nature medicine

Article

https://doi.org/10.1038/s41591-024-02931-w

# *APOE4* homozygozity represents a distinct genetic form of Alzheimer's disease

Received: 3 November 2023

Accepted: 19 March 2024

Published online: 06 May 2024

Juan Fortea[1,2,3,13], Jordi Pegueroles[1,2], Daniel Alcolea[1,2], Olivia Belbin[1,2], Oriol Dols-Icardo[1,2], Lídia Vaqué-Alcázar[1,4], Laura Videla[1,2,3], Juan Domingo Gispert[5,6,7,8,9], Marc Suárez-Calvet[5,6,7,8,9], Sterling C. Johnson[10], Reisa Sperling[11], Alexandre Bejanin[1,2], Alberto Lleó[1,2] & Víctor Montal[1,2,12,13]

This study aimed to evaluate the impact of *APOE4* homozygosity on Alzheimer's disease (AD) by examining its clinical, pathological and biomarker changes to see whether *APOE4* homozygotes constitute a distinct, genetically determined form of AD. Data from the National Alzheimer's Coordinating Center and five large cohorts with AD biomarkers were analyzed. The analysis included 3,297 individuals for the pathological study and 10,039 for the clinical study. Findings revealed that almost all *APOE4* homozygotes exhibited AD pathology and had significantly higher levels of AD biomarkers from age 55 compared to *APOE3* homozygotes. By age 65, nearly all had abnormal amyloid levels in cerebrospinal fluid, and 75% had positive amyloid scans, with the prevalence of these markers increasing with age, indicating near-full penetrance of AD biology in *APOE4* homozygotes. The age of symptom onset was earlier in *APOE4* homozygotes at 65.1, with a narrower 95% prediction interval than *APOE3* homozygotes. The predictability of symptom onset and the sequence of biomarker changes in *APOE4* homozygotes mirrored those in autosomal dominant AD and Down syndrome. However, in the dementia stage, there were no differences in amyloid or tau positron emission tomography across haplotypes, despite earlier clinical and biomarker changes. The study concludes that *APOE4* homozygotes represent a genetic form of AD, suggesting the need for individualized prevention strategies, clinical trials and treatments.

AD is a genetically complex disorder with both rare and common genetic variants involved in its pathogenesis[1,2]. Mutations in three genes, *APP*, *PSEN1* and *PSEN2*, cause early-onset autosomal dominant Alzheimer's disease (ADAD)[3], whereas variants in dozens of other genes have been associated with an increased risk of developing the more common sporadic (late-onset) form of the disease[1]. Among these genes, *APOE* is considered the strongest genetic risk factor. The three main characteristics of genetically determined AD with respect to sporadic AD are the near-full penetrance of the disease, the predictability of the age at symptom onset and a predictable sequence of pathological, biomarker and clinical changes.

*APOE4* homozygotes have a lifetime risk for AD dementia that can reach 60% at age 85, markedly increased compared to heterozygotes or noncarriers[4]. The recognition of this very high lifetime risk is much higher than the low-risk common alleles identified by genome-wide association studies in AD[1], and comparable to that found in Mendelian diseases[4].

A full list of affiliations appears at the end of the paper. e-mail: jfortea@santpau.cat; victor.montal@protonmail.com

**Table 1 | Demographic, clinical and biomarker data from the multisite clinical cohort**

| | Overall | *APOE4/4* | *APOE3/4* | *APOE3/3* | *APOE2/X* |
|---|---|---|---|---|---|
| NACC | | | | | |
| *n* sample | 3,297 | 273 | 1,088 | 1,565 | 371 |
| **Sex** | | | | | |
| Female | 1,553 (47%) | 130 (48%) | 511 (47%) | 723 (46%) | 189 (51%) |
| Male | 1,744 (53%) | 143 (52%) | 577 (53%) | 842 (54%) | 182 (49%) |
| **ADNC** | | | | | |
| High ADNC | 1,590 (48%) | 225 (82%) | 696 (64%) | 551 (35%) | 118 (32%) |
| Intermediate ADNC | 675 (20%) | 35 (13%) | 225 (21%) | 347 (22%) | 68 (18%) |
| Low ADNC | 598 (18%) | 10 (3.7%) | 137 (13%) | 360 (23%) | 91 (25%) |
| Not AD | 434 (13%) | 3 (1.1%) | 30 (2.8%) | 307 (20%) | 94 (25%) |
| **Alzheimer's dementia** | 2,099 (64%) | 240 (88%) | 820 (75%) | 840 (54%) | 199 (54%) |
| **Age at symptom onset** | 71 (11.2) | 65 (8.2) | 70 (6.5) | 74 (6.8) | 74 (8.6) |
| **Age at MCI** | 79 (9.1) | 72 (10) | 77 (8.4) | 82 (8.2) | 82 (9.7) |
| **Age at dementia onset** | 81 (9.4) | 74 (12.2) | 79 (9.1) | 83 (9.9) | 84 (11.4) |
| **Age at death** | 83 (10.5) | 80 (6.7) | 85 (7.8) | 89 (8.8) | 88 (7.3) |
| **Clinical cohorts** | | | | | |
| ***n* sample** | 10,036 | 519 | 3,142 | 5,139 | 1,236 |
| **Age** | 71 (7) | 69 (7) | 71 (7) | 71 (7) | 72 (7) |
| **Sex** | | | | | |
| Female | 5,666 (56%) | 285 (55%) | 1,762 (56%) | 2,946 (57%) | 673 (54%) |
| Male | 4,370 (44%) | 234 (45%) | 1,380 (44%) | 2,193 (43%) | 563 (46%) |
| **DX** | | | | | |
| HC | 8,218 (83%) | 289 (57%) | 2,418 (78%) | 4,426 (87%) | 1,085 (90%) |
| MCI | 1,045 (11%) | 113 (22%) | 384 (12%) | 464 (9·1%) | 84 (7.0%) |
| AD | 618 (6.3%) | 105 (21%) | 290 (9.4%) | 184 (3.6%) | 39 (3.2%) |
| **CSF A$\beta_{1-42}$ (pg per ml)** | 1,024 (538) (*n*=1,966) | 631 (319) (*n*=190) | 898 (445) (*n*=730) | 1,178 (571) (*n*=871) | 1,208 (560) (*n*=175) |
| **CSF pTau181 (pg per ml)** | 24 (14) (*n*=2,115) | 31 (17) (*n*=189) | 27 (14) (*n*=749) | 22 (11) (*n*=977) | 21 (12) (*n*=200) |
| **Hippo volume** | 0.0047 (0.0008) (*n*=5,253) | 0.0045 (0.0008) (*n*=410) | 0.0046 (0.0008) (*n*=1,970) | 0.0048 (0.0008) (*n*=2,306) | 0.0049 (0.0008) (*n*=565) |
| **Centiloid** | 21 (37) (*n*=7,562) | 56 (41) (*n*=364) | 35 (42) (*n*=2,345) | 12 (31) (*n*=3,896) | 9 (28) (*n*=955) |
| **Plasma pTau181 (pg per ml)** | 16 (16) (*n*=1,278) | 20 (11) (*n*=113) | 16 (10) (*n*=475) | 15 (22) (*n*=563) | 14 (9) (*n*=127) |
| **Plasma NfL (pg per ml)** | 32 (23) (*n*=2,086) | 34 (19) (*n*=182) | 31 (22) (*n*=762) | 32 (24) (*n*=940) | 32 (26) (*n*=202) |
| **Tau-PET (SUVr) Braak 1/2** | 1.19 (0.16) (*n*=1,289) | 1.33 (0.20) (*n*=86) | 1.21 (0.17) (*n*=462) | 1.16 (0.14) (*n*=600) | 1.14 (0.13) (*n*=139) |
| **Tau-PET (SUVr) Braak 3/4** | 1.20 (0.15) (*n*=1,289) | 1.33 (0.21) (*n*=86) | 1.21 (0.16) (*n*=462) | 1.18 (0.14) (*n*=600) | 1.16 (0.11) (*n*=139) |
| **Tau-PET (SUVr) Braak 5/6** | 1.14 (0.15) (*n*=1,289) | 1.23 (0.23) (*n*=86) | 1.14 (0.15) (*n*=462) | 1.12 (0.14) (*n*=600) | 1.11 (0.12) (*n*=139) |

DX, diagnosis; HC, healthy controls; Hippo, hippocampal. Data are shown as: mean (standard deviation).

The predictability of the age at symptom onset in genetically determined AD has facilitated clinical trials and is the cornerstone of counseling mutation carriers and their families. However, no previous study has assessed the predictability of symptom onset in *APOE4* homozygotes, and, consequently, the statistical approaches commonly used in ADAD, including the concept of estimated years to symptom onset (the predicted time until an individual with a disease-causing mutation starts showing AD, based on family history), have not been used.

The predictable sequence of pathological, biomarker and clinical changes in both ADAD and Down syndrome has provided unique insights into the pathophysiology of AD[3,5,6]. Many biomarker studies have assessed the impact of *APOE* on the biomarker changes. However, mainly due to sample size limitations, the majority combine *APOE4* heterozygotes and homozygotes in one '*APOE4* carriers' category. Of note, the studies that analyze *APOE4* heterozygotes and homozygotes have found a gene dose response on AD biomarkers[7,8]. Nevertheless, no study has comprehensively analyzed the gene dose effect across the amyloid, tau, neurodegeneration framework (AT(N))[9] biomarker categories with age and estimated years to symptom onset in *APOE4* homozygotes.

Taking advantage of the large dataset from the National Alzheimer's Coordinating Center (NACC) for pathological data ($n > 3,200$) and collecting data from five large multicenter cohorts of subjects with AD biomarkers published to date ($n > 10,000$), we aimed to assess the clinical, pathological and biomarker changes in *APOE4* homozygotes to test the hypothesis that they can be considered as another form of genetically determined dementia[5]; in fact constituting one of the most frequent Mendelian diseases.

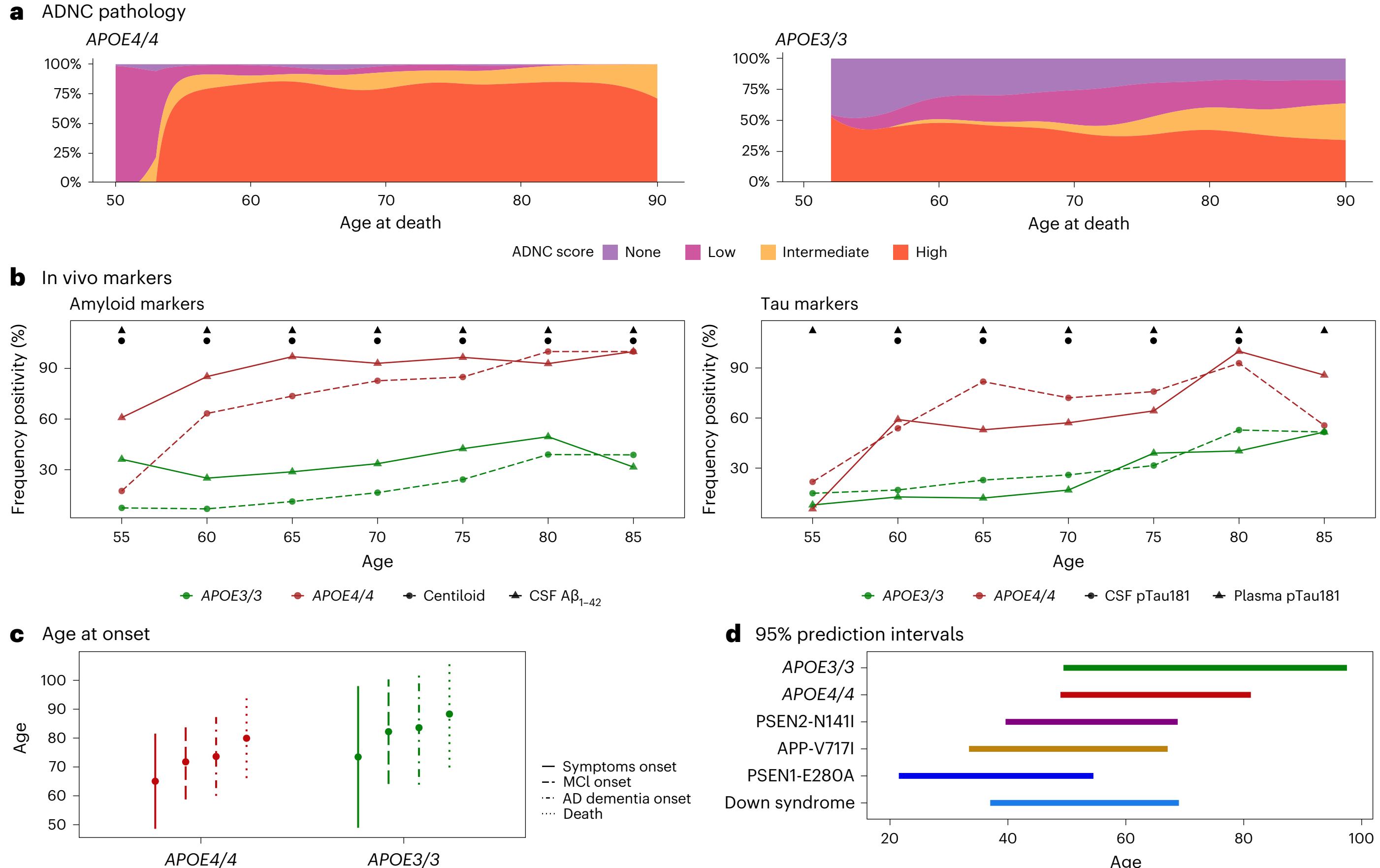


**Fig. 1 | Penetrance and predictive value of *APOE4* homozygosity in AD.** **a**, Distribution of ADNC scores by age of death, comparing *APOE4* homozygotes (*APOE4/4*) with *APOE3* homozygotes (*APOE3/3*). **b**, Frequency of positive amyloid and tau biomarkers across 5-year age intervals for both *APOE4* and *APOE3* homozygotes and the statistical significance of the difference. It demonstrate that *APOE4* homozygotes consistently exhibit higher levels of abnormal biomarkers; by age 65, nearly all subjects in *APOE4* homozygotes show abnormal levels of CSF amyloid-β. Black triangles and dots indicate significant differences between *APOE4* homozygotes and *APOE3* homozygotes in that age interval. **c**, Overview of the mean and 95% confidence interval ages at which symptoms, MCI and dementia manifest in *APOE4* homozygotes compared to *APOE3* homozygotes. It highlights that *APOE4* homozygotes experience significantly earlier onset ages and have narrower 95% confidence intervals for these milestones than *APOE3* homozygotes ($n$ = 240 *APOE4* homozygotes for symptom onset, $n$ = 55 for MCI, $n$ = 48 for AD dementia and $n$ = 48 for death; $n$ = 832 *APOE3* homozygotes for symptom onset, $n$ = 369 for MCI, $n$ = 265 for AD dementia and $n$ = 268 for death). **d**, The 95% prediction intervals for various genetically determined forms of AD. The panel highlights a similar variability (or predictability) for disease onset in *APOE4* homozygotes and both ADAD mutation carriers or individuals with Down syndrome, but a significantly wider prediction interval in the *APOE3* homozygotes compared to all other examined groups. (The data on ADAD and Down syndrome have been taken from ref. 13.)

## Results

We collected data from 3,297 brain donors from the NACC cohort[10] and 10,039 individuals from the clinical cohorts. Table 1 presents the demographic characteristics, *APOE* haplotype, neuropathological data and the biomarker data. There were no significant differences in sex distribution between haplotypes (both NACC and clinical cohorts). A summary of demographic, clinical and biomarker data for each clinical cohort data is available in Supplementary Table 1.

### Biological penetrance of AD in *APOE4* homozygotes

We analyzed the biological penetrance of AD in both postmortem data from NACC and in vivo biomarker results in the clinical cohorts. Concretely, for the postmortem data, we studied the profile of Alzheimer's disease neuropathological change (ADNC) scores (a measure of neuropathology load)[11] along the age span. Remarkably, nearly all *APOE4* homozygotes exhibited either high or intermediate ADNC scores, while this was the case for approximately 50% of *APOE3* homozygotes (Fig. 1a). Of note, the neuropathological findings in *APOE4* homozygotes were consistent regardless of their age at the time of death.

We then analyzed the biological penetrance using in vivo biomarkers from the clinical cohorts. Concretely, we binarized as positive or negative each participant's data for amyloid (Centiloid, cerebrospinal fluid (CSF) amyloid-β peptide 1–42 ($A\beta_{1-42}$) and tau (CSF, phosphorylated tau at residue 181 (pTau181)). The frequency of positive amyloid and tau biomarkers across 5-year age intervals in *APOE4* and *APOE3* homozygotes showed that *APOE4* homozygotes consistently exhibit higher levels of abnormal biomarkers than *APOE3* homozygotes starting at age 55. By age 65, nearly all *APOE4* homozygote participants show abnormal levels of CSF $A\beta_{1-42}$ and 75% had positive amyloid scans. The biological penetrance of AD increased with age for the other biomarkers, reaching 88% for all amyloid and tau biomarkers at age 80, despite the selection bias in this population toward cognitively unimpaired individuals (Fig. 1b). Of note, the biological penetrance profile was similar when splitting by sex (Supplementary Fig. 1).

Further details, illustrating the neuropathological variations based on *APOE* genotype, age and clinical diagnosis, can be found in Supplementary Fig. 2, which shows a clear *APOE* gene dose effect on AD neuropathology, as previously described[12].

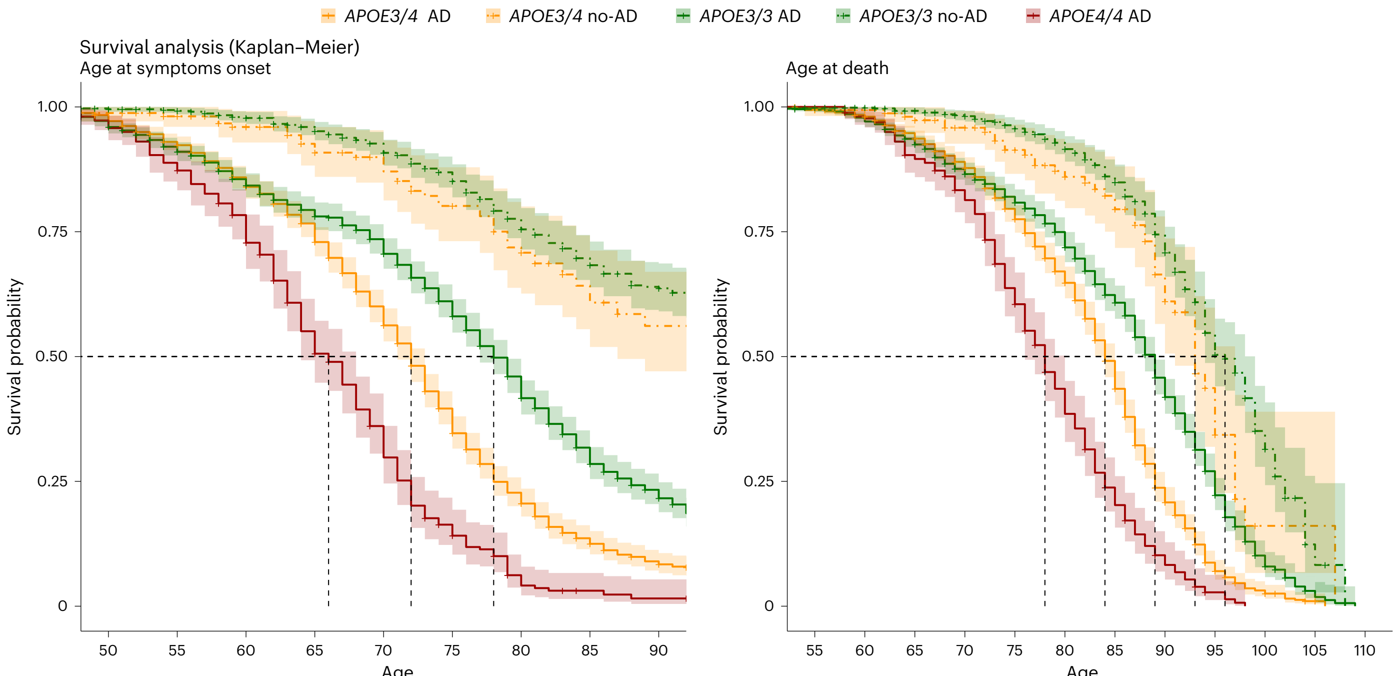


**Fig. 2 | Comparative survival analysis of clinical symptom onset and mortality by age, across different *APOE* haplotypes.** *APOE4* homozygotes are shown in red, *APOE4* heterozygotes in orange and *APOE3* homozygotes in green. Dashed lines indicate participants with low or no ADNC, whereas solid lines represent those with medium to high levels of AD pathology (ADNC medium or high). Confidence intervals are depicted by the shaded areas surrounding each line. The analysis contrasts the timing of clinical symptom onset (left) with age at death (right) as the event of interest.

### Predictability of symptomatology in *APOE4* homozygotes

Table 1 and Fig. 1c show the age at symptom onset, clinical diagnosis of mild cognitive impairment (MCI), dementia and death (±2 s.d.) in *APOE4* and *APOE3* homozygotes from the postmortem cohort. *APOE4* homozygotes started experiencing AD symptoms at age 65.6, MCI at 71.8, dementia at 73.6 and death at 77.2, approximately 7–10 years earlier than *APOE3* homozygotes ($P < 0.05$ differences). We also performed Kaplan–Meier survival analysis that confirmed the gene dosage effect on both the age at symptom onset and age at death, as illustrated in Fig. 2a,b.

We then studied the variability on the age at symptom onset of *APOE4* homozygotes compared to other genetically determined forms of AD. Consequently, we calculated the 95% prediction intervals of symptom onset (that is, the age range within which we expect symptom onset to start with 95% confidence). We found the same variability or predictability in *APOE4* homozygotes (32 years; Fig. 1d) compared to the *PSEN1* (33 years) and Down syndrome (32 years) (versus *PSEN1*: $z$-score = 0.92; $P$ value = 0.35 and versus Down syndrome: $z$-score = 1.19; $P$ value = 0.23), whereas it was significantly higher in *APOE3* homozygotes (versus *PSEN1*: $z$-score = 1.99; $P$ value = 0.04 and versus Down syndrome: $z$-score = 3.36; $P$ value < 0.01)[13]. The predictability of the age at symptom onset was similar when splitting by sex (Supplementary Fig. 1).

### Natural history of AD biomarker changes in *APOE4* homozygotes

To explore the timing of changes in biomarkers, we employed the concept of 'estimated years to symptom onset', setting the baseline age at 65.6 years to zero (age at symptom onset in *APOE4* homozygotes), as is commonly found in the study of other genetically determined forms of AD. We compared the trajectories of several biomarkers of the AT(N) framework with age and with respect to the estimated years to symptom onset in *APOE4* and *APOE3* homozygotes (Fig. 3). We assessed the age at which divergence occurs by visually examining the locally estimated scatterplot smoothing curves. The onset of CSF $A\beta_{1-42}$ concentrations in *APOE4* homozygotes cannot be ascertained as there were already differences in the youngest individuals in their late 40s. The increases in Centiloid scores started visually before age 50 years (15 years from symptom onset). CSF pTau and plasma pTau concentrations in *APOE4* homozygotes followed similar trajectories, with their concentrations visually starting to increase in participants in their early 50s, around 10–15 years before symptom onset. Regarding neurodegeneration biomarkers, plasma concentrations of neurofilament light chain (NfL) showed a steep increase in all groups in a pattern similar to that of hippocampal atrophy. Of note, the start of the hippocampal atrophy was difficult to ascertain as hippocampal volumes showed a steep decrease with age at all ages and for all haplotypes, probably reflecting the neurodevelopmental impact of *APOE4* on the medial temporal lobe. In any case, *APOE4* homozygotes clearly presented different volumes at the end of the sixth decade. Supplementary Fig. 3 shows the results for tau positron emission tomography (tau-PET) standardized uptake value ratio (SUVr) and Supplementary Fig. 4 shows the stratified analyses by sex.

### Comparison with ADAD and Down syndrome

We then constructed and integrated a model for the biomarker changes using the standardized differences between *APOE4* homozygotes and cognitively unimpaired *APOE3* homozygotes to better characterize the order and rate of pathophysiological changes in *APOE4* homozygotes and to compare them to the same model described in ADAD and Down syndrome (Fig. 4). The combined models clearly show the similarities in the temporality of biomarker changes in all three genetic conditions. The main difference between *APOE4* homozygotes and ADAD was found in the hippocampal atrophy, which showed smaller volumes at all ages included in this study.

### Similar biomarker changes at AD dementia stage across haplotypes

In addition, we investigated the changes in biomarkers with age among patients with a diagnosis of AD dementia. Consistent with the

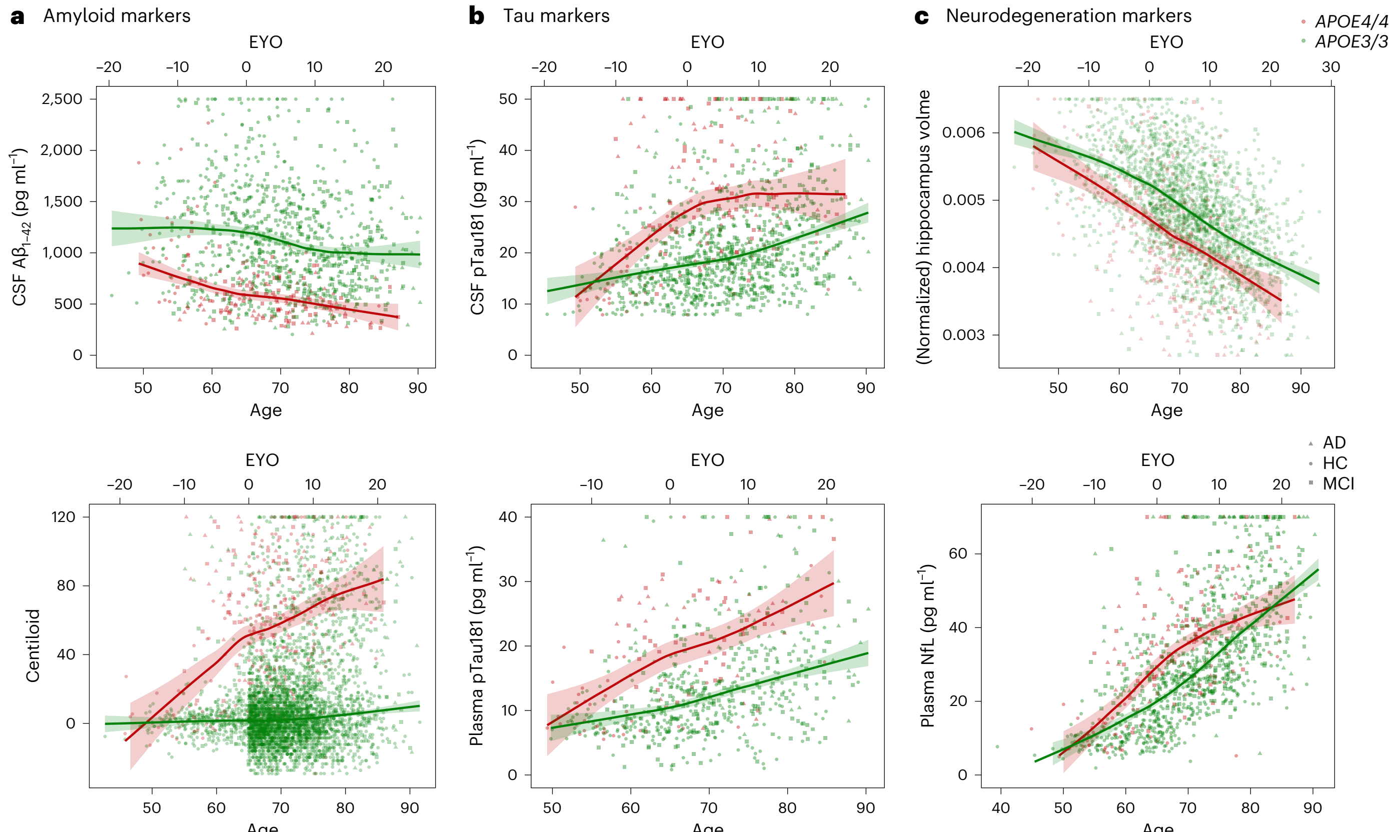


**Fig. 3 | Changes in AD biomarkers in *APOE3* and *APOE4* homozygotes with age.** Locally esimated scatterplot smoothing curves representing biomarker changes with age by *APOE* haplotype using the AT(N) classification system are shown. *APOE4* homozygotes are shown in red and *APOE3* homozygotes in green. There is a predictable sequence of biomarker changes in *APOE4* homozygotes. **a**, Trajectory of *APOE4* and *APOE3* homozygotes clearly deviate starting with amyloid biomarkers, which were already decreased in the youngest participants (50 years of age). **b**, Tau markers. **c**, Neurodegeneration biomarkers. Of note, *APOE4* homozygotes also showed hippocampal atrophy from the offset, probably reflecting the neurodevelopmental impact on the medial temporal lobe. Circles represent cognitively unimpaired participants, triangles represent patients with MCI and squares represent patients with AD dementia. The vertical dashed lines at 65 years represent the age of expected symptom onset in *APOE4* homozygotes. Shading represents 95% confidence intervals. Outliers have been visually truncated for ease of interpretation, while all statistical analyses were performed using the complete dataset. EYO, estimated years to symptom onset.

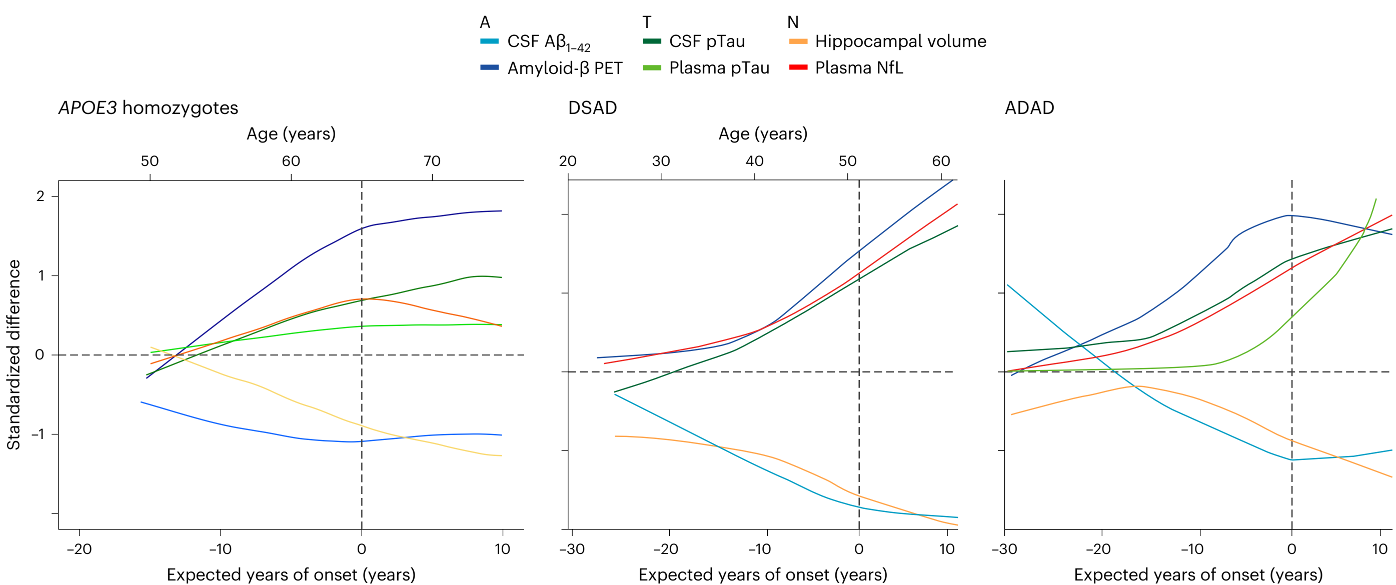


**Fig. 4 | Integrated model of the natural history of AD biomarker changes in *APOE4* homozygotes.** Biomarker changes (standardized differences) as a function of age and estimated years to symptom onset in *APOE4* homozygotes (left), DSAD (middle) and ADAD (right). The predictable sequence of biomarker changes is remarkably similar in *APOE4* homozygotes to that described in ADAD or DSAD. (The data are adapted from ref. 5).

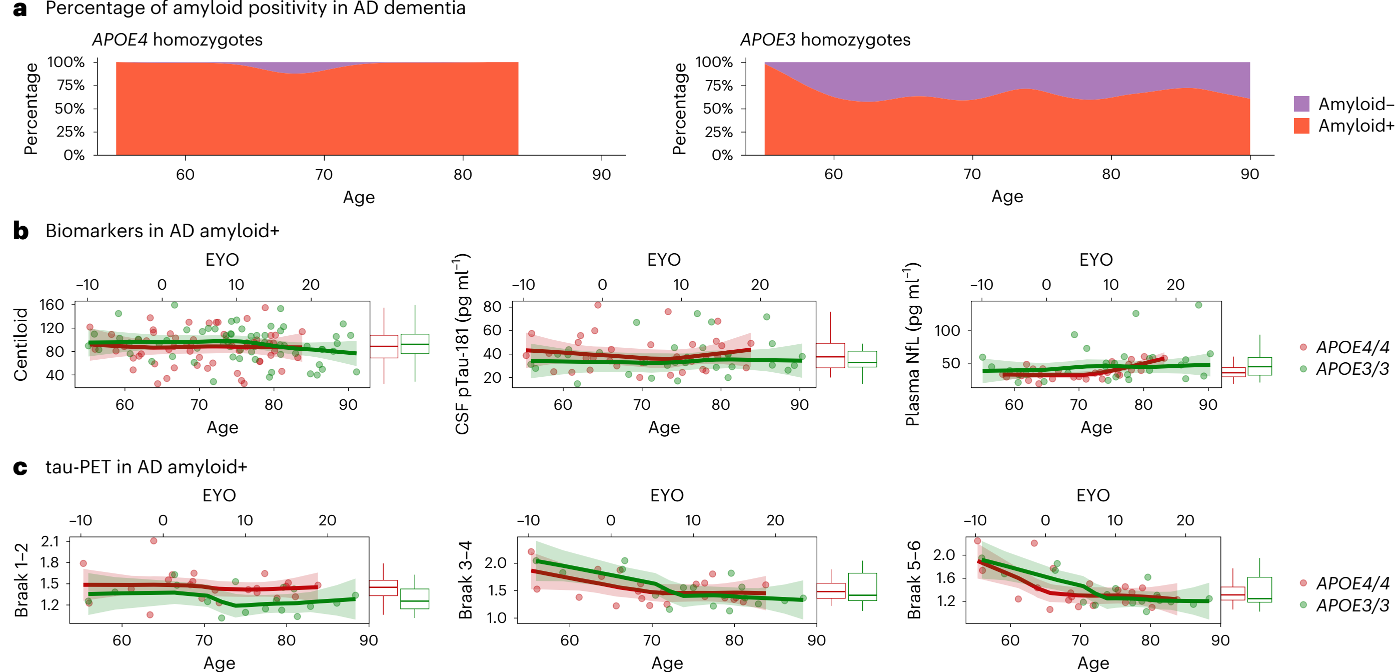


**Fig. 5 | Biomarker changes in patients with AD dementia. a**, Percentage of positivity in the amyloid PET scans by *APOE* haplotype. **b**, Biomarker levels (Centiloids, CSF pTau181 and plasma NfL levels) with age by *APOE* haplotype in individuals with AD dementia and a positive amyloid PET scan. **c**, The tau burden with age measured in the different Braak regions (1–2, 3–4 and 5–6) by *APOE* haplotype in individuals with AD dementia and a positive amyloid PET scan. We found no differences in the amyloid or tau-PET uptake between *APOE4* and *APOE3* homozygotes in these patients. Shading indicates the 95% confidence intervals. *APOE4* homozygotes are represented in red and *APOE3/3* carriers in green. The solid line represents the median and the dashed lines represent the 25th and 75th percentiles.

neuropathological findings, nearly all patients with AD dementia who had at least one *APOE4* allele were amyloid PET positive (irrespective of age), whereas amyloid PET positivity decreased with age in *APOE3* homozygotes (Fig. 5a).

To assess whether there were differences in the AD biology across haplotypes and a potential association with age, we selected those with a diagnosis of AD dementia and a positive amyloid PET scan (to avoid non-AD cases in *APOE4* noncarriers) and examined the differences in tau and amyloid biomarkers. Interestingly, we found no significant differences across haplotypes. We did not find age-related differences in Centiloid or CSF pTau levels in individuals with AD dementia (Fig. 5b). There were no differences between the different haplotypes in tau-PET uptake either, but tau-PET uptake decreased with age in all Braak regions for all haplotypes (Fig. 5c). This suggests that the difference in tau uptake across haplotypes with disease severity (Supplementary Fig. 3) might be driven by an earlier onset in *APOE4* homozygotes with respect to the other haplotypes.

### Beyond homozygosity: *APOE4* gene dose effect

Our study principally investigates *APOE4* homozygotes, establishing their similarities with ADAD and Down syndrome. Nonetheless, we extend our analysis to *APOE3* and *APOE4* heterozygotes in the Supplementary Information. These analyses underscore a distinct gene dose effect of *APOE4* on neuropathological (Supplementary Fig. 2) and in vivo (Supplementary Fig. 5a) biological penetrance, age of cognitive alterations presentation and death (Supplementary Fig. 5b), and biomarker profiles (Supplementary Fig. 6). *APOE3* and *APOE4* heterozygotes consistently exhibit intermediate phenotypes between *APOE3* and *APOE4* homozygotes. This gradient effect is delineated in Supplementary Figs. 4–7 and Supplementary Tables 1 and 2.

## Discussion

This study provides comprehensive evidence to propose *APOE4* homozygotes as another form of genetically determined AD, similar to ADAD and Down syndrome associated Alzheimer's disease (DSAD). We leveraged the unique resources of the NACC cohort and gathered one of the largest multicenter cohorts with multimodal AD biomarkers ($n > 10{,}000$) that enabled us to analyze more than 500 *APOE4* homozygotes. Our work showed that *APOE4* homozygotes meet the three main characteristics of genetically determined AD, namely near-full penetrance, symptom onset predictability and a predictable sequence of biomarker and clinical changes.

We first showed that *APOE4* homozygotes present near-full penetrance of AD biology. In this respect, it is worth noting that AD is now considered a biological entity that can be diagnosed in vivo based on the presence of AD biomarkers, irrespective of the presence or not of clinical symptoms[9]. Second, although previous studies had already reported the impact of the *APOE* haplotype in advancing symptom onset and risk for the disease[14], we demonstrated that symptom onset in *APOE4* homozygotes was as predictable as in ADAD and DSAD[13] (and significantly higher than in *APOE3* homozygotes). As a consequence, we propose a reappraisal of the conceptual framework and statistical approaches to favor the use of those commonly utilized in genetically determined dementias rather than the conceptual and analytical approach used in sporadic AD studies (for example, estimated years to symptom onset versus odds ratios)[5,15]. Finally, most biomarker studies collapse *APOE4* carriers into one group, mainly due to sample size considerations. However, there have also been fewer studies with small sample sizes or using only one or two modalities (mainly amyloid) that have shown an *APOE* gene dose effect[7,8,16–19]. Using an integrated model we could establish a predictable sequence of biomarker changes that was remarkably similar in sequence to that described in ADAD or DSAD[3,6]. We found distinct hippocampal volume patterns in *APOE4*

homozygotes, probably reflecting the potential neurodevelopmental impact on the medial temporal lobe and a shift toward a limbic predominant phenotype in AD presentation in *APOE4* carriers. Interestingly, when we restricted the analyses to patients with AD dementia, we found that, despite the very different ages at symptom onset, the biomarker changes were similar across haplotypes in patients with dementia of the same age. Of note, although Supplementary Fig. 3 shows that *APOE4* homozygotes initially appear to have an increased tau burden, this is moderated when both age and clinical status are considered (Fig. 5). In agreement with other studies we found a lower burden of tau with age, probably due to a higher prevalence of other copathologies and/or reduced physiological resilience to any form of pathology at older ages[20].

We propose a reconceptualization of the genetic architecture of AD, which is usually divided into the sporadic and autosomal dominant forms[2]. *APOE* is considered a risk factor rather than a causal gene. However, Genin et al.[4] proposed an autosomal semidominant inheritance for *APOE4* in AD, based on estimates of the lifetime risk for AD dementia in *APOE4* homozygotes (and the intermediate risk in heterozygotes) that can exceed 60–80% (refs. 4,14,21) in the range of major genes in Mendelian diseases. We provide an integrated clinical, pathological and biomarker confirmation of this hypothesis, and propose *APOE4* homozygotes should be considered as another form of genetically determined AD, like ADAD and DSAD[5]. We would like to note that Down syndrome underwent a similar recent reappraisal[5] based on the demonstration of universal AD pathology, the predictable sequence of biomarkers and clinical changes[6], and a near-full penetrance for dementia in this population[13]. Much like in Down syndrome (but not in ADAD), the later onset of clinical symptoms in *APOE4* homozygotes has led to an underestimation of its penetrance. Competing age-related causes of death often precede the manifestation of AD symptoms, thus reducing its observed prevalence and masking the true extent of AD.

The reconceptualization of genetically determined AD, inclusive of conditions like *APOE4* homozygosity and Down syndrome, necessitates reevaluating established beliefs. Family history may not always be a reliable indicator, as parents can carry and transmit genetic conditions like trisomy 21 or the *APOE4* allele without manifesting them. Traditionally, ADAD is characterized by an early onset, typically before age 65. However, this criterion should not exclusively determine the genetic basis of the disease as it overlooks the complexity of genetic factors. Notably, certain mutations in presenilin 2, despite being definitively linked to ADAD, exhibit symptom onset around, or sometimes after, age 65. Moreover, the expected age for symptom onset in all forms of genetic AD exhibits considerable variability, potentially influenced by other genetic variants and lifestyle factors that can modify disease expression and progression.

This reconceptualization has profound consequences. First, *APOE4* homozygotes and heterozygotes should not be combined as is usually done, as they represent two distinct genetic risk profiles. There is a strong gene dose effect on clinical, pathological and biomarker data, with *APOE3* or *APOE4* heterozygotes consistently exhibiting intermediate phenotypes between *APOE3* and *APOE4* homozygotes, which supports the concept of autosomal semidominance as suggested by Genin et al.[4]. Second, given that the incidence of *APOE4* homozygotes is approximately 2% (with racial and geographical variations)[22], they would in fact constitute one of the most frequent Mendelian diseases. This will have consequences in counseling and the recommendations to screen for *APOE* in the population and in the study of patients with cognitive complaints. Nevertheless, it is important to note that our findings predominantly reflect the risk association of *APOE4* homozygosity (and heterozygosity) within European ancestry populations. Recognizing the paucity of data on individuals of African descent, we stress the recent findings suggesting differential *APOE4*-related risks across ancestries[23]. Future research must prioritize the inclusion of diverse populations to elucidate the full scope of effect of *APOE4* on AD risk, ensuring that genetic insights translate into benefits for all ethnicities. Third, *APOE4* homozygotes would share the unique opportunities for research and trials recognized for genetically determined AD. These opportunities start to be recognized (ClinicalTrials.gov registration: NCT04770220).

Our study has several limitations that must be acknowledged. One limitation of this study is the use of convenience samples, which may not accurately represent the broader population. Our separate analyses on the NACC dataset and clinical cohorts each come with their own biases. NACC leans toward symptomatic participants, whereas clinical cohorts have stringent inclusion criteria that favor cognitive health, possibly underestimating the true burden of AD. Despite these constraints, the consistent findings of near-full biological AD penetrance among *APOE4* homozygotes in both datasets bolster the robustness of our study's conclusions. The lack of $A\beta_{1-40}$ levels or the use of the ADNC, whose score is not a granular measure of specific neuropathological lesions, are other limitations. However, the ADNC is nevertheless broadly accepted as a general marker of AD neuropathology. The lack of centralized biomarker assessment across multiple centers is another limitation. Despite our efforts to standardize data by exclusively including studies using the same Roche platform and adding 'site' as a covariate in our statistical models, intersite variations could still introduce bias. Another major limitation is the cross-sectional design. Together with the biases of our convenience cohort, we were unable to calculate incidence or cumulative incidence of AD dementia for each haplotype. Future longitudinal studies considering competing mortality risks will allow for a more comprehensive understanding of the disease risk and its progression over time. The relationship between *APOE4* homozygosity and AD risk may be obscured by higher mortality from other age-related conditions[24]. Additionally, as mentioned, all participants came from the USA or Europe and were predominantly white. However, there are geographical differences in *APOE4* frequency and ethnic risk mitigation, with *APOE4* conferring a lesser risk in Black than in white populations[3]. Future studies should focus on population-based studies with diverse origins.

In conclusion, our study provides compelling evidence to propose that *APOE4* homozygotes represent a distinct, genetically determined form of AD, which has important implications for public health, genetic counseling of carriers and future research directions.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information, details of author contributions and competing interests, and statements of data and code availability are available at https://doi.org/10.1038/s41591-024-02931-w.

## References

1. Bellenguez, C. et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat. Genet.* **54**, 412–436 (2022).
2. Frisoni, G. B. et al. The probabilistic model of Alzheimer disease: the amyloid hypothesis revised. *Nat. Rev. Neurosci.* **23**, 53–66 (2022).
3. Bateman R. J. et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* **367**, 795–804 (2012).
4. Genin, E. et al. *APOE* and Alzheimer disease: a major gene with semidominant inheritance. *Mol. Psychiatry* **16**, 903–907 (2011).
5. Fortea, J. et al. Alzheimer's disease associated with Down syndrome: a genetic form of dementia. *Lancet Neurol.* **20**, 930–942 (2021).
6. Fortea, J. et al. Clinical and biomarker changes of Alzheimer's disease in adults with Down syndrome: a cross-sectional study. *Lancet* **395**, 1988–1997 (2020).

7. Jansen, W. J. et al. Prevalence of cerebral amyloid pathology in persons without dementia: a meta-analysis. *JAMA* **313**, 1924–1938 (2015).
8. Saddiki H. et al. Age and the association between apolipoprotein E genotype and Alzheimer disease: a cerebrospinal fluid biomarker-based case-control study. *PLoS Med.* https://doi.org/10.1371/JOURNAL.PMED.1003289 (2020).
9. Jack, C. R. et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* **14**, 535–562 (2018).
10. Beekly, D. L. et al. The National Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. *Alzheimer Dis. Assoc. Disord.* **18**, 270–277 (2004).
11. Montine, T. J. et al. National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. *Acta Neuropathol.* **123**, 1–11 (2012).
12. Reiman, E. M. et al. Exceptionally low likelihood of Alzheimer's dementia in *APOE2* homozygotes from a 5,000-person neuropathological study. *Nat. Commun.* **11**, 1–11 (2020).
13. Iulita M. F. et al. Association of Alzheimer disease with life expectancy in people with Down syndrome. *JAMA Netw. Open* https://doi.org/10.1001/JAMANETWORKOPEN.2022.12910 (2022).
14. Corder, E. H. et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* **261**, 921–923 (1993).
15. Fortea, J., Quiroz, Y. T. & Ryan, N. S. Lessons from Down syndrome and autosomal dominant Alzheimer's disease. *Lancet Neurol.* **22**, 5–6 (2023).
16. Therriault, J. et al. Frequency of biologically defined Alzheimer's disease in relation to age, sex, *APOE* ε4, and cognitive impairment. *Neurology* **96**, e975–e985 (2021).
17. Betthauser, T. J. et al. Multi-method investigation of factors influencing amyloid onset and impairment in three cohorts. *Brain* **145**, 4065–4079 (2022).
18. Snellman, A. et al. *APOE ε4* gene dose effect on imaging and blood biomarkers of neuroinflammation and beta-amyloid in cognitively unimpaired elderly. *Alzheimers Res. Ther.* **15**, 71 (2023).
19. Ghisays, V. et al. Brain imaging measurements of fibrillar amyloid-β burden, paired helical filament tau burden, and atrophy in cognitively unimpaired persons with two, one, and no copies of the *APOE ε4* allele. *Alzheimers Dement.* **16**, 598–609 (2020).
20. Mehta, R. I. & Schneider, J. A. What is 'Alzheimer's disease'? The neuropathological heterogeneity of clinically defined Alzheimer's dementia. *Curr. Opin. Neurol.* **34**, 237–245 (2021).
21. van der Lee, S. J. et al. The effect of *APOE* and other common genetic variants on the onset of Alzheimer's disease and dementia: a community-based cohort study. *Lancet Neurol.* **17**, 434–444 (2018).
22. Belloy, M. E., Napolioni, V. & Greicius, M. D. A quarter century of *APOE* and Alzheimera's disease: progress to date and the path forward. *Neuron* **101**, 820–838 (2019).
23. Belloy, M. E. et al. *APOE* genotype and Alzheimer disease risk across age, sex, and population ancestry. *JAMA Neurol.* **80**, 1284–1294 (2023).
24. Jack, C. R. et al. Long-term associations between amyloid positron emission tomography, sex, apolipoprotein E and incident dementia and mortality among individuals without dementia: hazard ratios and absolute risk. *Brain Commun.* **4**, fcac017 (2022).


**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2024

[1]Sant Pau Memory Unit, Hospital de la Santa Creu i Sant Pau - Biomedical Research Institute Sant Pau, Barcelona, Spain. [2]Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas. CIBERNED, Barcelona, Spain. [3]Barcelona Down Medical Center, Fundació Catalana Síndrome de Down, Barcelona, Spain. [4]Department of Medicine, Faculty of Medicine and Health Sciences, Institute of Neurosciences, University of Barcelona, Barcelona, Spain. [5]BarcelonaβBeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain. [6]Neurosciences Programme, IMIM - Hospital del Mar Medical Research Institute, Barcelona, Spain. [7]Department of Medicine and Life Sciences, Universitat Pompeu Fabra, Barcelona, Spain. [8]Centro de Investigación Biomédica en Red Bioingeniería, Biomateriales y Nanomedicina. Instituto de Salud carlos III, Madrid, Spain. [9]Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain. [10]Wisconsin Alzheimer's Disease Research Center, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI, USA. [11]Brigham and Women's Hospital Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA. [12]Barcelona Supercomputing Center, Barcelona, Spain. [13]These authors contributed equally: Juan Fortea, Víctor Montal. ✉e-mail: jfortea@santpau.cat; victor.montal@protonmail.com

# Methods

## Study design

We used two distinct sources of human data: (1) a neuropathological study using data from NACC[10], and (2) an in vivo study using data from five clinical cohorts that included multiple biomarkers. A description of each of the cohorts can be found in the Supplementary Information.

The study was approved by the Sant Pau Ethics Committee in accordance with the recommendations of the Declaration of Helsinki. Each cohort involved in the study obtained approval from its respective Institutional Review Board committee.

## Participants

**National Alzheimer's Coordinating Center.** We included participants with a neuropathological evaluation[11], *APOE* haplotype information (data accessed on 1 August 2022) and a clinical evaluation[25]. Most participants also had information on the age at disease onset (symptom onset, MCI and/or dementia). Further details on the participants' diagnosis and a description of the neuropathological scoring can be found in the Supplementary Information.

**Clinical cohorts.** Our study included cross-sectional data from five multisite cohorts: (1) Alzheimer's Disease Neuroimaging Initiative ($n = 2{,}214$)[26]; (2) The A4 Study ($n = 5{,}477$)[27]; (3) the ALFA Study ($n = 418$)[28]; (4) the Wisconsin Register for Alzheimer's Prevention ($n = 612$)[29] and (5) the OASIS3 Project ($n = 1{,}317$)[30]. All five cohorts include a diverse set of biomarkers across the AD continuum, with a special emphasis on preclinical AD. We included all the available data from participants with at least one available biomarker of interest, a clinical diagnosis and an *APOE* haplotype.

## *APOE* genotyping

We included participants with available *APOE* haplotype as reported at each site[26–30].

## Biochemical analysis

A subset of 1,665 participants from three sites (ADNI, Alfa + and Wisconsin Register for Alzheimer's Prevention) underwent biofluid measurements. All sites followed a similar processing pipeline, and protein levels were quantified using the same technology across all cohorts[26,28,29]. Specifically, Elecsys was used to measure CSF $A\beta_{1-42}$ and pTau181 levels, and SIMOA was used to measure plasma pTau and NfL levels. We used the biofluid quantification directly provided by each cohort. Of note, three of the five clinical sites do not have $A\beta_{1-40}$ measurements, and we did not include the $A\beta_{1-42}$ or $A\beta_{1-40}$ ratio.

## Brain imaging

A subgroup of 5,108 participants underwent assessments of hippocampal volume using T1-weighted MRI data and we used the mean volume of the bilateral hippocampus normalized by the total intracranial volume as reported at each site. Another subset of 7,490 participants underwent amyloid PET imaging using the Pittsburgh compound B, Florbetapir or Flutemetamol tracers. SUVr measures were reported by each site and transformed into Centiloid scale scores to integrate data from the different tracers. To classify participants as amyloid-positive, we used a threshold of 24.4 Centiloids[31]. Another subset of 1,267 participants underwent tau-PET imaging with flortaucipir. We quantified the SUVr in the different Braak stage regions (Braak 1–2, 3–4 and 5–6) using regions of interest from Freesurfer Desikan Atlas. A detailed description on the tau-PET quantification pipeline can be found in the Supplementary Information.

## Statistical analyses

All statistical analyses were conducted using R (v.4.2.2), utilizing the 'survival', 'survminer' and 'statsExpressions' packages. Demographic differences across groups were evaluated using chi-square tests for categorical variables and Kruskal–Wallis tests for continuous variables. Pairwise comparisons followed, deploying the Dwass–Steel–Critchlow–Fligner method.

We used the NACC dataset to analyze the distribution of ADNC scores across different ages at death, comparing *APOE4* and *APOE3* homozygotes, and to analyze the impact of the *APOE* haplotype on disease onset and its predictability (as assessed by the 95% prediction interval). We statistically compared the different age at onset between haplotypes using Kruskal–Wallis tests (followed by pairwise comparisons using the Dwass–Steel–Critchlow–Fligner test). We also used data from ref. 13 to compare predictive intervals for *APOE4* and *APOE3* homozygotes, *PSEN1* and Down syndrome[13]. We calculated the average age and standard deviation for symptom onset in *PSEN1* and Down syndrome, then *z*-normalized the predictive intervals for *APOE4* and *APOE3* to assess statistical similarity based on a normal distribution.

We also conducted a Kaplan–Meier survival analysis to ascertain the gene dosage effect on the age at symptom onset and age at death. We assessed the probability of remaining free from a dementia diagnosis and surviving across different time points using the Kaplan–Meier method, which was followed by a Cox regression model for further insight. Participants from the NACC were stratified based on ADNC status into two categories: ADNC positive (high or intermediate) and ADNC negative (none or low). We then evaluated the survival probabilities for each group. Incorporating sex as a covariate, the Cox regression analysis yielded statistically significant differences in survival outcomes between the groups, underscoring the influence of AD pathology on disease progression and mortality. All remaining analyses were performed using the clinical cohorts' dataset. To determine the order and temporality of the biomarker changes across haplotypes, we first compared the frequency of positive amyloid and tau biomarkers across 5-year age intervals for both *APOE4* and *APOE3* homozygotes. We used a previously reported cohort-specific threshold to binarize amyloid and tau biomarkers into positive and negative. We also compared the biomarker levels at each age interval using Welch's *t*-test and including the clinical cohort as a covariate. We also fitted a first-order locally estimated scatterplot smoothing curve in each haplotype independently using a standard tricubic weight function with a span parameter to 0.75 (refs. 3,6). As in our previous study, we defined biomarker change as the age at which the groups appear to start diverging visually, because the exact age at which the confidence intervals diverge is dependent on intrinsic limitations of studies assessing the natural history of biomarkers, such as the nature of the variable, the sensitivity of the assay, the slope of the association and, in our study, the uneven sample sizes for the different biomarkers[6]. To compare the timing of changes in *APOE4* homozygotes to those in Down syndrome and ADAD, we constructed a model of the standardized difference between all *APOE4* and healthy control *APOE* homozygotes as a function of estimated years from expected symptom onset[3,6]. Concretely, we included the whole set of *APOE4* homozygotes, independently of their cognitive status, and normalized their biomarker scores by the values from cognitively unimpaired *APOE3* homozygotes.

Finally, to compare the AD biology across haplotypes in the dementia stage, we compared the biomarker changes to age in those participants with a diagnosis of AD dementia and a positive amyloid scan to avoid the bias introduced by the differential risk of AD biology across haplotypes.

Concretely, we used Welch's *t*-test to compare between *APOE* haplotypes. We performed several post hoc sensitivity analyses to investigate the possibility of site-specific variations, which may have been caused by differences in imaging protocols or biochemical biomarker analysis protocols. These analyses are available in the Supplementary Information. The Supplementary Information also provides an in-depth explanation of the statistical methods used to generate the Supplementary Results.

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Access to tabular data from ADNI (https://adni.loni.usc.edu/), OASIS (https://oasis-brains.org/), A4 (https://ida.loni.usc.edu/collaboration/access/appLicense.jsp) and NACC (https://naccdata.org/) can be requested online, as publicly available databases. All requests will be reviewed by each study's scientific board. Concrete inquiries to access the WRAP (https://wrap.wisc.edu/data-requests-2/) and ALFA + (https://www.barcelonabeta.org/en/alfa-study/about-the-alfa-study) cohort data can be directed to each study team for concept approval and feasibility consultation. Requests will be reviewed to verify whether the request is subject to any intellectual property.

## Code availability

All statistical analyses and raw figures were generated using R (v.4.2.2). We used the open-sourced R packages of ggplot2 (v.3.4.3), dplyr (v.1.1.3), ggstream (v.0.1.0), ggpubr (v.0.6), ggstatsplot (v.0.12), Rmisc (v.1.5.1), survival (v.3.5), survminer (v.0.4.9), gtsummary (v.1.7), epitools (v.0.5) and statsExpression (v.1.5.1). Rscripts to replicate our findings can be found at https://gitlab.com/vmontalb/apoe4-asdad (ref. 32). For neuroimaging analyses, we used Free Surfer (v.6.0) and ANTs (v.2.4.0).

## References


25. Morris, J. C. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* **43**, 2412–2414 (1993).
26. Weiner, M. W. et al. The Alzheimer's Disease Neuroimaging Initiative 3: continued innovation for clinical trial improvement. *Alzheimer's Dement.* **13**, 561–571 (2017).
27. Sperling R. A. et al. The A4 Study: stopping AD before symptoms begin? *Sci. Transl. Med.* https://doi.org/10.1126/scitranslmed.3007941 (2014).
28. Molinuevo, J. L. et al. The ALFA project: a research platform to identify early pathophysiological features of Alzheimer's disease. *Alzheimer's Dement.: Transl. Res. Clin. Interventions* **2**, 82–92 (2016).
29. Johnson, S. C. et al. The Wisconsin Registry for Alzheimer's Prevention: a review of findings and current directions. *Alzheimer's Dement.: Diagnosis, Assess. Dis. Monit.* **10**, 130–142 (2018).
30. LaMontagne P. J. et al. OASIS-3: longitudinal neuroimaging, clinical and cognitive dataset for normal aging and Alzheimer disease. Preprint at *MedRxiv* https://doi.org/10.1101/2019.12.13.19014902 (2019).
31. La Joie, R. et al. Multisite study of the relationships between antemortem [$^{11}$C]PIB-PET Centiloid values and postmortem measures of Alzheimer's disease neuropathology. *Alzheimers Dement.* **15**, 205–216 (2019).
32. Montal, V. APOE4-ASDAD. *GitLab* https://gitlab.com/vmontalb/apoe4-asdad (2024).


## Acknowledgements


We acknowledge the contributions of several consortia that provided data for this study. We extend our appreciation to the NACC, the Alzheimer's Disease Neuroimaging Initiative, The A4 Study, the ALFA Study, the Wisconsin Register for Alzheimer's Prevention and the OASIS3 Project. Without their dedication to advancing Alzheimer's disease research and their commitment to data sharing, this study would not have been possible. We also thank all the participants and investigators involved in these consortia for their tireless efforts and invaluable contributions to the field. We also thank the institutions that funded this study, the Fondo de Investigaciones Sanitario, Carlos III Health Institute, the Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas and the Generalitat de Catalunya and La Caixa Foundation, as well as the NIH, Horizon 2020 and the Alzheimer's Association, which was crucial for this research. Funding: National Institute on Aging. This study was supported by the Fondo de Investigaciones Sanitario, Carlos III Health Institute (INT21/00073, PI20/01473 and PI23/01786 to J.F., CP20/00038, PI22/00307 to A.B., PI22/00456 to M.S.-C., PI18/00435 to D.A., PI20/01330 to A.L.) and the Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas Program 1, partly jointly funded by Fondo Europeo de Desarrollo Regional, Unión Europea, Una Manera de Hacer Europa. This work was also supported by the National Institutes of Health grants (R01 AG056850; R21 AG056974, R01 AG061566, R01 AG081394 and R61AG066543 to J.F., S10 OD025245, P30 AG062715, U54 HD090256, UL1 TR002373, P01 AG036694 and P50 AG005134 to R.S.; R01 AG027161, R01 AG021155, R01 AG037639, R01 AG054059; P50 AG033514 and P30 AG062715 to S.J.) and ADNI (U01 AG024904), the Department de Salut de la Generalitat de Catalunya, Pla Estratègic de Recerca I Innovació en Salut (SLT006/17/00119 to J.F.; SLT002/16/00408 to A.L.) and the A4 Study (R01 AG063689, U24 AG057437 to R.A.S). It was also supported by Fundación Tatiana Pérez de Guzmán el Bueno (IIBSP-DOW-2020-151 o J.F.) and Horizon 2020–Research and Innovation Framework Programme from the European Union (H2020-SC1-BHC-2018-2020 to J.F.; 948677 and 847648 to M.S.-C.). La Caixa Foundation (LCF/PR/GN17/50300004 to M.S.-C.) and EIT Digital (Grant 2021 to J.D.G.) also supported this work. The Alzheimer Association also participated in the funding of this work (AARG-22-923680 to A.B.) and A4/LEARN Study AA15-338729 to R.A.S.). O.D.-I. receives funding from the Alzheimer's Association (AARF-22-924456) and the Jerome Lejeune Foundation postdoctoral fellowship.


## Author contributions

J.F. and V.M. conceptualized the research project and drafted the initial manuscript. V.M., J.P. and J.F. conducted data analysis, interpreted statistical findings and created visual representations of the data. O.B. and O.D.-I. provided valuable insights into the genetics of APOE. L.V., A.B. and L.V.-A. meticulously reviewed and edited the manuscript for clarity, accuracy and coherence. J.D.G., M.S.-C., S.J. and R.S. played pivotal roles in data acquisition and securing funding. A.L. and D.A. contributed to the study design, offering guidance and feedback on statistical analyses, and provided critical review of the paper. All authors carefully reviewed the manuscript, offering pertinent feedback that enhanced the study's quality, and ultimately approved the final version.

## Competing interests

S.C.J. has served at scientific advisory boards for ALZPath, Enigma and Roche Diagnostics. M.S.-C. has given lectures in symposia sponsored by Almirall, Eli Lilly, Novo Nordisk, Roche Diagnostics and Roche Farma, received consultancy fees (paid to the institution) from Roche Diagnostics and served on advisory boards of Roche Diagnostics and Grifols. He was granted a project and is a site investigator of a clinical trial (funded to the institution) by Roche Diagnostics. In-kind support for research (to the institution) was received from ADx Neurosciences, Alamar Biosciences, Avid Radiopharmaceuticals, Eli Lilly, Fujirebio, Janssen Research & Development and Roche Diagnostics. J.D.G. has served as consultant for Roche Diagnostics, receives research funding from Hoffmann–La Roche, Roche Diagnostics and GE Healthcare, has given lectures in symposia sponsored by Biogen, Philips Nederlands, Esteve and Life Molecular Imaging and serves on an advisory board for Prothena Biosciences. R.S. has received personal consulting fees from Abbvie, AC Immune, Acumen, Alector, Bristol Myers Squibb, Janssen, Genentech, Ionis and Vaxxinity outside the submitted work.

O.B. reported receiving personal fees from Adx NeuroSciences outside the submitted work. D.A. reported receiving personal fees for advisory board services and/or speaker honoraria from Fujirebio-Europe, Roche, Nutricia, Krka Farmacéutica and Esteve, outside the submitted work. A.L. has served as a consultant or on advisory boards for Almirall, Fujirebio-Europe, Grifols, Eisai, Lilly, Novartis, Roche, Biogen and Nutricia, outside the submitted work. J.F. reported receiving personal fees for service on the advisory boards, adjudication committees or speaker honoraria from AC Immune, Adamed, Alzheon, Biogen, Eisai, Esteve, Fujirebio, Ionis, Laboratorios Carnot, Life Molecular Imaging, Lilly, Lundbeck, Perha, Roche and outside the submitted work. O.B., D.A., A.L. and J.F. report holding a patent for markers of synaptopathy in neurodegenerative disease (licensed to Adx, EPI8382175.0). The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41591-024-02931-w.

**Correspondence and requests for materials** should be addressed to Juan Fortea or Víctor Montal.

**Peer review information** *Nature Medicine* thanks Naoyuki Sato, Yadong Huang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Jerome Staal, in collaboration with the *Nature Medicine* team.

**Reprints and permissions information** is available at www.nature.com/reprints.

Corresponding author(s): Juan Fortea, Victor Montal

Last updated by author(s): Dec 30, 2023

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed | |
|---|---|---|
| ☐ | ☒ | The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ | A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☐ | ☒ | The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☐ | ☒ | A description of all covariates tested |
| ☐ | ☒ | A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ | A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☐ | ☒ | For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ | For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ | For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ | Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | N/A. We did not collect data for the current study |
| Data analysis | All statistical analyses and raw figures were generated using R (v). Concretely, we used the open-sourced R packages of ggplot2, dlpyr, ggstram, ggpubr, ggstatsplot, Rmisc, survival, survminer, gtsummary, epitools and statsExpressions. Rscripts to replicate the findings can be requested to corresponding authors. |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Access to tabular data from ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/)

can be requested online. All requests will be reviewed by each study scientific board. Concrete inquiries to access the ALFA+ cohort data can be directed to M.S.C and/or J.D.G authors. Requests will be reviewed to verify whether the request is subject to any intellectual property.

## Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

| | |
|---|---|
| Reporting on sex and gender | Sex was determined based on self-report and was considered in the study design and the analyses. We do also report all our main findings segregating by sex |
| Population characteristics | *Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."* |
| Recruitment | A detailed description and protocol for paritcipant recruitment can be found for each of the cohorts at their webpage. Concretely, ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/), Alfa+ (https://fpmaragall.org/investigacion-alzheimer/estudio-alfa-contra-alzheimer/) and A4 (https://a4study.org/) |
| Ethics oversight | The study was approved by the Sant Pau Ethics Committee in accordance with the recommendations of the Declaration of Helsinki. Each cohort involved in the study obtained approval from its respective Institutional Review Board committee. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | 3,297 datapoints from brain donors (NACC) and 10,039 from clinical cohorts |
| Data exclusions | We excluded participants that do not have an APOE genotype screening. |
| Replication | Observational study. We did not include a replication cohort due to small propensity of grup of interest (APOE-4/4) |
| Randomization | NO randomization |
| Blinding | NO Blinding design |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

### Materials & experimental systems

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | Antibodies |
| ☒ | ☐ | Eukaryotic cell lines |
| ☒ | ☐ | Palaeontology and archaeology |
| ☒ | ☐ | Animals and other organisms |
| ☐ | ☒ | Clinical data |
| ☒ | ☐ | Dual use research of concern |

### Methods

| n/a | Involved in the study | |
|---|---|---|
| ☒ | ☐ | ChIP-seq |
| ☒ | ☐ | Flow cytometry |
| ☐ | ☒ | MRI-based neuroimaging |

## Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

| | |
|---|---|
| Clinical trial registration | N/A . Not a clinical trial |
| Study protocol | N/A . Not a clinical trial |
| Data collection | Each of the included cohorts follow a specific protocol for data collection. We direct the reader to each of the cohorts for a full description: ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/), Alfa+ (https://fpmaragall.org/investigacion-alzheimer/estudio-alfa-contra-alzheimer/) and A4 (https://a4study.org/) |
| Outcomes | N/A . Not a clinical trial. We assessed biofluids markers (CSF amyloid 42 and pTau, Plasma NfL, pTau), MRI measures (hippocampus volume) and amyloid and tau PET measures. We did also used CDR scores in supplementary analyses. |

## Magnetic resonance imaging

### Experimental design

| | |
|---|---|
| Design type | N/A |
| Design specifications | N/A |
| Behavioral performance measures | N/A |

### Acquisition

| | |
|---|---|
| Imaging type(s) | Structural MRI |
| Field strength | 3T |
| Sequence & imaging parameters | 1mm isotropic. Site-specific variations of TR and TE (controlled by ad-hoc supplementary analyses to assess sequence influence on the results) |
| Area of acquisition | whole brain |

Diffusion MRI ☐ Used ☒ Not used

### Preprocessing

| | |
|---|---|
| Preprocessing software | Freesurfer, FSL, or ANTs (site/cohort specific) |
| Normalization | Site pipeline specific. We direct the reader to each cohort details: ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/), Alfa+ (https://fpmaragall.org/investigacion-alzheimer/estudio-alfa-contra-alzheimer/) and A4 (https://a4study.org/) |
| Normalization template | Site pipeline specific. We direct the reader to each cohort details: ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/), Alfa+ (https://fpmaragall.org/investigacion-alzheimer/estudio-alfa-contra-alzheimer/) and A4 (https://a4study.org/) |
| Noise and artifact removal | Site pipeline specific. We direct the reader to each cohort details: ADNI (https://adni.loni.usc.edu/), WRAP (https://wrap.wisc.edu/), OASIS (https://oasis-brains.org/) and NACC (https://naccdata.org/), Alfa+ (https://fpmaragall.org/investigacion-alzheimer/estudio-alfa-contra-alzheimer/) and A4 (https://a4study.org/) |
| Volume censoring | N/A |

### Statistical modeling & inference

| | |
|---|---|
| Model type and settings | *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).* |
| Effect(s) tested | *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.* |

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

| | |
|---|---|
| Anatomical location(s) | Segmented the hippocampal and the whole brain volumes using Freesurfer, FSL or ANTS (cohort-speciic). We perfromed statistical analyses (see above) based on the normalized hippocampal volume by whole brain volume to control for intra-individual effects. |

| | |
|---|---|
| Statistic type for inference (See Eklund et al. 2016) | ROI-based. Two-sample t-test for statistical inference. Sub-analyses by sex and site. |
| Correction | *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).* |

## Models & analysis

| n/a | Involved in the study |
|---|---|
| ☐ | ☒ Functional and/or effective connectivity |
| ☐ | ☒ Graph analysis |
| ☐ | ☒ Multivariate modeling or predictive analysis |

| | |
|---|---|
| Functional and/or effective connectivity | *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).* |
| Graph analysis | *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).* |
| Multivariate modeling and predictive analysis | *Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.* |