2.4 Coronary Artery Disease & Alzheimer's Disease discovery GWASs**

To investigate the PRS contribution due to AD we utilised the largest currently available meta-GWAS of AD featuring 35,274 clinical and autopsy-documented AD cases and 59,163 controls<sup>35</sup>. Summary statistics were downloaded from The National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS; July 2020) comprising 11,480,633 SNPs.

In order to investigate the PRS contribution due to CAD we use a meta-analysis of 60,801 CAD cases and 123,504 controls<sup>36</sup>. Summary statistics were downloaded from the

CARDIoGRAMplusC4D (Coronary ARtery DIsease Genome-wide Replication and Meta-analysis (CARDIoGRAM) plus The Coronary Artery Disease (C4D) Genetics) consortium <http://www.cardiogramplusc4d.org/data-downloads/> in July 2020 comprising 8,624,384 variants.

Given the very large effect size of APOE-e4 upon AD pathology, we removed the APOE region so as to explore genetic effects beyond this risk factor. The block removed from chromosome 19 (hg19 coordinates) comprises SNPs 44 400 375 (rs430308) to 46 500 052 (rs62113435).

## 2.5 Polygenic Risk Scores

The Polygenic Risk Score (PRS) is a weighted sum of allele counts where the weights are odds ratios (effect sizes) from a discovery GWAS and represent the strength between the variant and the trait it is associated with. Polygenic Risk Scores were computed using PRSice v2.1.9<sup>37</sup>. To ensure that loci are isolated and independent, LD clumping was applied so that SNPs in LD with one another are removed such that the SNP with the lowest p-value within each LD block of correlated SNPs is held for analysis. LD clumping was conducted with clumping window of 250kb on either side of the index SNP, with an  $r^2$  threshold of 0.1 and p-value threshold of 1.

The PRS for each subject was computed using the “--score avg” setting in PRSice by summing up the product of each variant by the number of effective alleles observed divided by the number of alleles included for that individual. This last divisor makes the PRS scores more comparable between subjects in the presence of missing SNPs. PRS was calculated for each subject for two p-value cutoffs  $p=1e-5$  (i.e., genome-wide suggestive loci) capturing oligogenic effects and  $p=0.5$  capturing polygenic effects, which have been identified as sufficient to encompass threshold-variation<sup>17</sup>.

Of note, it is imperative when working with PRS that there is no sample overlap between discovery GWAS (here Kunkle et al. 2019) and the target cohort (here the combined ADNI1, ADNIGO and ADNI2). We did not encounter any sample overlap as the discovery GWAS use The Alzheimer's Disease Genetics Consortium (ADGC) summary statistics, which only use HC and AD cases from ADNI1 comprising of 1.5T scans. However, from the 846



individuals with imaging and genetic data, we excluded 116 subjects with non-central European ancestry, resulting in the final sample size of 730 for our study. Moreover, demographic genetic differences arising from ancestral population structure have been found to bias PRS scores<sup>38</sup>. This is the case particularly for high p-value cut-offs (e.g.,  $p=0.5$ ) leading to PRS with many 1000s of SNPs. To account for this effect our linear regression models include the first five principal components of population structure as covariates.

## **2.6 Second-order grey matter changes**

By plotting longitudinal whole brain volume change, or BSI (from baseline in units of cubic cm or ml where  $1\text{cm}^3=1\text{ml}$ ), for each subject taken at three, six, twelve and twenty-four months post baseline, the gradient of the line of best fit gives a second order change (the second derivative) in whole brain volume, or the rate of change of BSI per subject. We term this rate of change the “atrophy rate”, which provides a measure of how fast the whole brain volume is declining. By way of illustration, if, for example, there is a BSI of 6ml between months 0-24 then the gradient would be  $0.25\text{ml/month}^2$  (which equals  $3\text{ml/yr}^2$ ) and would have an intercept of zero.

## **2.7 Statistical analysis**

We first investigated polygenic risk scores for AD (AD PRS) as well as polygenic risk scores for CAD (CAD PRS) and their relationship with total whole brain volume WBV over a range of timepoints (which are normalised by dividing by Intracranial Volume, ICV), total White Matter Hyperintensities WMH (tot) at baseline (which are log-transformed) and CSF amyloid beta measurements. We used linear regressions within each diagnostic group adjusting for age, sex, education, APOE-e4 burden and the first five principal components of population structure. In the case of the WMH (tot) we also included the ICV as an additional covariate. Regressions provided t-values as well as p-values for the association between PRS and biomarkers; given the large number of comparisons ( $36 = 2 \text{ PRS thresholds} \times 2 \text{ PRS} \times 3 \text{ diagnoses} \times 3 \text{ biomarkers}$ ) we have adjusted p-values for multiple comparisons using an FDR-correction (rate=5%). Given that both increasing whole brain atrophy, WMHs as well as BSI values are all known to lead to worse health outcomes all p-values are based on one-tailed tests. For the regional WMH data we carried out an identical analysis but over five lobar regions: frontal, parietal, occipital, temporal and a combined basal ganglia, thalami and

infratentorial volume. Each of these regions was normalised by their corresponding regional volume. Finally, the occipital region was explored in more detail through four radial layers of increasing distance from the ventricular system, such that the first two layers represent the inner periventricular WMH loads and the final two layers the deep WMH loads.

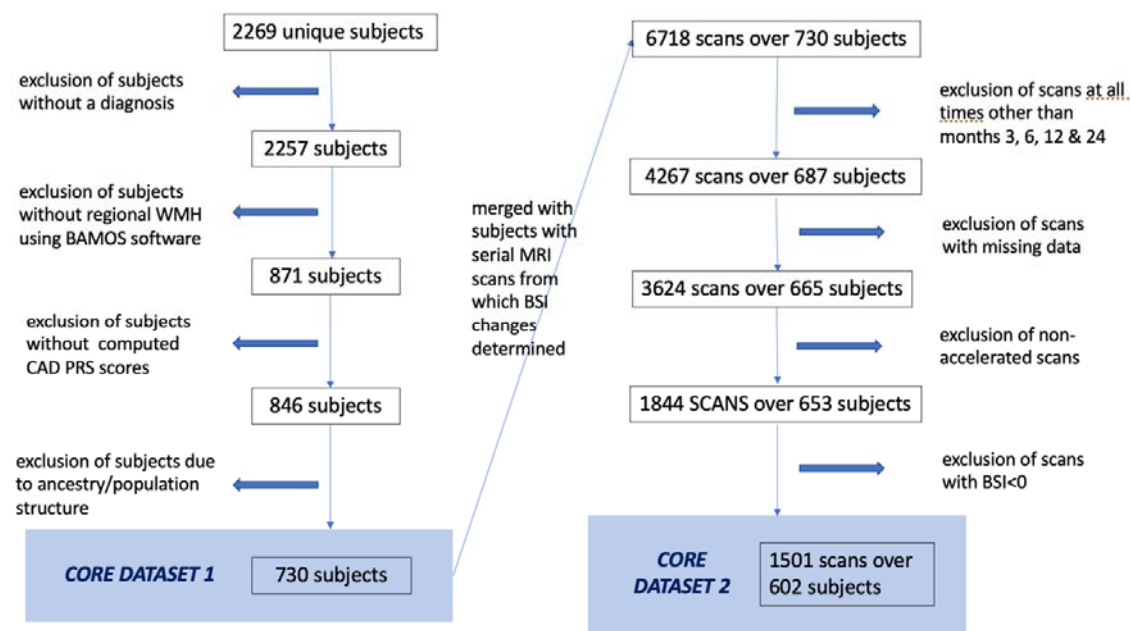
In addition, for the *atrophy accelerations* we also used linear mixed effects models (with subject as random intercept and time-since-baseline as random slope across diagnostic groups; R library lme4) across both p-value PRS-thresholds. In addition, to explore further how *atrophy accelerations* are associated with polygenic risk we also utilised linear mixed effects models. Here subject is the random effect so the atrophy acceleration is calculated per subject. In this way we allow for a positive correlation among measurements for the same individual once the fixed effects have been accounted for. Time or months since baseline is also included as a random effect. For both of these variables there is one intercept. In order to get the influence of PRS on the 'rate of atrophy' (atrophy acceleration) the model includes the PRS-by-time interaction (*PRS\*time*). This allows different PRSs to influence the rate of BSIs. This is all implemented using the R library function lme4 with dependent variable the BSI measurement and the independent variable *PRS\*time* or polygenic score-by-time interaction, where time is the months since baseline) across both p-value PRS-thresholds:  
***full\_model = lmer(atrophy ~ PRS \* time + AGE + SEX + EDUCATION + APOE4 + PC1...PC5 + (1 + time | ID))***

Similar covariates as used in previous linear regressions and with the *PRS\*time* comprise the fixed effects. We used a likelihood ratio test to compare full models and reduced models (without the polygenic score-by-time interaction in the reduced model) to compute model p-values:

***reduced\_model = lmer(atrophy ~ PRS + time + AGE + SEX + EDUCATION + APOE4 + PC1...PC5 + (1 + time | ID))***  
***anova(full\_model, reduced\_model)***

We used a likelihood ratio test to compare full models and reduced models (without the polygenic score-by-time interaction) to compute model p-values.

### 3. Results



**Figure 1: Illustration of the selection of subjects filtered for analysis leaving 730 subjects that comprise CORE DATASET 1 and 602 subjects that comprise CORE DATASET 2.**

#### 3.1 Filtering core datasets

A total of 2269 unique subjects were extracted from ADNIMERGE of which 2257 had available baseline diagnoses. To maximise the numbers per diagnostic group, sub-diagnostic groups were merged: (1) "CN" (Control/Normal) and "SMC" (Significant Memory Concern) to "HC" (Healthy Control); (2) "EMCI" (Early Mild Cognitive Impairment) and "LMCI" (Late Mild Cognitive Impairment) to Mild Cognitive Impairment label "MCI"; (3) "AD" (Alzheimer's Disease) remaining unchanged. A subset of these subjects with computed total and regional WMH using BaMoS (see below) results in a cohort of 871 individuals. Further subjects were excluded following addition of AD PRS and later CAD PRS (134,456 and 135,584 variants respectively) and principal components of population structure, leaving  $N=730$  subjects that comprise CORE DATASET 1 (Figure 1).

We augmented CORE DATASET 1 with available boundary shift integral (BSI) data (ADNIMERGE *foxlabbsi* table) to investigate atrophy accelerations, i.e., changes in whole brain volumes over time. Given the paucity of scans at months 8, 18, 36, 48, 60, 72, 84, 96,

120, 132, 144 and 156 (see supplementary *Figure S2*) only scans from months 3, 6, 12 and 24 were retained (*Figure S3*). The remaining atrophy values comprised 824 instances of negative total KN-BSIs (across all months), implying a growth in whole brain volume over time. As this is unlikely these scans were considered to be inaccurate and removed from the sample which resulted in a loss of 51 subjects. This formed CORE DATASET 2 which comprises 1501 scans over N=602 subjects (*Figure 1*).

### 3.1.1 Core dataset summaries

*Table 1* shows the demographics of CORE DATASET 1 and 2 respectively. There were significant differences in age and years of education between the diagnostic categories

**Table 1: Baseline demographics of Table 1a (top) CORE DATASET 1 and Table 1b (bottom) CORE DATASET 2.** For CORE DATASET 1 more than half of the subjects fall in the MCI diagnostic category, almost a third HC and less than a fifth AD. The HC group has a higher fraction of females and as expected the AD participants are slightly older on average and there is a slight decrease in time spent in education for the AD subset (one-way ANOVA - age:  $p=1.49e-05$ ; education:  $p=0.0064$ ). In CORE DATASET 2 we see a similar breakdown of sample sizes, sex, age and education by diagnostic group one-way ANOVA - age:  $p=2.51e-05$ ; education:  $p=0.06$ ).

	HC	MCI	AD	TOTAL
N	218 (29.9%)	390 (53.4%)	122 (16.7%)	730
Female (%)	117(53.7%)	171 (43.8%)	51(41.8%)	339 (46.4%)
Age (SD)	73.4 (5.9)	71.5 (7.5)	74.8 (8.0)	72.6 (7.2)
Education (SD)	16.7 (2.5)	16.3 (2.6)	15.7 (2.7)	16.3 (2.6)

	HC	MCI	AD	TOTAL
N	172 (28.6%)	336 (55.8%)	94 (15.6%)	602
Female (%)	91 (52.9%)	151 (44.9%)	37 (39.4%)	279 (46.3%)
Age (SD)	73.7 (6.1)	71.7 (7.3)	75.1 (7.5)	72.8 (7.1)
Education (SD)	16.6 (2.5)	16.4 (2.6)	15.8 (2.7)	16.3 (2.6)

*Table S1* (and *Figure S3*) show the number of follow-up MRI scans available within each diagnostic group. The proportions match the sample-size distributions. However, only a few scans were available at the 24-month mark for the AD subgroup.

### 3.2 White matter lesions

WMH volume was greater in the AD group (Tukey multiple comparisons of means, AD-to-MCI: $p_{\text{adj}}=0.007$ , AD-to-HC: $p_{\text{adj}}=0.03$ ) and marginally greater in the MCI group over the healthy controls (*Figure S1, a relationship which holds when WMH is corrected for subject age*; WMH volume was also greater in the AD group when broken down by sex; Table S2). Following an exploration of regional WMH (including total frontal, parietal, temporal and occipital WMH volumes (results not shown) the most compelling result was the combined basal ganglia and infratentorial WMH volumes in male subjects, which showed a clear difference between diagnostic groups (*Figure S4*).

### 3.3 Polygenic risk

Our linear regression analysis for an effect between AD PRS or CAD PRS (over both PRS thresholds) on baseline GM, total WMH or CSF amyloid did not find any statistically significant results ( $P_{\text{FDR}} > 0.05$ ; Table 2). However, for individuals with an AD diagnosis both AD PRS ( $P_{\text{FDR}}=0.03$ ) and CAD PRS ( $P_{\text{FDR}}=0.04$ ) are associated with occipital lobe WMH at the PRS-threshold of  $p=0.5$  (Table 4).

**Table 2: Results from CORE DATASET 1 using the low p-cut-off 1e-05 and the high p-cut-off 0.5.** t-values (p-values, one-tailed FDR-corrected) to 2sf. following linear regression with confounders for age, sex, education, APOE-e4 burden and first five principal components of population structure (additional ICV confounder for log transformed WMH regression) across diagnostic groups for AD PRS and CAD PRS; WBV normalised by ICV; AD PRS and CAD PRS both exclude APOE region. For amyloid beta, subjects with missing data were removed from the linear model (n=665).

		HC		MCI		AD	
		AD PRS	CAD PRS	AD PRS	CAD PRS	AD PRS	CAD PRS
P=1e-05	WBV	0.61 (0.30)	1.9 (0.27)	-2.2 (0.07)	1.8 (0.34)	1.0 (0.49)	-0.78 (0.47)
	WMH (tot)	0.21 (0.44)	0.31 (0.39)	-1.08 (0.34)	-1.43 (0.48)	1.93 (0.09)	-0.78 (0.13)
	Amyloid Beta	-0.89 (0.20)	2.4 (0.16)	0.1 (0.48)	1.1 (0.47)	-0.35 (0.38)	0.13 (0.45)
P=0.5	WBV	-0.8 (0.30)	0.85 (0.26)	-0.96 (0.31)	0.24 (0.42)	0.31 (0.49)	-0.23 (0.47)
	WMH (tot)	0.68 (0.36)	0.48 (0.39)	-0.73 (0.34)	-0.12 (0.48)	1.63 (0.09)	1.35 (0.13)
	Amyloid Beta	-0.89 (0.20)	0.13 (0.48)	-0.46 (0.48)	-0.49 (0.47)	-1.2 (0.20)	1.8 (0.12)

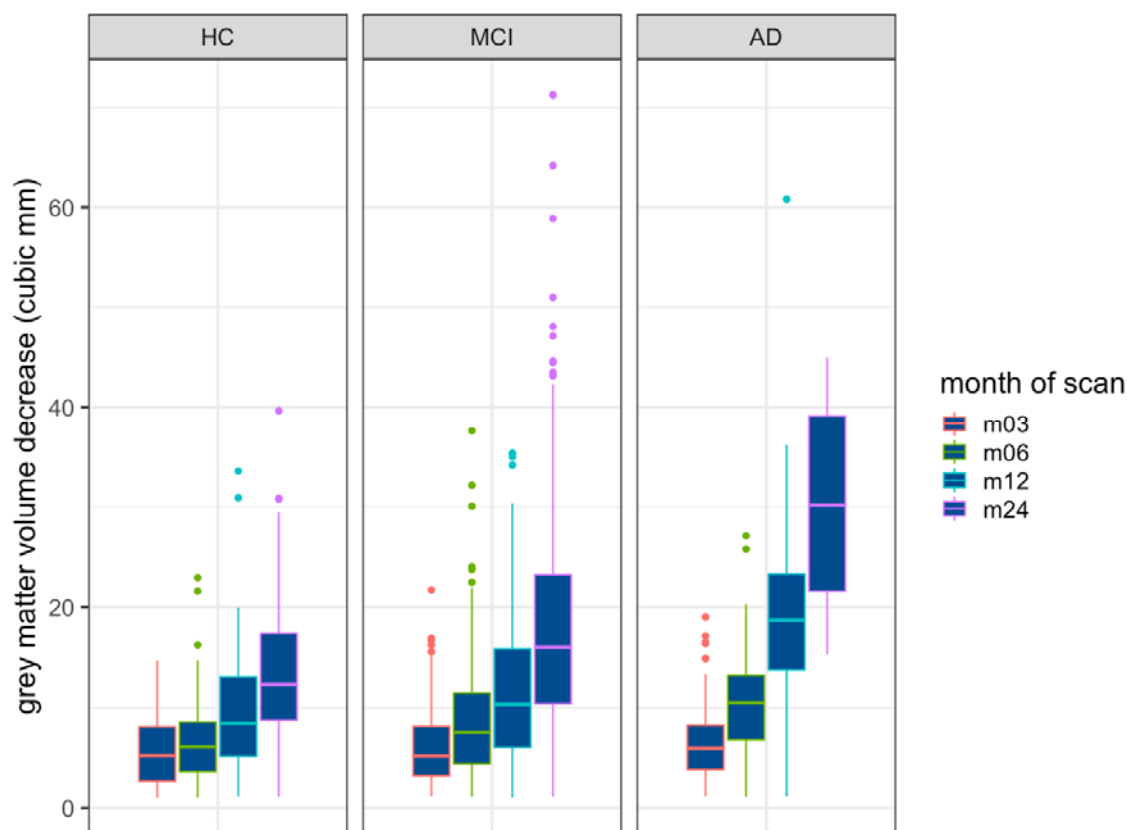
**Table 3: Regional WMH analysis from CORE DATASET 1 using the low p-cut-off 1e-05 and the high p-cut-off 0.5.** T-values (p-values, one-tailed FDR-corrected) to 2dp. following linear regression with confounders for age, sex, education, APOE-e4 burden and first five principal components of population structure and natural log of regional WMH normalised by regional volume.

		HC		MCI		AD	
		AD	CAD	AD	CAD	AD	CAD
		PRS	PRS	PRS	PRS	PRS	PRS
P=1e-05	<b>frontal</b>	0.29 (0.45)	0.24 (0.41)	-0.98 (0.34)	-2.0 (0.21)	1.8 (0.29)	-1.2 (0.25)
	<b>parietal</b>	-0.15 (0.49)	0.52 (0.45)	-1.0 (0.49)	-0.95 (0.49)	1.8 (0.12)	-1.5 (0.16)
	<b>occipital</b>	0.66 (0.45)	0.29 (0.46)	-1.1 (0.17)	-1.6 (0.5)	1.4 (0.06)	-0.96 (0.2)
	<b>temporal</b>	-0.03 (0.44)	1.2 (0.31)	-0.13 (0.42)	-1.3 (0.45)	2.3 (0.38)	-0.69 (0.28)
	<b>basal ganglia + thalami + infratentorial</b>	0.28 (0.46)	-0.15 (0.47)	-1.0 (0.11)	-0.53 (0.46)	0.93 (0.49)	-0.21 (0.47)
P=0.5	<b>frontal</b>	0.64 (0.38)	0.76 (0.27)	-0.26 (0.34)	-0.10 (0.49)	1.1 (0.29)	0.80 (0.25)
	<b>parietal</b>	0.52 (0.49)	0.29 (0.45)	-0.47 (0.47)	-0.09 (0.49)	1.5 (0.12)	1.2 (0.16)
	<b>occipital</b>	0.48 (0.34)	0.27 (0.46)	-0.92 (0.17)	0.03 (0.5)	1.7 (0.03)	2.0 (0.04)
	<b>temporal</b>	0.53 (0.34)	0.69 (0.31)	-1.3 (0.20)	-0.31 (0.45)	2.0 (0.39)	1.3 (0.17)
	<b>basal ganglia + thalami + infratentorial</b>	0.71 (0.46)	-0.42 (0.45)	-0.71 (0.05)	0.15 (0.46)	0.76 (0.49)	0.35 (0.47)

### 3.3 Whole brain atrophy

Given that a snapshot of whole brain volume at baseline may not be sufficiently informative, we investigated the effect of total WMH at baseline and PRS on longitudinal changes in whole brain volumes. Following the baseline scan, KN-BSI increases (representing a decrease in whole brain volume) with passing months in all three diagnostic groups (*Figure 2*), but this is most pronounced in the AD cohort (see *Figure S4* for multiple comparison p-

values). Using linear regression, we found a statistically significant effect ( $P_{FDR} < 0.05$ ) of baseline WMH on WBV atrophy in all diagnostic groups over differing times (*Table 4*). This association also held in a corresponding regional analysis of frontal, parietal, occipital, temporal and combined basal ganglia, thalami and infratentorial WMH volumes (*Table S3*) with WBV drop being particularly correlated over months 12 and 24 (especially in the MCI cohort).



**Figure 2: whole brain atrophy (KN-BSI) in each diagnostic group across all time intervals – m03, m06, m12 and m24 represent 3, 6, 12 and 24 months following baseline scan respectively.**

AD PRS is associated with whole brain atrophy (KN-BSI) in HC and MCI, but not in subjects with AD. Moreover, in the MCI cohort AD PRS is correlated with brain atrophy at each time-point but only for the  $p=1e-5$  PRS threshold. Finally, in the HC cohort the KN-BSI and AD PRS are correlated in the later month 12 and month 24.



There is no statistically significant correlation between CAD PRS and KN-BSI in both the HC and AD diagnostic groups. However, in the MCI cohort we see a correlation in the later month 12 and month 24, but only for the  $p=0.5$  threshold. So in the  $p=0.5$  threshold for month 12 and month 24, the CAD PRS is more correlated with whole brain atrophy than the AD PRS (month 12:  $t=2.2$ ,  $P_{FDR}=0.03$ ; month 24:  $t=2.1$ ,  $P_{FDR}=0.04$ ).

**Table 4: Results from CORE DATASET 2.** T-values (p-values) for a range of KN-BSI whole brain atrophy over scans at months 3, 6, 12 and 24 with respect to natural log of WMH (at baseline or month 0), CAD PRS and AD PRS in different diagnostic groups; linear regression with confounders for age, sex, education, APOE-e4 burden and five principal components of population structure (including ICV confounder for WMH regression) across diagnostic groups for two PRS thresholds; APOE region excluded in AD PRS and CAD PRS. All PRS regression p-values are corrected for multiple testing and are one-tailed.

KNBSI	PRS cut-off	HC			MCI			AD		
		WMH	AD PRS	CAD PRS	WMH	AD PRS	CAD PRS	WMH	AD PRS	CAD PRS
month3	P=1e-5	1.9 (0.053)	0.89 (0.48)	0.68 (0.49)	1.5 (0.14)	2.7 (0.01)	0.17 (0.46)	2.9 (0.0039)	0.84 (0.48)	1.2 (0.27)
	P=0.5		-0.18 (0.47)	-0.04 (0.49)		-0.36 (0.4)	1.9 (0.07)		-0.18 (0.48)	0.95 (0.27)
month6	P=1e-5	1.9 (0.062)	1.8 (0.2)	1.1 (0.3)	2.3 (0.019)	3.1 (0.004)	-0.004 (0.5)	2.49 (0.013)	1.5 (0.32)	0.58 (0.3)
	P=0.5		0.57 (0.39)	0.6 (0.3)		-0.19 (0.49)	1.8 (0.08)		0.2 (0.49)	1.7 (0.1)
month12	P=1e-5	4.2 (3.7e-05)	2.6 (0.04)	0.64 (0.49)	2.5 (0.013)	2.8 (0.01)	-0.007 (0.5)	3 (0.0024)	1.6 (0.27)	-0.23 (0.46)
	P=0.5		1.7 (0.06)	-0.03 (0.49)		-0.49 (0.35)	2.2 (0.03)		-0.77 (0.28)	1.5 (0.15)
month24	P=1e-5	2.7 (0.0076)	3.3 (0.007)	0.63 (0.45)	1.3 (0.2)	3.1 (0.004)	-0.28 (0.4)	0.19 (0.85)	2.3 (0.14)	-0.25 (0.4)
	P=0.5		2.5 (0.01)	-0.26 (0.45)		-0.5 (0.37)	2.1 (0.04)		-0.77 (0.3)	0.6 (0.36)

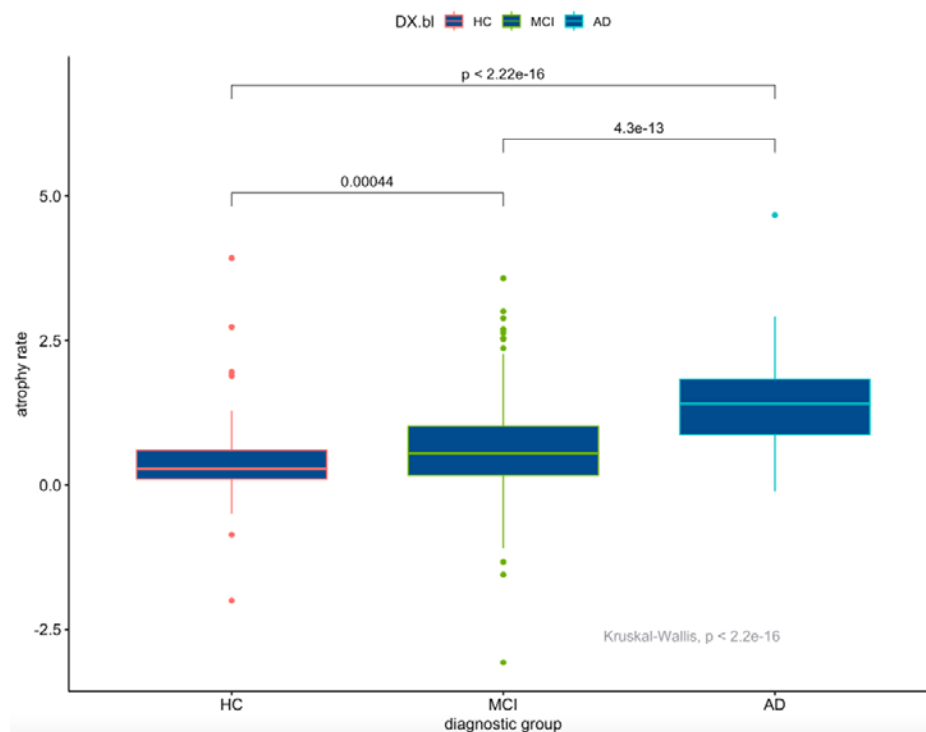
### 3.4 whole brain atrophy acceleration

To probe the *rate of change* of whole brain volume over longitudinal scans per subject with respect to the underlying genetics we enumerate the ‘atrophy acceleration’ in *Table 5a*.

*Atrophy acceleration* is significantly greater in the AD cohort compared to HC and MCI ( $p<4.3e-13$ ; *Figure 3*). While there is no correlation between CAD PRS and atrophy acceleration (*Table 5a*), there is a statistically significant correlation between AD PRS and atrophy acceleration for the  $p=0.5$  threshold in the healthy controls (HC). It is also notable that at the  $p=1e-5$  threshold this correlation shows a statistical trend ( $P_{FDR}=0.06$ ). Our mixed effect analysis (*Table 5b*) showed that there is no association between longitudinal changes in grey matter (quantified by the BSI) and time-by-CAD polygenic score interaction, however,



time-by-AD polygenic score interaction is correlated with longitudinal WBV changes in both the HC ( $P_{FDR} = 0.01$ ) and MCI ( $P_{FDR} = 0.03$ ) subsets at the  $p = 1e-5$  PRS-threshold.



**Figure 3: Median atrophy acceleration in cubic cm/month<sup>2</sup> across diagnostic groups.** AD subjects whole brain volume is decreasing much faster than MCI, which are decreasing faster than HC. Multiple comparison Kruskal-Wallis p-values shown.

**Table 5: Table 5a (top) T-values (p-values) for atrophy acceleration, or, rate of change of whole brain volume (gradient from line of best fit over serial scans at months 3, 6, 12 and 24 per subject) with respect to AD PRS and CAD PRS.** Following linear regression with confounders for age, sex, education, APOE-e4 burden and five principal components of population structure across diagnostic groups for two PRS thresholds; APOE region excluded in AD PRS and CAD PRS. All PRS regression p-values are corrected for multiple testing and are one-tailed. **Table 5b (bottom) P-values from linear mixed effects model with subject as random intercept and time-since-baseline as random slope.** Across diagnostic groups and oligogenic and polygenic thresholds for atrophy acceleration or rate of change whole brain volume (gradient from line of best fit over serial scans at months 3, 6, 12 and 24 per subject) with respect to CAD PRS- and AD PRS-by-time interaction.

Diagnostic group	PRS cut-off	CAD PRS	AD PRS
HC	<b>P=1e-5</b>	-1.35 (0.28)	1.84 (0.06)

<b>MCI</b>	<b>P=0.5</b>	0.58 (0.35)	<b>2.06 (0.04)</b>
	<b>P=1e-5</b>	-1.98 (0.26)	0.77 (0.49)
<b>AD</b>	<b>P=0.5</b>	-0.49 (0.34)	-0.16 (0.49)
	<b>P=1e-5</b>	0.35 (0.47)	-1.21 (0.49)
	<b>P=0.5</b>	-0.3 (0.47)	0.03 (0.49)

Diagnostic group	PRS cut-off	CAD PRS	AD PRS
<b>HC</b>	<b>P=1e-5</b>	0.46	<b>0.01</b>
	<b>P=0.5</b>	0.51	0.05
<b>MCI</b>	<b>P=1e-5</b>	0.27	<b>0.03</b>
	<b>P=0.5</b>	0.58	0.16
<b>AD</b>	<b>P=1e-5</b>	0.31	0.97
	<b>P=0.5</b>	0.86	0.72

#### 4. Discussion [1771 words]

In this study we aimed to investigate possible associations between cardiac genetics and Alzheimer's dementia. We focused upon the genetic variants outside of the APOE locus that have been shown to be associated with AD, and separately, with CAD genetic variants. In particular we looked at their impacts upon white matter lesions and whole brain volume changes. We found a correlation between coronary artery genetic risk and whole brain atrophy suggesting that many small effect-size variants contribute to neuronal loss in mildly cognitively impaired (MCI) individuals. Surprisingly, under certain PRS-thresholds and at certain times in the disease-course the underlying polymorphisms for coronary artery disease may well have more of an impact than those of AD. We have also highlighted how whole brain atrophy acceleration is associated with Alzheimer's disease polygenic risk score outside the APOE locus in healthy individuals. Further, we have shown for the first time that genetic variants that contribute to both AD and CAD beyond APOE are a strong predictor of white matter lesions in the occipital lobe in subjects already diagnosed with AD.

White matter lesions were clearly higher in the AD group, something that is driven in part by age (*Figure S5*). While all linear models treated age as a covariate, disentangling white matter lesion increase as a function of time separate to any age-related AD pathologies is challenging. This may well be of great import if early pharmaceutical intervention is necessary to curtail AD in later life as WMH extent could be utilised as a surrogate biomarker for downstream dementia<sup>39</sup>.

No associations were found between AD PRS or CAD PRS and baseline total WBV and WMH volumes across all diagnostic groups, although PRSs have been shown to be predictive of AD risk<sup>40</sup>. That said, AD PRS beyond APOE is correlated with white matter load in the occipital lobe in AD subjects upon the inclusion of many variants (polygenic  $p=0.5$ ). In fact, our regional analysis of WMH showed statistically significant association between both AD PRS and CAD PRS with occipital lobe WMH lesion volume in the AD group at a PRS-threshold of  $p=0.5$ . This confirms the presence of significant WMH lesions in this lobe seen in other dementia cohorts<sup>41</sup> and the finding that occipital WMH is correlated with reduced executive function<sup>42</sup>. This may also be related to the cerebral amyloid angiopathy pathway. We were unable to establish this white matter load association by occipital lobe layer, however, this is to our knowledge the first confirmation of a genetic link with this regional WMH and a genetic link to coronary artery disease.

There are also no statistically significant correlations with CSF amyloid beta and CAD or AD PRS, presumably we did not find evidence for an association with AD PRS because amyloid biomarkers are mainly influenced by the APOE-e4 genotype<sup>21</sup>. Moreover, the analysis was conducted within disease groups and not between-disease-groups where stronger differences are known to exist<sup>43</sup>.

When it came to investigating whole brain changes over time our analysis confirmed an association of baseline WMH with WBV decline across all three diagnostic groups. WMH here was derived from scans at baseline and not at the time of repeat scans, thus, WMH burden may indeed serve as an indicator of near-term or future decreases in whole brain volume. The strength of these associations was driven by sample size and time between baseline and follow-up WBV measure. These results confirm earlier studies<sup>41,44</sup> that showed higher white matter lesion load is correlated with decreasing whole brain volume, albeit using cross-sectional WBV measures. Earlier studies investigating longitudinal whole brain volume change only found an association with WMH in healthy controls<sup>45</sup> or hippocampal atrophy in MCI subjects<sup>46</sup>, in contrast to our analysis which extends this correlation to participants with AD.

In the matter of whether PRS are predictive of longitudinal decline in WBV, our analysis showed correlations between AD PRS and WBV decline in the HC group (later months) and MCI group (all months), but not the AD group. Decline in the healthy controls group reflects regional WBV decline with respect to AD PRS observed previously<sup>47</sup>. The lack of a polygenic effect in the AD group, may originate from either limited statistical power, the overall extent of WBV damage in AD or that while AD is advanced WBV volume decrease is driven by other factors beyond AD-risk variants. As in the HC group, the AD PRS, like baseline WMH, is correlated with WBV decrease in later months 12 and 24, which raises the question as to whether genetic effects become more influential with time. Before this can be answered we would have to account for there being more samples in months 12 and 24, and the fact that picking up changes in WBV volumes over months 3 and 6 will be more difficult as they will likely be more subtle. Interestingly, the AD PRS correlation with WBV atrophy in the MCI group was significant at all time points for the oligogenic ( $p=1e-5$ ) PRS-threshold but not the polygenic ( $p=0.5$ ) PRS-threshold. This suggests that there are some larger effect-size, common variants (i.e., the peaks in the Manhattan plot outside the *APOE* locus) that are

predictive of brain atrophy in persons that already display some cognitive impairment. This result confirms the hypothesis that a small number of AD-SNPs allow maximal predictive power in AD-related subjects<sup>48</sup>. MCI, however, is a heterogeneous group of which only a subset exhibit MCI due to an underlying AD pathology. This is also the most bimodal group in terms of whether they have AD (progressors vs. not progressors). Thus, the AD PRS may correctly predict WBV decline in subjects who have MCI due to AD.

The motivation for this study was an exploration of the genetic effects of coronary artery disease on Alzheimer's disease. The CAD PRS was only correlated with WBV decline in the MCI group reflecting incidence-based relationships seen in other studies<sup>49</sup>. Again, the PRS-WBV atrophy association was strongest in the later months 12 and 24 but showed a statistical trend in the earlier months 3 and 6. In the MCI group, CAD PRS is significantly correlated with WBV atrophy for polygenic effects ( $p=0.5$  threshold) but AD PRS is significantly correlated with WBV atrophy for oligogenic effects ( $p=1e-5$  threshold). One interpretation might be that the more genetic variants included, the more important the role of cardiac genetics over AD genetics in individuals with mild cognitive impairment. Another interpretation may be that AD PRS is contributing to individuals with MCI due to developing AD dementia, whereas CAD PRS is bestowing subjects a more vascular component to their cognitive loss.

Recent work<sup>50,51</sup> investigating the genetic architecture of AD with regards to PRS thresholds argues that the polygenic threshold is optimal. Their work shows that studies using an oligogenic threshold ignore the fact that there will be fewer APOE-e4 carrying individuals in the older category biasing results against high p-value threshold variants. In our study we removed the APOE locus so as to explore genetic changes outside this genetic region, however, all linear regressions include APOE-e4 burden as a confounding variable. Investigators have shown that APOE's effect on AD is greater in older cohorts and suggest that variants outside of APOE could contribute to AD in older persons<sup>52</sup>. As APOE has been considered a target for both the treatment of coronary heart disease and AD these two pathologies may both be affected by similar pathways. We examined regression models with and without the APOE-e4 covariate using ANOVA. The inclusion of the APOE-e4 burden improved all models significantly ( $p_{adj} < 0.01$ ) with the exception being WBV volume decline in months 3 ( $p_{adj}=0.8$ ) and 6 ( $p_{adj}=0.1$ ) in the HC group (results not shown).

Brain volume is decreasing faster in AD subjects than MCI subjects, where in turn it is decreasing faster than in healthy individuals (Figure 3). The study of polygenic risk suggests that many small effect-size AD variants beyond APOE are a predictor of atrophy acceleration in healthy individuals (and perhaps also MCI subjects). This is in many ways a surprising result as one would expect the AD cohort (and perhaps the MCI cohort) to be comprised of persons whose WBV decrease is accelerating as the pathology develops. Some caution is warranted here given that AD cohort subjects have fewer longitudinal scans thereby biasing the computed gradient. That said, the HC correlation may indicate that genetic effects modulate the speed at which brain cells atrophy during healthy stages. What we do not know is whether a larger atrophy acceleration while healthy makes AD inevitable, more likely or in no way indicative of downstream pathology. However, it is reassuring that the AD PRS in healthy controls is corroborated in the mixed effects models, which are more flexible, embodying both random intercepts and random slopes so as to more realistically capture the subject-level heterogeneity.

The BSI data available through ADNI are not mid-point symmetric: WBV change going from scan A-to-C is not the same as adding WBV volume changes in scan A-to-B and scan B-to-C. The implication is that one has to be cautious about saying anything about the time-varying nature of these non-symmetric BSIs. We investigated this effect using symmetric BSI data<sup>53</sup> for the same ADNIGO and ADNI2 subjects (n=572; 3T, accelerated, identical scanner protocol, removing BSI<0, only months 3, 6, 12 and 24 post baseline). Repeated analysis on this symmetric BSI data resulted in the same statistically significant outcomes as described in this work (results not shown). The only exception to this is that the correlation between baseline WMH and WBV atrophy in the HC group was no longer replicated. We are also cautious about over-interpreting our results: while the BSI does measure changes at the border of the brain, it does not mean that the change has actually occurred there. Tissue may have been lost from the middle of the white matter with this change being measured at the edge of the brain.

Given that both whole brain atrophy and whole brain atrophy acceleration are both shown to be correlated with polygenic scores it may be that utilising such genetic summary information (perhaps alongside other routinely collected health measures) can one day be used at birth (or middle age) as a predictor late-life cognitive problems<sup>54</sup>. If so, and if such risk is driven in part by cardiac health, it may be that this risk can be reduced through wellness and behaviour

changes in, early to mid-life<sup>55</sup>. Of course, this does depend on whether healthy individuals with such brain changes go on to develop dementia, something only larger longitudinal studies over many years can answer.

## 5. Conclusions

The link between CAD and AD is gaining more attention with one recent study identifying 23 brain regions associated with both cardiovascular disease and Alzheimer's dementia<sup>56</sup>. WMH loads have been linked to hypertension, hypercholesterolemia and BMI in middle-aged subjects, all of which also contribute to cardiac health<sup>57</sup>. While our work offers some support for the importance of genetic variants known to effect coronary heart disease also being involved in cognitive decline in the elderly, more data is needed to validate these findings. Similarly, the usefulness of WMHs and atrophy accelerations need to be investigated further. In particular serial WMH measures alongside WBV changes would be a more realistic guide to cerebral changes. Further understanding of this pathway from gene to brain will come from cardiac imaging of the heart linked to CAD genetics alongside routinely measured biomarkers over the life-course of many individuals such as those found in large scale, data-rich, longitudinal cohort studies.



## Acknowledgements

This research was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre. Andre Altmann holds an MRC eMedLab Medical Bioinformatics Career Development Fellowship. This work was supported by the Medical Research Council [grant number MR/L016311/1].

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i-Select chips was funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut

Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD/PERADES was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01-AG-12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC-10-196728.

### **Funding**

E.deS. acknowledges support from the National Institute for Health Research (NIHR) University College London Hospitals Biomedical Research Centre (UCLH BRC). A.A. holds a Medical Research Council eMedLab Medical Bioinformatics Career Development Fellowship. This work was supported by the Medical Research Council [grant number MR/L016311/1]. C.H.S. is supported by an Alzheimer's Society Junior Fellowship (AS-JF-17-011). J.B. was supported by an Alzheimer's Research United Kingdom Senior Research Fellowship. M.A.S. acknowledges financial support the Engineering and Physical Sciences Research Council (EPSRC)-funded UCL Centre for Doctoral Training in Medical Imaging (EP/L016478/1).

### **Competing Interests**

The authors declare no competing interests.

## References

1. Eggink E, Moll van Charante EP, van Gool WA, Richard E. A Population Perspective on Prevention of Dementia. *Journal of clinical medicine*. 2019;8(6):834. doi:10.3390/jcm8060834
2. Gaugler J, James B, Johnson T, Marin A, Weuve J, Assoc As. 2019 Alzheimer's disease facts and figures. *Alzheimers & Dementia*. Mar 2019;15(3):321-387. doi:10.1016/j.jalz.2019.01.010
3. Villemagne VL, Dore V, Burnham SC, Masters CL, Rowe CC. Imaging tau and amyloid-beta proteinopathies in Alzheimer disease and other conditions. *Nat Rev Neurol*. Apr 2018;14(4):225-236. doi:10.1038/nrneurol.2018.9
4. Edwards FA. A Unifying Hypothesis for Alzheimer's Disease: From Plaques to Neurodegeneration. *Trends in Neurosciences*. May 2019;42(5):310-322. doi:10.1016/j.tins.2019.03.003
5. Abeyasinghe A, Deshapriya R, Udawatte C. Alzheimer's disease; a review of the pathophysiological basis and therapeutic interventions. *Life Sci*. Sep 1 2020;256:117996. doi:10.1016/j.lfs.2020.117996
6. Francis PT, Ramirez MJ, Lai MK. Neurochemical basis for symptomatic treatment of Alzheimer's disease. *Neuropharmacology*. Sep-Oct 2010;59(4-5):221-9. doi:10.1016/j.neuropharm.2010.02.010
7. Fierini F. Mixed dementia: Neglected clinical entity or nosographic artifice? *J Neurol Sci*. Mar 15 2020;410:116662. doi:10.1016/j.jns.2019.116662
8. Ylikoski A, Erkinjuntti T, Raininko R, Sarna S, Sulkava R, Tilvis R. White matter hyperintensities on MRI in the neurologically nondiseased elderly. Analysis of cohorts of consecutive subjects aged 55 to 85 years living at home. *Stroke*. Jul 1995;26(7):1171-7. doi:10.1161/01.str.26.7.1171
9. Garde E, Mortensen EL, Krabbe K, Rostrup E, Larsson HBW. Relation between age-related decline in intelligence and cerebral white-matter hyperintensities in healthy octogenarians: a longitudinal study. *Lancet*. Aug 19 2000;356(9230):628-634. doi:10.1016/S0140-6736(00)02604-0
10. Snyder HM, Corriveau RA, Craft S, et al. Vascular contributions to cognitive impairment and dementia including Alzheimer's disease. *Alzheimers Dement*. Jun 2015;11(6):710-7. doi:10.1016/j.jalz.2014.10.008
11. Frey BM, Petersen M, Mayer C, Schulz M, Cheng B, Thomalla G. Characterization of White Matter Hyperintensities in Large-Scale MRI-Studies. *Front Neurol*. 2019;10:238. doi:10.3389/fneur.2019.00238
12. Khera AV, Kathiresan S. Genetics of coronary artery disease: discovery, biology and clinical translation. *Nat Rev Genet*. Jun 2017;18(6):331-344. doi:10.1038/nrg.2016.160

13. Van Giau V, Bagyinszky E, Yang YS, Youn YC, An SSA, Kim SY. Genetic analyses of early-onset Alzheimer's disease using next generation sequencing. *Scientific Reports*. Jun 10 2019;9doi:ARTN 8368  
10.1038/s41598-019-44848-2
14. Gatz M, Reynolds CA, Fratiglioni L, *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry*. Feb 2006;63(2):168-74.  
doi:10.1001/archpsyc.63.2.168
15. Ridge PG, Hoyt KB, Boehme K, *et al.* Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol Aging*. May 2016;41:200 e13-200 e20.  
doi:10.1016/j.neurobiolaging.2016.02.024
16. Purcell SM, Wray NR, Stone JL, *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. Aug 6 2009;460(7256):748-752.  
doi:10.1038/nature08185
17. Escott-Price V, Sims R, Bannister C, *et al.* Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain*. Dec 2015;138(Pt 12):3673-84.  
doi:10.1093/brain/awv268
18. Chaudhury S, Brookes KJ, Patel T, *et al.* Alzheimer's disease polygenic risk score as a predictor of conversion from mild-cognitive impairment. *Transl Psychiatry*. May 24 2019;9(1):154. doi:10.1038/s41398-019-0485-7
19. Morgan AR, Touchard S, O'Hagan C, *et al.* The Correlation between Inflammatory Biomarkers and Polygenic Risk Score in Alzheimer's Disease. *Journal of Alzheimers Disease*. 2017;56(1):25-36. doi:10.3233/Jad-160889
20. Martiskainen H, Helisalmi S, Viswanathan J, *et al.* Effects of Alzheimer's disease-associated risk loci on cerebrospinal fluid biomarkers and disease progression: a polygenic risk score approach. *J Alzheimers Dis*. 2015;43(2):565-73. doi:10.3233/JAD-140777
21. Altmann A, Scelsi MA, Shoai M, *et al.* A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE. *Brain Communications*. 2020;2(1)doi:10.1093/braincomms/fcz047
22. Lupton MK, Strike L, Hansell NK, *et al.* The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. *Neurobiol Aging*. Apr 2016;40:68-77. doi:10.1016/j.neurobiolaging.2015.12.023
23. Sabuncu MR, Buckner RL, Smoller JW, *et al.* The Association between a Polygenic Alzheimer Score and Cortical Thickness in Clinically Normal Subjects. *Cerebral Cortex*. Nov 2012;22(11):2653-2661. doi:10.1093/cercor/bhr348
24. Cruchaga C, Del-Aguila JL, Saef B, *et al.* Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimers Dement*. Feb 2018;14(2):205-214. doi:10.1016/j.jalz.2017.08.013

25. Chasioti D, Yan JW, Nho K, Saykin AJ. Progress in Polygenic Composite Scores in Alzheimer's and Other Complex Diseases. *Trends in Genetics*. May 2019;35(5):371-382. doi:10.1016/j.tig.2019.02.005
26. Sudre CH, Cardoso MJ, Bouvy WH, Biessels GJ, Barnes J, Ourselin S. Bayesian model selection for pathological neuroimaging data applied to white matter lesion segmentation. *IEEE Trans Med Imaging*. Oct 2015;34(10):2079-102. doi:10.1109/TMI.2015.2419072
27. Walsh P, Sudre CH, Fiford CM, *et al.* CSF amyloid is a consistent predictor of white matter hyperintensities across the disease course from aging to Alzheimer's disease. *Neurobiology of Aging*. 2020/07/01/ 2020;91:5-14. doi:<https://doi.org/10.1016/j.neurobiolaging.2020.03.008>
28. Cardoso MJ, Modat M, Wolz R, *et al.* Geodesic Information Flows: Spatially-Variant Graphs and Their Application to Segmentation and Fusion. *IEEE Transactions on Medical Imaging*. 2015;34(9):1976-1988. doi:10.1109/TMI.2015.2418298
29. Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE Trans Med Imaging*. Oct 1997;16(5):623-9. doi:10.1109/42.640753
30. Leung KK, Clarkson MJ, Bartlett JW, *et al.* Robust atrophy rate measurement in Alzheimer's disease using multi-site serial MRI: Tissue-specific intensity normalization and parameter selection. *Neuroimage*. Apr 1 2010;50(2):516-523. doi:10.1016/j.neuroimage.2009.12.059
31. Manning EN, Leung KK, Nicholas JM, *et al.* A Comparison of Accelerated and Non-accelerated MRI Scans for Brain Volume and Boundary Shift Integral Measures of Volume Change: Evidence from the ADNI Dataset. *Neuroinformatics*. Apr 2017;15(2):215-226. doi:10.1007/s12021-017-9326-0
32. Saykin AJ, Shen L, Foroud TM, *et al.* Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. *Alzheimers Dement*. May 2010;6(3):265-273. doi:10.1016/j.jalz.2010.03.013
33. Scelsi MA, Khan RR, Lorenzi M, *et al.* Genetic study of multimodal imaging Alzheimer's disease progression score implicates novel loci. *Brain*. Jul 1 2018;141(7):2167-2180. doi:10.1093/brain/awy141
34. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8
35. Kunkle BW, Grenier-Boley B, Sims R, *et al.* Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A beta, tau, immunity and lipid processing. *Nature Genetics*. Mar 2019;51(3):414-+. doi:10.1038/s41588-019-0358-2

36. Nikpay M, Goel A, Won HH, *et al.* A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*. Oct 2015;47(10):1121+. doi:10.1038/ng.3396
37. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience*. Jul 1 2019;8(7)doi:10.1093/gigascience/giz082
38. Kerminen S, Martin AR, Koskela J, *et al.* Geographic Variation and Bias in the Polygenic Scores of Complex Diseases and Traits in Finland. *American Journal of Human Genetics*. Jun 6 2019;104(6):1169-1181. doi:10.1016/j.ajhg.2019.05.001
39. d'Arbeloff T, Elliott ML, Knodt AR, *et al.* White matter hyperintensities are common in midlife and already associated with cognitive decline. *Brain Communications*. 2019;1(1)doi:10.1093/braincomms/fcz041
40. Baker E, Escott-Price V. Polygenic Risk Scores in Alzheimer's Disease: Current Applications and Future Directions. Mini Review. *Frontiers in Digital Health*. 2020-August-11 2020;2(14)doi:10.3389/fdgh.2020.00014
41. Brugulat-Serrat A, Salvadó G, Operto G, *et al.* White matter hyperintensities mediate gray matter volume and processing speed relationship in cognitively unimpaired participants. *Human Brain Mapping*. 2020;41(5):1309-1322. doi:<https://doi.org/10.1002/hbm.24877>
42. Jiang J, Paradise M, Liu T, *et al.* The association of regional white matter lesions with cognition in a community-based cohort of older individuals. *NeuroImage: Clinical*. 2018/01/01/ 2018;19:14-21. doi:<https://doi.org/10.1016/j.nicl.2018.03.035>
43. Saddiki H, Fayosse A, Cognat E, *et al.* Age and the association between apolipoprotein E genotype and Alzheimer disease: A cerebrospinal fluid biomarker-based case-control study. *PLOS Medicine*. 2020;17(8):e1003289. doi:10.1371/journal.pmed.1003289
44. Wang Y, Yang Y, Wang T, Nie S, Yin H, Liu J. Correlation between White Matter Hyperintensities Related Gray Matter Volume and Cognition in Cerebral Small Vessel Disease. *Journal of Stroke and Cerebrovascular Diseases*. 2020/12/01/ 2020;29(12):105275. doi:<https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105275>
45. Barnes J, Carmichael OT, Leung KK, *et al.* Vascular and Alzheimer's disease markers independently predict brain atrophy rate in Alzheimer's Disease Neuroimaging Initiative controls. *Neurobiology of Aging*. 2013/08/01/ 2013;34(8):1996-2002. doi:<https://doi.org/10.1016/j.neurobiolaging.2013.02.003>
46. Fiford CM, Manning EN, Bartlett JW, *et al.* White matter hyperintensities are associated with disproportionate progressive hippocampal atrophy. *Hippocampus*. Mar 2017;27(3):249-262. doi:10.1002/hipo.22690

47. Harrison TM, Mahmood Z, Lau EP, *et al.* An Alzheimer's Disease Genetic Risk Score Predicts Longitudinal Thinning of Hippocampal Complex Subregions in Healthy Older Adults. *eneuro*. 2016;3(3):ENEURO.0098-16.2016. doi:10.1523/eneuro.0098-16.2016
48. Zhang Q, Sidorenko J, Couvy-Duchesne B, *et al.* Risk prediction of late-onset Alzheimer's disease implies an oligogenic architecture. *Nature Communications*. 2020/09/23 2020;11(1):4799. doi:10.1038/s41467-020-18534-1
49. Elman JA, Panizzon MS, Logue MW, *et al.* Genetic risk for coronary heart disease alters the influence of Alzheimer's genetic risk on mild cognitive impairment. *Neurobiology of Aging*. 2019/12/01/ 2019;84:237.e5-237.e12. doi:<https://doi.org/10.1016/j.neurobiolaging.2019.06.001>
50. Ganna L, Emily B, Josua S-H, *et al.* Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores is most accurate when using all genetic information. *Nature Portfolio*. 2021/06/16 2021;doi:10.21203/rs.3.rs-137252/v1
51. Leonenko G, Baker, Emily, *et al.* Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores is most accurate when using all genetic information. *Nature Portfolio*. 2021/06/16 2021;doi:10.21203/rs.3.rs-137252/v1
52. Bellou E, Baker E, Leonenko G, *et al.* Age-dependent effect of APOE and polygenic component on Alzheimer's disease. *Neurobiology of Aging*. 2020/09/01/ 2020;93:69-77. doi:<https://doi.org/10.1016/j.neurobiolaging.2020.04.024>
53. Leung KK, Ridgway GR, Ourselin S, Fox NC. Consistent multi-time-point brain atrophy estimation from the boundary shift integral. *Neuroimage*. Feb 15 2012;59(4):3995-4005. doi:10.1016/j.neuroimage.2011.10.068
54. Khera AV, Chaffin M, Aragam KG, *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*. Sep 2018;50(9):1219-+. doi:10.1038/s41588-018-0183-z
55. Abell JG, Kivimäki M, Dugravot A, *et al.* Association between systolic blood pressure and dementia in the Whitehall II cohort study: role of age, duration, and threshold used to define hypertension. *European Heart Journal*. 2018;39(33):3119-3125. doi:10.1093/eurheartj/ehy288
56. Lamar M, Boots EA, Arfanakis K, Barnes LL, Schneider JA. Common Brain Structural Alterations Associated with Cardiovascular Disease Risk Factors and Alzheimer's Dementia: Future Directions and Implications. *Neuropsychol Rev*. Oct 3 2020;doi:10.1007/s11065-020-09460-6
57. Salvadó G, Brugulat-Serrat A, Sudre CH, *et al.* Spatial patterns of white matter hyperintensities associated with Alzheimer's disease risk factors in a cognitively healthy middle-aged cohort. *Alzheimer's Research & Therapy*. 2019/01/24 2019;11(1):12. doi:10.1186/s13195-018-0460-1