

## ***APOE* ε4 effects on hippocampal atrophy in the healthy elderly reflect future cognitive decline**

Linda Zhang<sup>a</sup>, Miguel Calero<sup>b,c,d</sup>, Miguel Medina<sup>c,d</sup>, Bryan A. Strange<sup>a,e</sup>

<sup>a</sup>Department of Neuroimaging, Alzheimer's Disease Research Unit, CIEN Foundation, Queen Sofia Foundation Alzheimer Centre, Madrid, Spain.

<sup>b</sup>Chronic Disease Programme (UFIEC), Instituto de Salud Carlos III, Madrid, Spain.

<sup>c</sup>Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain

<sup>d</sup>Alzheimer Disease Research Unit, CIEN Foundation, Queen Sofia Foundation Alzheimer Centre, Madrid, Spain

<sup>e</sup>Laboratory for Clinical Neuroscience, CTB, Universidad Politécnica de Madrid, Madrid, Spain

### **Corresponding Author:**

Linda Zhang  
Calle de Valderrebollo, 5  
Centro Alzheimer Fundación Reina Sofía  
28031, Madrid, Spain  
Phone: +34 913-852-200 ext. 5007  
Email: [zhangl87@connect.hku.hk](mailto:zhangl87@connect.hku.hk)

### **Author Emails:**

Miguel Calero [mcalero@isciii.es](mailto:mcalero@isciii.es)  
Miguel Medina [mmedina@ciberned.es](mailto:mmedina@ciberned.es)  
Bryan A. Strange [bryan.strange@upm.es](mailto:bryan.strange@upm.es)

## **Abstract**

The *APOE*  $\epsilon 4$  allele is the primary genetic risk factor for late onset Alzheimer's disease (AD). A cardinal problem in determining *APOE*  $\epsilon 4$ 's effect on cognition and brain structure in older individuals is dissociating prodromal changes – linked to increased AD risk – from potential phenotypic differences. To address this, we used cognitive and neuroimaging data from a large cohort of cognitively normal 69-86 year-olds with up to 8 yearly follow-ups to investigate cross-sectional and longitudinal differences between *APOE*  $\epsilon 3/\epsilon 3$  homozygotes and  $\epsilon 3/\epsilon 4$  heterozygotes. Although we found a significant age-by-genotype interaction in right hippocampal volume, once our analyses were conditionalised by future diagnosis to account for prodromal mild cognitive impairment (MCI) and AD, this effect was no longer observed. Likewise, longitudinally, rate of hippocampal atrophy was determined not by genotype, but by future diagnosis. Thus, we provide direct evidence in support of the prodromal hypothesis of *APOE*  $\epsilon 4$  on brain structure.

**Keywords:** *APOE*  $\epsilon 4$ , hippocampal volumetry, longitudinal cohort, prodromal hypothesis, MRI, cognitively normal elderly

## 1. Introduction

Alzheimer's disease (AD) is the leading cause of dementia and of global concern given our ageing population. As AD pathology accumulates many years before detectable symptoms (Jack et al., 2013), there has been increased focus on finding at-risk individuals in nondemented populations to allow for early intervention. The apolipoprotein E (*APOE*)  $\epsilon 4$  allele is highly associated with AD susceptibility (Corder et al., 1993; Farrer et al., 1997; Liu et al., 2013), where each  $\epsilon 4$  allele represents an approximately 4-fold increased risk for developing sporadic AD (Mayeux et al., 1993; Tsai et al., 1994). It is therefore important to understand the mechanisms by which *APOE*  $\epsilon 4$  increases AD risk, which may inform the pathophysiology of AD and lead to improved biomarkers and earlier interventions. However, there is an inherent difficulty in determining if observed *APOE* effects on cognition or brain structure are due to an underlying *APOE* endophenotype, or the presence of prodromal AD pathology, particularly in middle- to older-aged individuals.

The 'phenotype hypothesis' (Greenwood and Parasuraman, 2003) posits that all  $\epsilon 4$  carriers suffer from certain deficits associated with the production of apolipoprotein-E protein, a lipid carrier in the brain, creating an *APOE*  $\epsilon 4$  cognitive phenotype independent of future AD diagnosis. This is supported by studies of young or middle-aged adults, where cognitive differences are observed between  $\epsilon 4$  carriers and noncarriers at an age before any AD pathology is likely to occur (Nao et al., 2017; Salvato et al., 2016). Conversely, the 'prodromal hypothesis' (Smith et al., 1998) proposes that any cognitive or structural differences observed between  $\epsilon 4$  carriers and noncarriers are driven by preclinical AD pathology, which  $\epsilon 4$  carriers are more likely to develop. As such, studies in elderly normal populations should not find significant genotype effects once future diagnosis is taken into account (Foster et al., 2013; Knight et al., 2014).

Systematic reviews of *APOE* effects on cognition (O'Donoghue et al., 2018) and functional MRI (Trachtenberg et al., 2012) have noted that methodological issues, such as limited sample size, broad age range, differences in cognitive assessment batteries or fMRI paradigms, or lack of longitudinal follow-up to account for future dementia, have led to inconsistent findings. This has also been true of studies conducted using structural MRI (den Heijer et al., 2002; Honea et al., 2009; Schuff et al., 2009; Wishart et al., 2006), of which there have been few large longitudinal studies focusing on elderly cognitively normal subjects. Of those with relatively large sample sizes, one study on nondemented  $\epsilon 4$  carriers ( $\epsilon 3/\epsilon 4$  genotype  $n=189$ ,  $\epsilon 4$  noncarriers  $n=270$ , mean age~76 years) found greater hippocampal atrophy in  $\epsilon 4$  carriers over 2 years (Shi et al., 2014). However, the comparison was conducted in a mixed group, where subjects with mild cognitive impairment (MCI) were combined with cognitively normal subjects. In contrast, another study on the same population, but separated into controls, MCI, and AD, found no significant genotype differences in controls (Manning et al., 2014).

In view of these conflicting results, we tested the effect of *APOE*  $\epsilon 4$  on neuropsychological scores and brain structure in a large longitudinal cohort of cognitively healthy 69-86 year-olds, with annual follow-ups over 8 years. We conducted cross-sectional and longitudinal analyses and – to specifically test the prodromal hypothesis – we conditionalised these effects based on future diagnosis (remained normal or converted to MCI/AD).

## **2. Materials and methods**

### ***2.1 Participants***

All participants are part of the Vallecas Project, a single-centre longitudinal community-based study, currently in its tenth year and undergoing the ninth annual visit. Genetic testing was conducted at baseline visit, and annual clinical, neuropsychological, and MRI

assessments remain ongoing. Inclusion and exclusion criteria have been described previously (Olazaran et al., 2015). From a pool of 1213 participants, we set additional criteria, including only those with MRI scanning and without a diagnosis of MCI or dementia, leaving a total of 916 participants on the first visit for cross-sectional analyses. Subjects for longitudinal analyses were selected from this group and are described below. All participants provided written informed consent. The Vallecas project was approved by the Ethics Committee of the Instituto de Salud Carlos III.

## ***2.2 APOE genotyping***

Total DNA was isolated from peripheral blood following standard procedures. Genotyping of *APOE* polymorphisms (rs429358 and rs7412) was determined by Real-Time PCR (Calero et al., 2009). Failure rate of genotyping was 0.3%. The frequency of the *APOE*  $\epsilon$ 4 allele in our cohort is 17.6%, consistent with previous findings in the Spanish and Southern European population (Calero et al., 2011; Mattsson et al., 2018).

## ***2.3 Neuropsychological assessment***

Participants underwent a comprehensive neuropsychological battery including the following tests: Cognitive performance: Mini Mental State Exam (MMSE), Free and Cued Selective Reminding Test (FCSRT, total delayed recall), Rey-Osterrieth Complex Figure (ROCF, delayed recall score), and phonological and semantic verbal fluency; Depression and Anxiety: Geriatric Depression Scale (GDS), State-Trait Anxiety Inventory (STAI); Functional scales: Clinical Dementia Rating (CDR), Functional Activities Questionnaire (FAQ).

## ***2.4 Structural and diffusion MRI image acquisition***

All T1-weighted images (3D fast spoiled gradient echo with inversion recovery preparation) were acquired using a 3 Tesla MRI (Signa HDxt GEHC, Waukesha, USA) with a phased array 8 channel head coil and the following parameters: repetition time (TR) 10 ms, echo

time (TE) 4.5 ms, inversion time (TI) 600 ms, field of view (FOV) 240 mm, matrix 288x288 and slice thickness 1 mm, yielding 0.5x0.5x1 mm<sup>3</sup> voxel size. All MRI scans were reported by a neuroradiologist, including Fazekas scoring (Fazekas et al., 1987). The latter was recorded as the higher value of periventricular or deep white matter hyperintensities score, and was assessed using fluid-attenuated inversion recovery (FLAIR) images (image parameters: TR 9000 ms, TE 130 ms, TI 2100 ms, FOV 24 mm, slice thickness 3.4 mm). Diffusion-weighted images were single shot spin echo echo-planar imaging (SE-EPI), with the following parameters: TR 9200 ms, TE 80 ms, b-value 800 s/mm<sup>2</sup> and 21 gradient directions, FOV 240 mm, matrix 128x128, slice thickness 3 mm.

## **2.5 MRI image analysis**

### *2.5.1 Comparison of grey matter density between genotypes using voxel-based morphometry*

Two voxel-based morphometry (VBM) (Good et al., 2001) analysis pipelines were employed to compare grey matter density (GMD) between *APOE* ε3/ε3 vs. *APOE* ε3/ε4 genotypes. The first was performed using SPM12 software (<http://www.fil.ion.ucl.ac.uk/spm/>). For each subject, the T1-weighted structural image was first segmented into grey matter, white matter and cerebrospinal fluid (CSF), and a nonlinear spatial registration technique (DARTEL) applied to grey matter tissue maps. Segmented grey matter images were normalised to the Montreal Neurological Institute (MNI) standard anatomical space and then smoothed with a Gaussian kernel of 6 mm full width at half maximum (FWHM). Ensuing images were entered into a two-sample *t*-test, comparing whole-brain differences in GMD of *APOE* ε3/ε3 vs. *APOE* ε3/ε4 genotypes. In a first analysis, individual total intracranial volume (TIV) values, obtained by summing the volumes of the grey matter, white matter and CSF, were included as a covariate of no interest. The two-sample *t*-test was then repeated, including

TIV, age, sex, and years of education as covariates of no interest. Lastly, we tested for a genotype-by-age interaction. All covariates were mean-centred prior to model estimation.

The second VBM analysis was implemented using FSL-VBM (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FSLVBM>), an optimised VBM protocol (Good et al., 2001) carried out with FSL tools (Smith et al., 2004). Two independent software analysis pipelines (SPM and FSL) were used as these can sometimes show disparate results (Rajagopalan and Pioro, 2015). Structural images were first brain-extracted and grey matter-segmented before being registered to the MNI standard space using non-linear registration (Andersson et al., 2007). The resulting images were averaged and flipped along the x-axis to create a left-right symmetric, study-specific grey matter template. All native grey matter images were then non-linearly registered to this study-specific template and "modulated" to correct for local expansion (or contraction) due to the non-linear component of the spatial transformation. The modulated grey matter images were then smoothed with an isotropic Gaussian kernel with a sigma of 3 mm (equivalent FWHM of approximately 7 mm). Finally, a voxelwise general linear model (GLM) was applied using permutation-based non-parametric testing, correcting for multiple comparisons across space. As with the SPM analyses, TIV was initially included as a covariate of no interest, followed by the inclusion of age, sex, and years of education.

### *2.5.2 Comparison of white matter integrity between genotypes using tract-based statistics*

To study structural integrity of white matter tracts, diffusion weighted images were processed using Tract-Based Spatial Statistics (TBSS) (Smith et al., 2006), within FSL. Raw diffusion tensor imaging (DTI) images were preprocessed to remove the effect of head movement and eddy current distortions. Fractional anisotropy (FA) images were then created by fitting a tensor model to the preprocessed diffusion data using the FMRIB's Diffusion Toolbox (FDT),

which were then skull-stripped using the Brain Extraction Tool (BET). FA data from all subjects were aligned to a common space using nonlinear registration (Andersson et al., 2007). Finally, a mean FA image was created and thinned to create the mean FA skeleton, which represents the centres of all tracts common to the group. The aligned FA images of each subject were then projected onto the mean FA skeleton and these data were used in voxelwise between-group GLM analyses. TIV, age, sex, and years of education were included as covariates. The same methodology was used to obtain and analyse mean diffusivity (MD) images.

### *2.5.3 Hippocampal volumetry*

Automatic segmentation of the hippocampus was performed on each participant's T1-weighted image using FreeSurfer v.6.0 (<https://surfer.nmr.mgh.harvard.edu/>). Technical details of the whole-brain segmentation methods have been described previously (Fischl et al., 2002). Hippocampal volumes were extracted using the hippocampal subfields module in FreeSurfer 6.0 (Iglesias et al., 2015), and total hippocampal volumes were a sum of the following subfields: subiculum, CA1-4, molecular layers, hippocampal tail, dentate gyrus, granular cell layer, and the hippocampus-amygdala transition area (HATA). Segmentations for all participants were visually inspected for accuracy.

### *2.5.4 Longitudinal analysis of hippocampal volume*

As well as yearly detailed assessments of current participants, a follow-up telephone interview was conducted in the 9<sup>th</sup> year of the Vallecas Project for all participants, including those who had dropped out of the study in the intervening years. It was therefore possible to obtain current diagnostic status for all subjects. In-person assessment was not possible due to the Covid-19 pandemic. Of the 804 subjects included in the cross-sectional analysis, 702 remained cognitively normal and 102 had converted to MCI or AD (of which 29 were

diagnosed or reported via telephone interview). Of the cognitively normal group, 433 had at least 2 yearly MRIs available (*APOE*  $\epsilon 3/\epsilon 3$ :  $n=369$  (67% female), *APOE*  $\epsilon 3/\epsilon 4$ :  $n=64$  (52% female); average number of MRIs available: 4.8). 57 of the converters had at least 2 yearly MRIs available, with at least 1 MRI before conversion to MCI (*APOE*  $\epsilon 3/\epsilon 3$ :  $n=37$  (57% female), average number of years pre-conversion: 2.6; *APOE*  $\epsilon 3/\epsilon 4$ :  $n=20$  (65% female), average number of years pre-conversion: 2.4). These subjects were analysed longitudinally to determine whether genotype affected hippocampal atrophy rates, and if this occurred differentially between those who remained cognitively normal and those who subsequently developed MCI. As we were interested in  $\epsilon 4$  effects in those with normal cognition, MRIs after conversion were not analysed.

To extract reliable volume estimates, subject T1s were automatically processed with the longitudinal stream in FreeSurfer v.6.0, the methods of which were described in detail previously (Reuter et al., 2012). Specifically, an unbiased within-subject template space and image is created using robust, inverse consistent registration. Several processing steps, such as skull stripping, Talairach transforms, atlas registration, as well as spherical surface maps and parcellations are then initialised with common information from the within-subject template, significantly increasing reliability and statistical power. Likewise, the longitudinal hippocampal subfields module uses the same within-subject template to produce robust subfield volume estimates (Iglesias et al., 2016). The longitudinal hippocampal volumes were comprised of the same subfields as for the cross-sectional analysis.

## ***2.6 Statistical analyses***

All demographic, neuropsychological, and hippocampal volumetry data were analysed using Stata v.15 (StataCorp, 2017, Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC). Power calculation was performed using G\*Power 3.1.9.2

(<http://www.gpower.hhu.de/>). Correction for multiple comparisons was conducted via false discovery rate (FDR) estimation within RStudio v1.1.442 (RStudio Team 2016. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. <http://www.rstudio.com>) using the *padjust* function. Hardy-Weinberg equilibrium and sex distribution was tested with the  $\chi^2$  test. One-way ANOVAs were conducted on continuous variables (age, years of education, neuropsychological test scores, hippocampal volumes, total intracranial volumes) to test for differences between genotypes. For neuropsychological test scores, age, sex, and years of education were included as covariates. For hippocampal volumes, left and right hippocampi were tested separately, with age, sex, years of education, and TIV (derived from SPM12 as described in section 2.5.1) included as covariates. In the cross-sectional analysis, as well as testing for a main effect of genotype, a genotype-by-age interaction was tested with the same covariates. In the presence of genotype-by-age interactions, subjects were divided into age-at-baseline groups (69 to <75, 75 to <80, 80 to 86) and a one-way ANOVA was performed between the age groups for *APOE*  $\epsilon 3/\epsilon 3$  and *APOE*  $\epsilon 3/\epsilon 4$  separately.

For longitudinal hippocampal volume analysis, linear mixed-effects models were implemented in Stata v.15, with hippocampal volume as the dependent variable, age, genotype (1 for  $\epsilon 4$  carrier, 0 for noncarriers), future diagnosis (0 for control, 1 for future converter), sex, years of education, and TIV as fixed effects, and subject as a random effect. Age and TIV were scaled to avoid convergence errors. Two-way interactions of genotype-by-age, diagnosis-by-age, and sex-by-age were tested, as well as three-way interactions of genotype-by-diagnosis-by-age, genotype-by-sex-by-age, and diagnosis-by-sex-by-age. Each model was adjusted for years of education and TIV, and sex, genotype, and diagnosis if they were not included in an interaction. To account for whole brain atrophy, all models were also fitted adjusting for brain volume (summed cortical and subcortical grey and white matter) instead of TIV.

### 3. Results

#### 3.1 Genotyping

Of the 916 initial subjects, 3 individuals have the *APOE*  $\epsilon 2/\epsilon 2$  genotype (0.3%), 97 are  $\epsilon 2/\epsilon 3$  (10.6%), 657 are  $\epsilon 3/\epsilon 3$  (71.7%), 7 are  $\epsilon 2/\epsilon 4$  (0.8%), 147 are  $\epsilon 3/\epsilon 4$  (16.0%), and 5 are  $\epsilon 4/\epsilon 4$  (0.5%). The distribution of *APOE* genotypes in our subjects was in Hardy-Weinberg equilibrium ( $\chi^2=2.176$ ;  $p=0.537$ ). As in other studies (Lyall et al., 2016; Wisdom et al., 2011), participants with the  $\epsilon 2$  allele were excluded from neuropsychological and neuroimaging analyses due to the potential protective effects of the  $\epsilon 2$  allele, both in cognition and cortical thinning (Kim et al., 2017). We also excluded the  $\epsilon 4/\epsilon 4$  genotype given the very small number of these homozygotes in our sample. Thus, our final analyses were conducted on those participants with the *APOE*  $\epsilon 3/\epsilon 3$  or  $\epsilon 3/\epsilon 4$  genotypes ( $n=804$ , 67% female).

#### 3.2 Demographics and neuropsychological data

As shown in **Table 1**, we found no significant baseline differences of genotype on any demographic or neuropsychological variable at an alpha of 0.05 uncorrected for multiple comparisons, apart from sex ( $\chi^2=4.38$ ,  $p=0.036$ ), which did not survive FDR correction. For independent *t*-tests, we performed a post-hoc power analysis on these comparisons. Given our sample size, an error probability  $\alpha=0.05$ , and a moderate effect size (Cohen's  $d=0.5$ ), the power ( $1 - \beta$ ) is 1. For a small effect size (Cohen's  $d=0.2$ ), the power is 0.57. For all neuropsychological scores, tests for genotype-by-age also failed to give significant effects.

#### 3.3 Neuroimaging

##### 3.3.1 Cross-sectional analyses

**Table 2** shows baseline comparisons of neuroimaging variables. Global measures (total grey matter volume, total white matter volume, TIV) were not significantly different between

genotypes, even when stratified by future diagnosis. Whole-brain grey matter density comparisons conducted using voxel-based morphometry, and diffusion-weighted imaging analyses to assess white matter microstructure showed no significant genotype effects and are described in further detail in **Supplementary Notes**.

Cross-sectionally, there was a main effect of  $\epsilon 4$  carrier status on total hippocampal volume ( $F(1, 798)=10.93, p=0.001, \eta^2p=0.0135$ ) and right hippocampal volume ( $F(1, 798)=12.25, p=0.0005, \eta^2p=0.0151$ ), where *APOE*  $\epsilon 3/\epsilon 3$  subjects have larger volumes than  $\epsilon 4$  carriers, after regressing out the effects of TIV, age, sex, and education (**Table 2**). A similar, though smaller, effect was found in the left hippocampus ( $F(1, 798)=7.95, p=0.0049, \eta^2p=0.0099$ ). No sex-by-genotype interaction was found in any measure.

A significant genotype-by-age interaction was found for the right hippocampus ( $F(1, 797)=6.98, p=0.0084, \eta^2p=0.0087$ ), where  $\epsilon 3$  homozygotes showed a greater reduction in hippocampal volume with increasing age compared to  $\epsilon 4$  carriers (**Fig. 1a**). A genotype-by-age interaction was not observed for left hippocampal volume ( $F(1, 797)=1.74, p=0.187$ ; **Fig. 1b**) or TIV ( $F(1, 798)=0.95, p=0.33$ ; **Fig. 1c**). A one-way ANOVA, after splitting into 3 age groups (<75, 75 to <80, 80+ years), confirmed a significant difference in right hippocampal volume between age groups in  $\epsilon 3$  homozygotes, but not in  $\epsilon 4$  carriers ( $\epsilon 3/\epsilon 3: F(2, 654)=38.57, p<0.0001, \eta^2p=0.1055$ ;  $\epsilon 3/\epsilon 4: F(2, 144)=1.41, p=0.248$ ). A Tukey post-hoc test on  $\epsilon 3$  homozygotes revealed that mean right hippocampal volume was significantly smaller with each increasing age group (<75 vs. 75 to <80:  $t=-6.81, p=3.5 \times 10^{-10}$ ; <75 vs. 80+:  $t=-7.10, p=4.0 \times 10^{-10}$ ; 75 to <80 vs. 80+:  $t=-2.75, p=0.017$ ; **Fig. 1d**). Finally, right hippocampal volume was significantly larger in the  $\epsilon 3$  homozygotes than  $\epsilon 4$  carriers only in the youngest (69 to <75) age group ( $t=3.04, p=0.0025$ ; **Fig. 1d**).

To test whether this genotype-by-age interaction on the right hippocampus may be driven by prodromal effects, the analysis was conducted again using future conversion status. Of the 804 subjects, 702 remained cognitively normal (*APOE*  $\epsilon 3/\epsilon 3$ :  $n=588$  (69% female), *APOE*  $\epsilon 3/\epsilon 4$ :  $n=114$  (56% female)) and 102 converted to MCI or Alzheimer's disease (*APOE*  $\epsilon 3/\epsilon 3$ :  $n=69$  (61% female), *APOE*  $\epsilon 3/\epsilon 4$ :  $n=33$  (70% female)) over the study period of 9 years.

Using Fisher's exact test, the distribution of *APOE*  $\epsilon 4$  carriers between controls and converters was found to be significantly different (odds ratio (OR)=2.47, 95% confidence interval (CI)=1.56-3.90,  $p=0.0003$ ), such that the proportion of  $\epsilon 4$  carriers who eventually convert is double that of  $\epsilon 4$  noncarriers (22% of  $\epsilon 4$  carriers convert compared to 10% of  $\epsilon 3$  homozygotes). When split into age groups, it became apparent that  $\epsilon 4$  carrier status was only significantly associated with future conversion before the age of 80 (<75: OR=2.54, 95% CI=1.28-5.05,  $p=0.011$ ; 75 to <80: OR=3.48, 95% CI=1.72-7.03,  $p=0.0009$ ; 80+: OR=0.86, 95% CI=0.23-3.24,  $p=1.00$ ; **Fig. 2**). Critically, when future converters were removed from the right hippocampal volume analysis, there was no longer a main effect of genotype ( $p=0.181$ ), and the cross-sectional genotype-by-age interaction (shown in **Fig. 1a**) was no longer significant as well ( $F(1, 695)=3.37$ ,  $p=0.07$ ).

### 3.3.2 Longitudinal analyses

Of the 804-subject cohort described above, a subgroup met inclusion criteria for longitudinal analyses ( $n=490$ , 64% females). Like the 804-subject cross-sectional sample, in this subgroup: (1) at baseline, there were no significant differences after FDR correction for any demographic and neuropsychological measures between genotypes (**Supplementary Table 1**); (2) right and total hippocampal volumes were significantly smaller in  $\epsilon 4$  carriers compared to noncarriers at baseline, but these effects were no longer present after stratifying by future diagnosis (**Supplementary Table 2**); and (3) converters in the longitudinal

subgroup were more associated with being  $\epsilon 4$  carriers than  $\epsilon 3$  homozygotes (OR=3.12, 95% CI=1.71-5.68,  $p=0.0005$ ).

Cognitively, future converters performed worse over time than non-converters in all neuropsychological tests, but genotype effects were only observed in FCSRT delayed recall scores. *APOE*  $\epsilon 4$  carriers were associated with a greater decline in FCSRT delayed recall scores compared to  $\epsilon 3$  homozygotes over an average of 4.7 annual visits (estimate: -0.31,  $p=0.013$ ) (**Table 3**). When stratified by future diagnosis,  $\epsilon 4$  carriers showed a reduced learning effect in FCSRT delayed recall scores in non-converters only (estimate: -0.21,  $p=0.045$ ), though this did not survive correction for multiple comparisons.

For right hippocampal atrophy rates, significant differences were found for age-by-future diagnosis only (**Table 3**). When stratified by diagnosis, no significant genotype effects were found on right hippocampal atrophy, but when stratified by genotype, future diagnosis significantly contributed to differences in atrophy rate (**Fig. 3**). The same pattern was observed both when adjusting for TIV and for brain volume. No significant sex-by-genotype or sex-by-future diagnosis interactions were present in the overall sample, thus further interactions with sex were not tested for in the stratified samples.

#### 4. Discussion

To specifically test the prodromal hypothesis of *APOE* genotype effects on cognition and brain structure, we used future diagnostic status to conditionalise our baseline cross-sectional findings. In our cross-sectional analyses, the presence of one *APOE*  $\epsilon 4$  allele was not significantly associated with any cognitive or global brain structure changes in our cohort. However, we did observe a significant interaction between *APOE*  $\epsilon 4$  genotype and age in right hippocampal volume, where  $\epsilon 3$  homozygotes had larger hippocampal volumes than  $\epsilon 4$  carriers before, but not after, the age of 80. This was no longer apparent after taking into

account future diagnosis, suggesting that the smaller hippocampal volumes observed in  $\epsilon 4$  carriers under 80 were due to the larger proportion of future converters present in the younger age groups. Our study, therefore, highlights the risk of including prodromal MCI or AD subjects in cross-sectional analyses, which may artificially inflate genotype effects.

The cross-sectional results did not preclude the possibility that  $\epsilon 4$  carrier status could differentially affect the rate of hippocampal atrophy in cognitively normal subjects.

Comparing longitudinal hippocampal atrophy rates in subjects who remained cognitively normal throughout the study duration, versus subjects who converted to MCI or AD at some point during the study, revealed no effects of genotype on hippocampal atrophy rate. By contrast, future conversion significantly increased hippocampal atrophy rate in the overall sample and when stratified by genotype.

Some studies have suggested that carrying the  $\epsilon 4$  allele differentially affects males and females. Being an *APOE*  $\epsilon 4$  carrier has been associated with greater risk of developing AD in females than males (Farrer et al., 1997), with an earlier age of onset (Neu et al., 2017; Noguchi et al., 1993), and faster rate of cognitive decline in female MCI patients than males (Lin et al., 2015). In our study, we found no sex-by-genotype interactions in either the cross-sectional or longitudinal analyses. As being a female  $\epsilon 4$  carrier is highly linked to future diagnosis, and our study population consists of more females than males, it is difficult to determine if any potential sex-by-genotype effects have been obscured by overlapping prodromal effects. We did not attempt to look for an interaction between genotype, sex, and diagnosis as the statistical power would be too low for our sample size.

A limitation of our study is the lack of beta-amyloid ( $A\beta$ ) biomarkers, which has been associated with *APOE*  $\epsilon 4$  carrier status (Reiman et al., 2009; Rowe et al., 2007). In a study involving healthy controls aged  $\geq 70$  years, *APOE*  $\epsilon 4$  allele carriers were more likely to be

amyloid positive on PiB-PET imaging (49% in  $\epsilon 4$  carriers vs. 21% of noncarriers) (Rowe et al., 2010). Cognitively, only  $\epsilon 4$  carriers with subjective memory complaints were associated with increased PiB binding. In a longitudinal study of cognitive performance (Lim et al., 2018), *APOE*  $\epsilon 4$  carrier status was also only associated with a significant decline in memory performance in amyloid positive healthy elderly subjects. In studies of hippocampal volume, hippocampal atrophy rate was found to be affected by amyloid positivity (as determined by PiB-PET (Khan et al., 2017) or CSF  $A\beta_{1-42}$  (Schuff et al., 2009)) in  $\epsilon 4$ -carrying MCI patients. This suggests that the effects of the *APOE*  $\epsilon 4$  allele on memory function and hippocampal volume may manifest only in the presence of  $A\beta$  pathology, supporting our conclusion that the effect of *APOE*  $\epsilon 4$  on cognition and brain structure reflects a prodromal state.

We also found that the association between the  $\epsilon 4$  allele and future MCI conversion was only significant in subjects under the age of 80. The *APOE*  $\epsilon 4$  allele is associated with greater risk of developing AD as well as an earlier age of onset (Jack et al., 1999), with peak progression risk between ages 70-75 (Bonham et al., 2016). As not all  $\epsilon 4$  carriers under age 75 convert to MCI or AD, it is possible that there are additional risk factors or lack of protective factors which lead to hippocampal volume reduction and associated conversion. Likewise, in  $\epsilon 4$  carriers who are able to remain cognitively normal over the age of 80, they may be benefitting from a protective factor which attenuates the effects of the *APOE*  $\epsilon 4$  allele on hippocampal volume. As the rates of hippocampal atrophy in our  $\epsilon 3$  homozygotes are the same as the  $\epsilon 3/\epsilon 4$  heterozygotes, it could be inferred that they share a common underlying pathological process affected by the same risk or protective factors; it is the risk of expressing these factors that is modulated by the  $\epsilon 4$  allele.

**Author contributions:**

Conceptualisation: LZ, MM, and BAS.; Data Acquisition: MC; Data Analysis: LZ and BAS;

Writing and Revision: LZ and BAS; Supervision and Funding: MC, MM, and BAS.

**Acknowledgements:**

We thank the participants of the Vallecas Project and the staff of the CIEN Foundation.

**Funding:**

This work was supported by the CIEN Foundation and the Queen Sofia Foundation, grants from Carlos III Institute of Health and a grant from the Alzheimer's Association (2016-NIRG-397128) to BAS.

**Declarations of interest:** None.

## References

- Andersson, J.L., Jenkinson, M., Smith, S., 2007. Non-linear registration, aka Spatial normalisation FMRIB technical report TR07JA2. FMRIB Analysis Group of the University of Oxford 2, 1-21.
- Bonham, L.W., Geier, E.G., Fan, C.C., Leong, J.K., Besser, L., Kukull, W.A., Kornak, J., Andreassen, O.A., Schellenberg, G.D., Rosen, H.J., Dillon, W.P., Hess, C.P., Miller, B.L., Dale, A.M., Desikan, R.S., Yokoyama, J.S., 2016. Age-dependent effects of APOE epsilon4 in preclinical Alzheimer's disease. *Ann Clin Transl Neurol* 3(9), 668-677.
- Calero, O., Bullido, M.J., Clarimón, J., Frank-García, A., Martínez-Martín, P., Lleó, A., Rey, M.J., Rábano, A., Blesa, R., Gómez-Isla, T., 2011. Genetic cross-interaction between APOE and PRNP in sporadic Alzheimer's and Creutzfeldt-Jakob diseases. *PLoS One* 6(7), e22090.
- Calero, O., Hortigüela, R., Bullido, M.J., Calero, M., 2009. Apolipoprotein E genotyping method by Real Time PCR, a fast and cost-effective alternative to the TaqMan® and FRET assays. *Journal of Neuroscience Methods* 183(2), 238-240.
- Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G., Roses, A.D., Haines, J., Pericak-Vance, M.A., 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261(5123), 921-923.
- den Heijer, T., Oudkerk, M., Launer, L.J., van Duijn, C.M., Hofman, A., Breteler, M.M.B., 2002. Hippocampal, amygdalar, and global brain atrophy in different apolipoprotein E genotypes. *Neurology* 59(5), 746.
- Farrer, L.A., Cupples, L.A., Haines, J.L., Hyman, B., Kukull, W.A., Mayeux, R., Myers, R.H., Pericak-Vance, M.A., Risch, N., Van Duijn, C.M., 1997. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: a meta-analysis. *JAMA* 278(16), 1349-1356.
- Fazekas, F., Chawluk, J.B., Alavi, A., Hurtig, H.I., Zimmerman, R.A., 1987. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *Am. J. Roentgenol.* 149(2), 351-356.
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. *Neuron* 33(3), 341-355.
- Foster, J.K., Albrecht, M.A., Savage, G., Lautenschlager, N.T., Ellis, K.A., Maruff, P., Szoek, C., Taddei, K., Martins, R., Masters, C.L., Ames, D., Group, A.R., 2013. Lack of reliable evidence for a distinctive epsilon4-related cognitive phenotype that is independent from clinical diagnostic status: findings from the Australian Imaging, Biomarkers and Lifestyle Study. *Brain* 136(Pt 7), 2201-2216.
- Good, C., Johnsrude, I., Ashburner, J., Henson, R., Friston, K., Frackowiak, R., 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage* 14(1 Pt 1), 21-36.
- Greenwood, P.M., Parasuraman, R., 2003. Normal genetic variation, cognition, and aging. *Behav Cogn Neurosci Rev* 2(4), 278-306.
- Honea, R.A., Vidoni, E., Harsha, A., Burns, J.M., 2009. Impact of APOE on the healthy aging brain: a voxel-based MRI and DTI study. *J Alzheimers Dis* 18(3), 553-564.
- Iglesias, J.E., Augustinack, J.C., Nguyen, K., Player, C.M., Player, A., Wright, M., Roy, N., Frosch, M.P., McKee, A.C., Wald, L.L., Fischl, B., Van Leemput, K., Alzheimer's Disease Neuroimaging, I., 2015. A computational atlas of the hippocampal formation

- using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *Neuroimage* 115, 117-137.
- Iglesias, J.E., Van Leemput, K., Augustinack, J., Insausti, R., Fischl, B., Reuter, M., Alzheimer's Disease Neuroimaging, I., 2016. Bayesian longitudinal segmentation of hippocampal substructures in brain MRI using subject-specific atlases. *Neuroimage* 141, 542-555.
- Jack, C.R., Jr., Knopman, D.S., Jagust, W.J., Petersen, R.C., Weiner, M.W., Aisen, P.S., Shaw, L.M., Vemuri, P., Wiste, H.J., Weigand, S.D., Lesnick, T.G., Pankratz, V.S., Donohue, M.C., Trojanowski, J.Q., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol* 12(2), 207-216.
- Jack, C.R., Jr., Petersen, R.C., Xu, Y.C., O'Brien, P.C., Smith, G.E., Ivnik, R.J., Boeve, B.F., Waring, S.C., Tangalos, E.G., Kokmen, E., 1999. Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment. *Neurology* 52(7), 1397-1403.
- Khan, W., Giampietro, V., Banaschewski, T., Barker, G.J., Bokde, A.L., Buchel, C., Conrod, P., Flor, H., Frouin, V., Garavan, H., Gowland, P., Heinz, A., Ittermann, B., Lemaitre, H., Nees, F., Paus, T., Pausova, Z., Rietschel, M., Smolka, M.N., Strohle, A., Gallinat, J., Vellas, B., Soininen, H., Kloszewska, I., Tsolaki, M., Mecocci, P., Spenger, C., Villemagne, V.L., Masters, C.L., Muehlboeck, J.S., Backman, L., Fratiglioni, L., Kalpouzos, G., Wahlund, L.O., Schumann, G., Lovestone, S., Williams, S.C., Westman, E., Simmons, A., Alzheimer's Disease Neuroimaging, I., AddNeuroMed Consortium, A.I.B., Lifestyle Study Research, G., consortium, I., 2017. A Multi-Cohort Study of ApoE varepsilon4 and Amyloid-beta Effects on the Hippocampus in Alzheimer's Disease. *J Alzheimers Dis* 56(3), 1159-1174.
- Kim, Y.J., Seo, S.W., Park, S.B., Yang, J.J., Lee, J.S., Lee, J., Jang, Y.K., Kim, S.T., Lee, K.-H., Lee, J.M., Lee, J.-H., Kim, J.S., Na, D.L., Kim, H.J., 2017. Protective effects of APOE e2 against disease progression in subcortical vascular mild cognitive impairment patients: A three-year longitudinal study. *Scientific Reports* 7(1).
- Knight, R.G., Tsui, H.S., Abraham, W.C., Skeaff, C.M., McMahon, J.A., Cutfield, N.J., 2014. Lack of effect of the apolipoprotein E epsilon4 genotype on cognition during healthy aging. *J Clin Exp Neuropsychol* 36(7), 742-750.
- Lim, Y.Y., Kalinowski, P., Pietrzak, R.H., Laws, S.M., Burnham, S.C., Ames, D., Villemagne, V.L., Fowler, C.J., Rainey-Smith, S.R., Martins, R.N., Rowe, C.C., Masters, C.L., Maruff, P.T., 2018. Association of beta-Amyloid and Apolipoprotein E epsilon4 With Memory Decline in Preclinical Alzheimer Disease. *JAMA Neurol* 75(4), 488-494.
- Lin, K.A., Choudhury, K.R., Rathakrishnan, B.G., Marks, D.M., Petrella, J.R., Doraiswamy, P.M., Alzheimer's Disease Neuroimaging, I., 2015. Marked gender differences in progression of mild cognitive impairment over 8 years. *Alzheimers Dement (N Y)* 1(2), 103-110.
- Liu, C.C., Liu, C.C., Kanekiyo, T., Xu, H., Bu, G., 2013. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol* 9(2), 106-118.
- Lyall, D.M., Ward, J., Ritchie, S.J., Davies, G., Cullen, B., Celis, C., Bailey, M.E.S., Anderson, J., Evans, J., McKay, D.F., McIntosh, A.M., Sattar, N., Smith, D.J., Deary, I.J., Pell, J.P., 2016. Alzheimer disease genetic risk factor APOE e4 and cognitive abilities in 111,739 UK Biobank participants. *Age and Ageing* 45(4), 511-517.
- Manning, E.N., Barnes, J., Cash, D.M., Bartlett, J.W., Leung, K.K., Ourselin, S., Fox, N.C., Alzheimer's Disease Neuroimaging, I., 2014. APOE epsilon4 is associated with disproportionate progressive hippocampal atrophy in AD. *PLoS One* 9(5), e97608.

- Mattsson, N., Groot, C., Jansen, W.J., Landau, S.M., Villemagne, V.L., Engelborghs, S., Mintun, M.M., Lleo, A., Molinuevo, J.L., Jagust, W.J., Frisoni, G.B., Ivanoiu, A., Chetelat, G., Resende de Oliveira, C., Rodrigue, K.M., Kornhuber, J., Wallin, A., Klimkowicz-Mrowiec, A., Kandimalla, R., Popp, J., Aalten, P.P., Aarsland, D., Alcolea, D., Almdahl, I.S., Baldeiras, I., van Buchem, M.A., Cavado, E., Chen, K., Cohen, A.D., Forster, S., Fortea, J., Frederiksen, K.S., Freund-Levi, Y., Gill, K.D., Gkatzima, O., Grimmer, T., Hampel, H., Herukka, S.K., Johannsen, P., van Laere, K., de Leon, M.J., Maier, W., Marcusson, J., Meulenbroek, O., Mollergard, H.M., Morris, J.C., Mroczko, B., Nordlund, A., Prabhakar, S., Peters, O., Rami, L., Rodriguez-Rodriguez, E., Roe, C.M., Ruther, E., Santana, I., Schroder, J., Seo, S.W., Soininen, H., Spuru, L., Stomrud, E., Struyfs, H., Teunissen, C.E., Verhey, F.R.J., Vos, S.J.B., van Waalwijk van Doorn, L.J.C., Waldemar, G., Wallin, A.K., Wiltfang, J., Vandenberghe, R., Brooks, D.J., Fladby, T., Rowe, C.C., Drzezga, A., Verbeek, M.M., Sarazin, M., Wolk, D.A., Fleisher, A.S., Klunk, W.E., Na, D.L., Sanchez-Juan, P., Lee, D.Y., Nordberg, A., Tsolaki, M., Camus, V., Rinne, J.O., Fagan, A.M., Zetterberg, H., Blennow, K., Rabinovici, G.D., Hansson, O., van Berckel, B.N.M., van der Flier, W.M., Scheltens, P., Visser, P.J., Ossenkoppele, R., 2018. Prevalence of the apolipoprotein E epsilon4 allele in amyloid beta positive subjects across the spectrum of Alzheimer's disease. *Alzheimers Dement* 14(7), 913-924.
- Mayeux, R., Stern, Y., Ottman, R., Tatemichi, T.K., Tang, M.X., Maestre, G., Ngai, C., Tycko, B., Ginsberg, H., 1993. The apolipoprotein epsilon 4 allele in patients with Alzheimer's disease. *Ann Neurol* 34(5), 752-754.
- Nao, J., Sun, H., Wang, Q., Ma, S., Zhang, S., Dong, X., Ma, Y., Wang, X., Zheng, D., 2017. Adverse Effects of the Apolipoprotein E epsilon4 Allele on Episodic Memory, Task Switching and Gray Matter Volume in Healthy Young Adults. *Front Hum Neurosci* 11, 346.
- Neu, S.C., Pa, J., Kukull, W., Beekly, D., Kuzma, A., Gangadharan, P., Wang, L.S., Romero, K., Arneric, S.P., Redolfi, A., Orlandi, D., Frisoni, G.B., Au, R., Devine, S., Auerbach, S., Espinosa, A., Boada, M., Ruiz, A., Johnson, S.C., Kosciak, R., Wang, J.J., Hsu, W.C., Chen, Y.L., Toga, A.W., 2017. Apolipoprotein E Genotype and Sex Risk Factors for Alzheimer Disease: A Meta-analysis. *JAMA Neurol* 74(10), 1178-1189.
- Noguchi, S., Murakami, K., Yamada, N., 1993. Apolipoprotein E genotype and Alzheimer's disease. *Lancet* 342(8873), 737.
- O'Donoghue, M.C., Murphy, S.E., Zamboni, G., Nobre, A.C., Mackay, C.E., 2018. APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: A review. *Cortex* 104, 103-123.
- Olazaran, J., Valenti, M., Frades, B., Zea-Sevilla, M.A., Avila-Villanueva, M., Fernandez-Blazquez, M.A., Calero, M., Dobato, J.L., Hernandez-Tamames, J.A., Leon-Salas, B., Aguera-Ortiz, L., Lopez-Alvarez, J., Larranaga, P., Bielza, C., Alvarez-Linera, J., Martinez-Martin, P., 2015. The Vallecas Project: A Cohort to Identify Early Markers and Mechanisms of Alzheimer's Disease. *Front Aging Neurosci* 7, 181.
- Rajagopalan, V., Piore, E.P., 2015. Disparate voxel based morphometry (VBM) results between SPM and FSL softwares in ALS patients with frontotemporal dementia: which VBM results to consider? *BMC Neurology* 15, 32.
- Reiman, E.M., Chen, K., Liu, X., Bandy, D., Yu, M., Lee, W., Ayutyanont, N., Keppler, J., Reeder, S.A., Langbaum, J.B., Alexander, G.E., Klunk, W.E., Mathis, C.A., Price, J.C., Aizenstein, H.J., DeKosky, S.T., Caselli, R.J., 2009. Fibrillar amyloid-beta burden in cognitively normal people at 3 levels of genetic risk for Alzheimer's disease. *Proc Natl Acad Sci U S A* 106(16), 6820-6825.

- Reuter, M., Schmansky, N.J., Rosas, H.D., Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61(4), 1402-1418.
- Rowe, C.C., Ellis, K.A., Rimajova, M., Bourgeat, P., Pike, K.E., Jones, G., Fripp, J., Tochon-Danguy, H., Morandau, L., O'Keefe, G., Price, R., Raniga, P., Robins, P., Acosta, O., Lenzo, N., Szoek, C., Salvado, O., Head, R., Martins, R., Masters, C.L., Ames, D., Villemagne, V.L., 2010. Amyloid imaging results from the Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging. *Neurobiol Aging* 31(8), 1275-1283.
- Rowe, C.C., Ng, S., Ackermann, U., Gong, S.J., Pike, K., Savage, G., Cowie, T.F., Dickinson, K.L., Maruff, P., Darby, D., Smith, C., Woodward, M., Merory, J., Tochon-Danguy, H., O'Keefe, G., Klunk, W.E., Mathis, C.A., Price, J.C., Masters, C.L., Villemagne, V.L., 2007. Imaging beta-amyloid burden in aging and dementia. *Neurology* 68(20), 1718-1725.
- Salvato, G., Patai, E.Z., McCloud, T., Nobre, A.C., 2016. Apolipoprotein varepsilon4 breaks the association between declarative long-term memory and memory-based orienting of spatial attention in middle-aged individuals. *Cortex* 82, 206-216.
- Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L.M., Trojanowski, J.Q., Thompson, P.M., Jack, C.R., Jr., Weiner, M.W., Alzheimer's Disease Neuroimaging, I., 2009. MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain* 132(Pt 4), 1067-1077.
- Shi, J., Lepore, N., Gutman, B.A., Thompson, P.M., Baxter, L.C., Caselli, R.J., Wang, Y., Alzheimer's Disease Neuroimaging, I., 2014. Genetic influence of apolipoprotein E4 genotype on hippocampal morphometry: An N = 725 surface-based Alzheimer's disease neuroimaging initiative study. *Hum Brain Mapp* 35(8), 3903-3918.
- Smith, G.E., Bohac, D.L., Waring, S.C., Kokmen, E., Tangalos, E.G., Ivnik, R.J., Petersen, R.C., 1998. Apolipoprotein E genotype influences cognitive 'phenotype' in patients with Alzheimer's disease but not in healthy control subjects. *Neurology* 50(2), 355-362.
- Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage* 31(4), 1487-1505.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 23, S208-S219.
- Trachtenberg, A.J., Filippini, N., Mackay, C.E., 2012. The effects of APOE-epsilon4 on the BOLD response. *Neurobiol Aging* 33(2), 323-334.
- Tsai, M.S., Tangalos, E.G., Petersen, R.C., Smith, G.E., Schaid, D.J., Kokmen, E., Ivnik, R.J., Thibodeau, S.N., 1994. Apolipoprotein E: risk factor for Alzheimer disease. *Am J Hum Genet* 54(4), 643-649.
- Wisdom, N.M., Callahan, J.L., Hawkins, K.A., 2011. The effects of apolipoprotein E on non-impaired cognitive functioning: a meta-analysis. *Neurobiol Aging* 32(1), 63-74.
- Wishart, H.A., Saykin, A.J., McAllister, T.W., Rabin, L.A., McDonald, B.C., Flashman, L.A., Roth, R.M., Mamourian, A.C., Tsongalis, G.J., Rhodes, C.H., 2006. Regional brain atrophy in cognitively intact adults with a single APOE ε4 allele. *Neurology* 67(7), 1221.

**Table 1. Demographic and neuropsychological variables according to *APOE*  $\epsilon$  genotype for subjects included in cross-sectional analyses.**

	<i>APOE</i> $\epsilon$ 3/ $\epsilon$ 3 (657) Mean (stdev), range	<i>APOE</i> $\epsilon$ 3/ $\epsilon$ 4 (147) Mean (stdev), range	<b>Difference <i>APOE</i> <math>\epsilon</math>3/<math>\epsilon</math>3 vs. <math>\epsilon</math>3/<math>\epsilon</math>4</b> One-way ANOVA, Welch's <i>t</i> -test, or $\chi^2$ results, df in parentheses	<b>Genotype-by-age interaction</b> df in parentheses
Age (years)	74.80 (3.88) 68.9-86.9	74.47 (3.77) 69.2-85.7	$t_{(221)}=0.95$ $p=0.341$	n/a
Sex	209 M 448 F	60 M 87 F	$\chi^2_{(1)}=4.38$ $p=0.036^*$	n/a
Years of education	10.41 (5.83) 0-24	10.72 (5.94) 0-24	$t_{(214)}=-0.57$ $p=0.571$	n/a
MMSE (/30)	28.70 (1.48) 21-30	28.67 (1.54) 23-30	$F_{(1, 799)}=0.24$ $p=0.627$	$F_{(1, 798)}=1.62$ $p=0.204$
FAQ (/30)	0.37 (0.73) 0-5	0.46 (0.72) 0-3	$F_{(1, 799)}=1.18$ $p=0.277$	$F_{(1, 798)}=1.56$ $p=0.212$
FCSRT (/16)	14.41 (1.76) 7-16	14.29 (2.06) 5-16	$F_{(1, 798)}=0.94$ $p=0.331$	$F_{(1, 797)}=2.71$ $p=0.100$
Rey-Osterrieth Complex Figure (/36)	12.39 (5.97) 0-31	12.52 (6.81) 0.5-35	$F_{(1, 798)}=0.15$ $p=0.695$	$F_{(1, 797)}=0.00$ $p=0.995$
Phonological Verbal Fluency	13.49 (4.45) 3-27	13.87 (4.55) 4-25	$F_{(1, 799)}=0.56$ $p=0.454$	$F_{(1, 798)}=0.19$ $p=0.661$
Semantic Verbal Fluency	18.49 (4.85) 6-35	18.99 (4.85) 9-34	$F_{(1, 799)}=0.75$ $p=0.388$	$F_{(1, 798)}=0.48$ $p=0.488$
GDS (/15)	1.55 (2.19) 0-12	1.50 (2.26) 0-14	$t_{(213)}=0.27$ $p=0.791$	$F_{(1, 800)}=0.34$ $p=0.563$
STAI state (/60)	14.23 (9.03) 0-55	14.97 (9.46) 0-50	$t_{(210)}=-0.86$ $p=0.389$	$F_{(1, 799)}=1.38$ $p=0.240$
STAI trait (/60)	16.96 (9.81) 0-49	16.98 (9.29) 0-45	$t_{(226)}=-0.028$ $p=0.978$	$F_{(1, 799)}=1.00$ $p=0.319$

Comparisons between neuropsychological test scores have been adjusted for age, sex, and years of education. Abbreviations: FAQ: Functional Activities Questionnaire, FCSRT: Free and Cued Selective Reminding Test, GDS: Geriatric Depression Scale, M/F: Males/Females,

MMSE: Mini Mental State Examination, STAI: State-Trait Anxiety Inventory, stdev: standard deviation, df: degrees of freedom. \* indicates significance at  $p < 0.05$ . No comparison was significant at  $p < 0.05$  after FDR estimation.

**Table 2. Neuroimaging variables according to *APOE*  $\epsilon$  genotype for subjects included in the cross-sectional analyses.**

	<i>APOE</i> $\epsilon 3/\epsilon 3$ (657) Mean (stdev), range	<i>APOE</i> $\epsilon 3/\epsilon 4$ (147) Mean (stdev), range	<b>Difference <i>APOE</i> <math>\epsilon 3/\epsilon 3</math> vs. <math>\epsilon 3/\epsilon 4</math></b> One-way ANOVA or <i>t</i> -test results, df in parentheses	<b>Genotype-by-age interaction</b> df in parentheses
Hippocampal Volume (cm <sup>3</sup> )	5.57 (0.61) 4.07-7.83	5.45 (0.64) 3.74-7.16	F <sub>(1, 798)</sub> =10.93 <i>p</i> =0.0010*†	F <sub>(1, 797)</sub> =4.34 <i>p</i> =0.038*
Right Hippocampal Volume (cm <sup>3</sup> )	2.84 (0.32) 2.06-4.03	2.77 (0.34) 1.75-3.61	F <sub>(1, 798)</sub> =12.25 <i>p</i> =0.0005*†	F <sub>(1, 797)</sub> =6.98 <i>p</i> =0.0084*†
Left Hippocampal Volume (cm <sup>3</sup> )	2.73 (0.31) 1.89-3.84	2.69 (0.32) 1.75-3.55	F <sub>(1, 798)</sub> =7.95 <i>p</i> =0.0049*†	F <sub>(1, 797)</sub> =1.74 <i>p</i> =0.187
Fazekas score for white matter lesions (/3)	1.15 (0.79) 0-3	1.07 (0.92) 0-3	F <sub>(1, 801)</sub> =1.04 <i>p</i> =0.309	F <sub>(1, 800)</sub> =1.96 <i>p</i> =0.162
Total Grey Matter Volume (cm <sup>3</sup> )	577.67 (44.55) 468.37-714.70	581.08 (48.36) 468.12-727.99	F <sub>(1, 798)</sub> =0.08 <i>p</i> =0.779	F <sub>(1, 797)</sub> =2.87 <i>p</i> =0.091
Total White Matter Volume (cm <sup>3</sup> )	388.80 (49.10) 268.66-563.77	391.62 (34.17) 459.49-610.45	F <sub>(1, 798)</sub> =0.16 <i>p</i> =0.686	F <sub>(1, 797)</sub> =3.67 <i>p</i> =0.056
Total Intracranial Volume (cm <sup>3</sup> )	1948.65 (81.98) 1710.51-2232.95	1939.20 (67.10) 1715.11-2185.24	F <sub>(1, 799)</sub> =2.06 <i>p</i> =0.152	F <sub>(1, 798)</sub> =0.95 <i>p</i> =0.329

Comparisons between brain structural measures (hippocampal volumes, total grey and white matter volumes) have been adjusted for total intracranial volume, age, sex, and years of education. Abbreviations: stdev: standard deviation, df: degrees of freedom. \* indicates significance at  $p < 0.05$ ; † indicates significance at  $p < 0.05$  after FDR estimation.

**Table 3. Estimates from linear mixed-effects models in longitudinal analyses.**

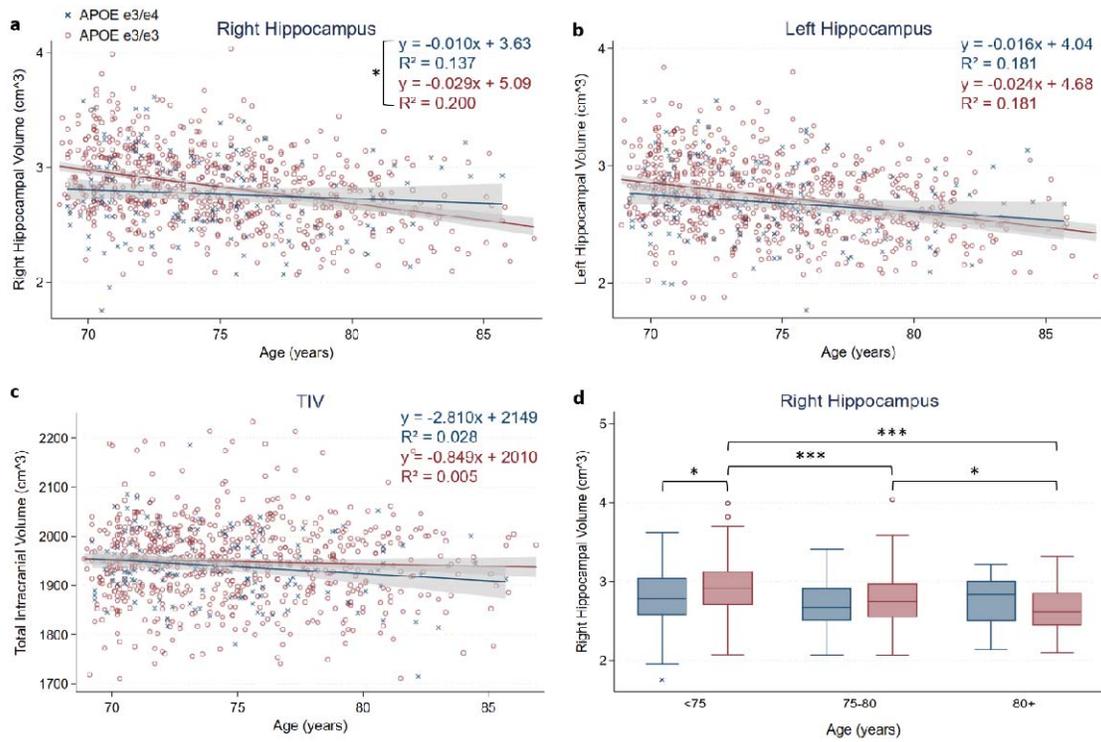
	FCSRT total delayed recall				Right hippocampal volume (adjusted by TIV)				Right hippocampal volume (adjusted by BV)			
	$\beta$	SE	95% CI	p	$\beta$	SE	95% CI	p	$\beta$	SE	95% CI	p
<i>All subjects (n=490)</i>												
APOE genotype ( $\epsilon 4=0$ or 1)	0.19	0.14	-0.08 – 0.46	0.163	-90.87	36.25	-161.91 – -19.83	0.012 *†	-88.21	30.80	-148.57 – -27.84	0.0042 *†
Future diagnosis (0=normal, 1=MCI/A D converter)	-2.55	0.17	-2.88 – -2.22	3.3 x 10 <sup>-52</sup> *†	-140.22	42.66	-223.82 – -56.61	0.001 *†	-145.75	36.37	-217.03 – -74.47	6.1 x 10 <sup>-5</sup> *†
Sex (0=male, 1=female)	0.15	0.11	-0.06 – 0.37	0.029 *	-175.61	28.51	-231.49 – -119.72	7.3 x 10 <sup>-10</sup> *†	22.26	25.70	-28.11 – 72.64	0.386
Age x $\epsilon 4$	-0.31	0.12	-0.55 – -0.07	0.013 *†	-20.78	10.66	-41.69 – 0.12	0.051	-15.75	9.73	-34.82 – 3.32	0.106
Age x future diagnosis	-1.33	0.13	-1.59 – -1.07	9.5 x 10 <sup>-24</sup> *†	-94.89	13.04	-120.46 – -69.33	3.5 x 10 <sup>-13</sup> *†	-71.02	12.15	-94.84 – -47.21	5.1 x 10 <sup>-9</sup> *†
Age x sex	-2.8 x 10 <sup>-4</sup>	0.09	-0.19 – 0.18	0.998	3.35	8.27	-12.86 – 19.56	0.685	-3.95	7.54	-18.73 – 10.83	0.601
Age x $\epsilon 4$ x future diagnosis	0.41	0.30	-0.18 – 0.99	0.171	-13.87	28.53	-69.78 – 42.04	0.627	-22.12	26.51	-74.08 – 29.83	0.404
Age x $\epsilon 4$ x sex	0.28	0.25	-0.22 – 0.77	0.272	-23.64	21.57	-65.92 – 18.65	0.273	-16.21	19.68	-54.77 – 22.36	0.410
Age x sex x future diagnosis	-0.27	0.27	-0.80 – 0.27	0.333	-20.45	26.72	-72.83 – 31.92	0.444	-4.03	24.83	-52.69 – 44.63	0.871
<i>Stratified by future diagnosis, converter=0 (n=433)</i>												
Age x $\epsilon 4$	-0.21	0.11	-0.42 – -0.005	0.045 *	-8.91	10.54	-29.57 – 11.74	0.397	-5.33	9.66	-24.27 – 13.60	0.581
<i>Stratified by future diagnosis, converter=1 (n=57)</i>												
Age x $\epsilon 4$	0.42	0.53	-0.62 – 1.46	0.420	-24.19	35.23	-93.24 – 44.86	0.492	-29.46	34.70	-97.48 – 38.55	0.396
<i>Stratified by genotype, <math>\epsilon 4=0</math> (n=406)</i>												
Age x future diagnosis	-1.46	0.16	-1.77 – -1.15	2.4 x 10 <sup>-20</sup> *†	-88.98	15.69	-119.74 – -58.24	1.4 x 10 <sup>-8</sup> *†	-62.92	14.60	-91.54 – -34.30	1.6 x 10 <sup>-5</sup> *†
<i>Stratified by genotype, <math>\epsilon 4=1</math> (n=84)</i>												
Age x future diagnosis	-0.92	0.25	-1.40 – -0.44	1.7 x 10 <sup>-4</sup> *†	-103.46	24.47	-151.42 – -55.50	2.4 x 10 <sup>-5</sup> *†	-86.04	22.93	-130.98 – -41.09	1.8 x 10 <sup>-4</sup> *†

Models were estimated for all subjects, as well as stratified by future diagnosis and genotype.

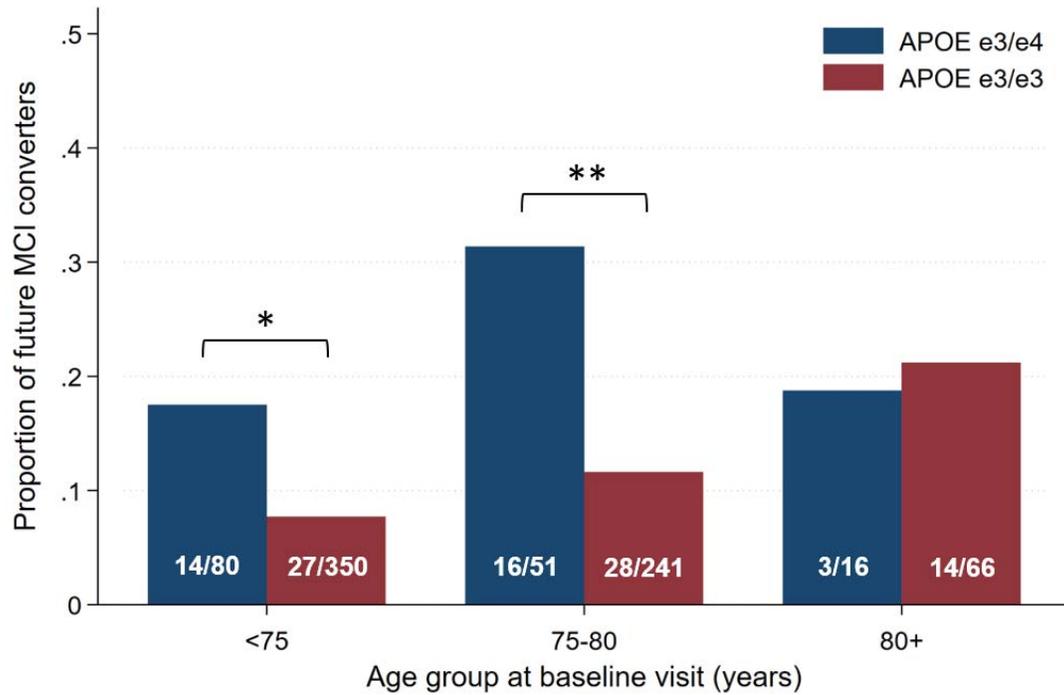
In addition to age, future diagnosis, genotype, and sex, FCSRT total delayed recall models

also included years of education as fixed covariates, and hippocampal volume models included years of education, and either total intracranial volume or brain volume as fixed covariates. Age, TIV, and BV were scaled. Interactions were tested in separate models. Abbreviations: SE: standard error, 95% CI: 95% confidence intervals. \* indicates significance at  $p < 0.05$ ; † indicates significance at  $p < 0.05$  after FDR estimation.

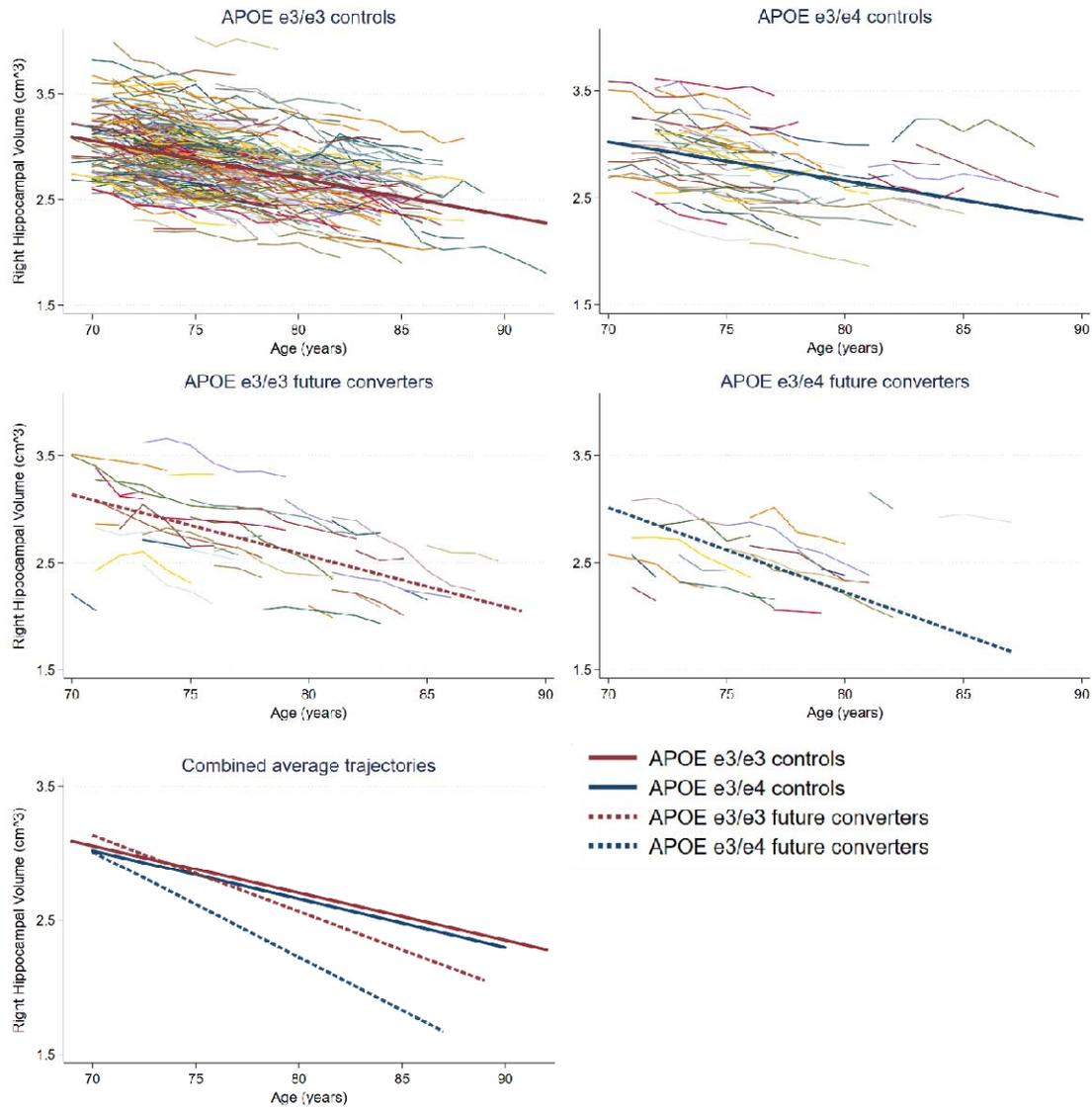
## Figures



**Fig. 1 - Cross-sectional analysis reveals an age-by-*APOE* genotype interaction for right hippocampal volume.** a) and d) Right hippocampal volume is reduced significantly more in *APOE* ε3/ε3 homozygotes than *APOE* ε3/ε4 carriers as age increases, but not b) left hippocampal volume or c) total intracranial volume. Lines are linearly fitted with 95% confidence intervals. For each boxplot in d), the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. Whiskers extend 1.5 times the interquartile range away from the top or bottom of the box, and outliers are plotted individually. \*  $p < 0.05$ , \*\*\*  $p < 0.0001$



**Fig. 2 - APOE ε4 carriers are more likely to be prodromal MCI than noncarriers before age 80.** Bars indicate proportion of future MCI converters in APOE ε4 carriers and ε3 homozygotes by baseline age group. Numbers within bars are actual numbers of future converters/subgroup *n*. \*  $p < 0.05$ , \*\*  $p < 0.001$



**Fig. 3 - APOE genotype has no effect on average rates of right hippocampal atrophy in either controls or converters.** The top 4 panels show spaghetti plots with average rates of right hippocampal atrophy, separated by diagnosis and genotype. Average slopes and intercepts (plotted separately in the top 4 panels, and together in the lower panel) were calculated as the average of each individual's linear estimate from regression with age and does not have confidence intervals.