Title: Towards cascading genetic risk in Alzheimer's disease

Andre Altmann[1], Leon M Aksman[2], Neil P Oxtoby[1], Alexandra Young[1], ADNI*, Daniel C Alexander[1], Frederik Barkhof[1,3,4], Maryam Shoai[3,5], John Hardy[3,5], Jonathan M Schott[5,6]

1 UCL Centre for Medical Image Computing, Department of Medical Physics and Biomedical Engineering & Department of Computer Science, University College London, UK

2 Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA

3 UCL Queen Square Institute of Neurology, University College London, UK

4 Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, The Netherlands

5 UK Dementia Research Institute, University College London, UK

6 Dementia Research Centre, UCL Queen Square Institute of Neurology, University College London, UK

Corresponding author:

Dr. Andre Altmann

90 High Holborn, 1st Floor

London, WC1V 6LJ

United Kingdom

Email: a.altmann@ucl.ac.uk

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

Abstract

Background: Alzheimer's disease (AD) typically progresses in stages, which have been defined by the presence of disease-specific biomarkers: Amyloid (A), Tau (T) and neurodegeneration (N). This progression of biomarkers has been condensed into the ATN framework, where each of the biomarkers can be either positive (+) or negative (-). Over the past decades genome wide association studies (GWAS) have implicated about 90 different loci involved with the development of late onset Alzheimer's disease (LOAD). Here we investigate whether genetic risk for AD contributes equally to the progression in different disease stages or whether it exhibits a stage-dependent effect.

Methods: Amyloid (A) and tau (T) status was defined using a combination of available PET and CSF biomarkers in the AD Neuroimaging Initiative (ADNI) cohort. In 425 participants with biomarker-confirmed A-T- status, we employed Cox proportional hazards models to estimate the contribution of *APOE* and polygenic risk scores (PRS; beyond *APOE*) to convert to A+T- status (68 conversions). Furthermore, we repeated the analysis in 402 participants with A+T- status and investigated the genetic contribution to conversion to A+T+ (46 conversions). Both survival analyses were adjusted for age, sex, and years of education.

Results: For progression from A-T- to A+T-, APOE-e4 burden showed significant effect (HR=2.71; 95% CI: 1.65-4.45; P<0.001), while PRS did not (HR=1.06; 95% CI: 0.82-1.40; P=0.65). Conversely, for the transition from A+T- to A+T+, the APOE-e4 burden contribution was reduced (HR=1.65 95% CI: 1.05-2.59; P=0.030), while the PRS showed an increased contribution (HR=1.67; 95% CI:1.22-2.28; P=0.0014). The marginal *APOE* effect was driven by e4 homozygotes (HR=2.81; 95% CI: 1.14-6.91; P=0.025) as opposed to e4 heterozygotes (HR=1.5; 95% CI: 0.75-2.99; P=0.25).

Conclusion: The genetic risk for LOAD unfolds in a disease stage-dependent fashion. A better understanding of the interplay between disease stage and genetic risk can lead to a more mechanistic understanding of transition between ATN stages, a better understanding of the molecular processes leading to AD as well as opening therapeutic windows for targeted interventions.

Introduction

Alzheimer's disease is characterized, on the neuropathological level, by the buildup of two proteins: amyloid plaques and neurofibrillary tangles of phosphorylated tau (Braak *et al.*, 1998). Both these pathological features can be observed long before the memory loss and decline in executive function that is characteristic of patients with Alzheimer's disease. Accumulation of the amyloid plaques in the brain predates the clinical symptoms of AD by two decades (Bateman *et al.*, 2012), while the spatial distribution of tau tangles reflects more closely the reported cognitive deficits and neurodegeneration (Ossenkoppele *et al.*, 2016).

The amyloid cascade hypothesis postulates that the deposition of the amyloid-β protein – the main component of the amyloid plaques – is the cause of AD, and that neurofibrillary tangles, cell loss, vascular damage and dementia are a direct consequence (Selkoe and Hardy, 2016). In keeping with the amyloid cascade hypothesis, a theoretical framework for the progression of biomarkers during the course of AD has been developed (Jack Jr *et al.*, 2013). Here, amyloid pathology is the first to appear, followed by tau pathology, neurodegeneration and finally cognitive decline. Support for this theoretical framework comes from a number of lines of evidence, including a variety of data-driven modeling approaches based on biomarker data (Fonteijn *et al.*, 2012; Jedynak *et al.*, 2012; Donohue *et al.*, 2014; Young *et al.*, 2014; Lorenzi *et al.*, 2019). In an attempt to operationalize this, a simplified ATN model has been proposed (Jack *et al.*, 2017). The components of the ATN model refer to the status of three different key biomarkers in AD: amyloid (A), tau (T) and neurodegeneration (N). In this approach, each of the three biomarkers can either be positive or negative. That is, exceeding the centiloid threshold in amyloid PET imaging would place a participant into the A positive (A+) group, while a scan just below the threshold would be considered amyloid negative (A-). One practical advantage (but also a major source of criticism) of this model is that the biomarker status can be assessed using a variety of methods: wet biomarkers (CSF or plasma) or brain imaging (centiloids or visual reads)(van der Flier and Scheltens, 2022). Although the progression from A-T-N- to A+T-N- to A+T+N- to A+T+N+ would be the most typical progression in AD and in keeping with the theoretical framework of biomarker progression,

all combinations of biomarker statuses emerge in observational cohorts (Ebenau *et al.*, 2020; Ingala *et al.*, 2021).

Genome-wide association studies (GWAS) have broadened the understanding of the genetic basis of late-onset AD (LOAD) over the last decades (Lambert *et al.*, 2023). Initially, these studies were restricted to cases with a clinical diagnosis of AD and healthy controls (Hollingworth *et al.*, 2011; Lambert *et al.*, 2013; Kunkle *et al.*, 2019). Recently, these GWAS have been enriched with participants with a family history of AD (DX-by-proxy) (Jansen *et al.*, 2019; Wightman *et al.*, 2021; Bellenguez *et al.*, 2022), leading to a drastic increase in sample size and expanding the set of genetic risk loci for AD to 90 (Lambert *et al.*, 2023). These loci have been linked to various molecular processes such as immunity, cholesterol processing and endocytosis (Sims *et al.*, 2020). Further studies investigated the genetic effects on AD-related biomarkers ranging from tau and amyloid levels in CSF (Deming et al., 2017, Jensen et al. 2022), in the brain (Ramanan *et al.*, 2020; Yan *et al.*, 2021) to MRI-based measures such as hippocampal volume (Hibar *et al.*, 2017) and phenotypes derived from disease progression modeling (Scelsi *et al.*, 2018). The strongest common genetic risk factor for AD is the e4 allele of the *APOE* gene: carriers of the e4 allele have a 2-4-fold increased risk to develop AD, while e4 homozygous subjects have an 8-12-fold increased risk (Belloy *et al.*, 2019). The genetic risk outside the *APOE* region is often summarized using polygenic scores, which have been shown to improve predictions of clinical diagnosis (Escott-Price *et al.*, 2015) and pathology-confirmed cases (Escott-Price *et al.*, 2017). Likewise, the effect of *APOE* as well as the polygenic risk on various imaging and non-imaging biomarkers have been investigated (Sabuncu *et al.*, 2012; Mormino *et al.*, 2016; Leonenko *et al.*, 2019; Altmann *et al.*, 2020), with ongoing work suggesting that risk accumulated along different molecular pathways exerts differential effects on different biomarkers in AD (Darst *et al.*, 2017; Leonenko *et al.*, 2019; Lorenzini *et al.*, 2023).

For AD and for other disorders, genetic risk is often considered as a time-invariant constant. That is, genetic risk identified through case-controls studies is assumed to affect both onset and progression. However, as AD is now understood to unfold in stages, we hypothesized that in line with the amyloid cascade hypothesis and the A/T/N framework that genetic risk in AD is disease-stage dependent. That is, some genetic risk factors will aid the transition from A- to A+, while other, distinct genetic risk factors will increase the risk to transition from T- to T+.

In this work we explore whether genetic vulnerability to Alzheimer's disease varies with disease stage. Using longitudinal data from the Alzheimer's disease neuroimaging imitative (ADNI) and survival analysis we show that *APOE* contributes to progression from T-A- to T-A+, but not from A+T- to A+T+. Conversely, polygenic risk contributes to the progression from A+T- to A+T+, but not from T-A- to T-A+.

Methods

Data

Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease. For up-to-date information, see www.adni-info.org. ADNI study data were accessed through the R-package ADNIMERGE (accessed: 20th July 2023).

Preparation of genetic data

The genetic data preparation followed the procedure described in Altmann et *al.* (2020). The additional genetic data contributed by the ADNI-3 cohort was integrated with the existing data using the same processing pipeline. Briefly, single nucleotide polymorphism (SNP) genotyping data were available for *n* = 2001 subjects across ADNI phases 1,2, GO and 3. Genotyping was conducted using four different platforms: Human610-Quad, HumanOmniExpress, Omni 2.5 M and Illumina Infinium Global Screening Array v2 (Illumina) (Saykin *et al.*, 2010). Prior to imputation, we applied subject-level quality control (QC) steps based on call rate (10% cutoff) and concordance between chip-inferred sex and self-reported sex separately for the four genotyping arrays. On SNP-level, we conducted standard QC steps ensuring compatibility with the reference panel used for imputation [strand consistency, allele names, position, Ref/Alt assignments and minor allele frequency discrepancy (MAF; 0.2 cutoff)]. Imputation was carried out using the Sanger Imputation Server (https://imputation.sanger.ac.uk/) with SHAPEIT for phasing (Delaneau *et al.*, 2013), Positional Burrows-Wheeler Transform (Durbin, 2014) for imputation and the entire Haplotype Reference Consortium (release 1.1) reference panel(McCarthy *et al.*, 2016). Data from the four different genotyping platforms were imputed separately. As part of post-imputation QC, multi-allelic variants and SNPs with imputation INFO score <0.3 were removed and genotype calls with posterior probability <0.9 were set to missing (i.e., hard called). Following the initial QC, genotypes from the four platforms were merged. Additional

information on the imputation and QC process is detailed in Scelsi *et al.* (2018) and https://rpubs.com/maffleur/452627. Using the merged data, we retained SNPs with MAF ≥1% and genotyping rate >0.9.

SNPweights (Chen *et al.*, 2013) was used to infer genetic ancestry from genotyped SNPs. The reference panel comprised Central European, Yoruba Africans and East Asian samples from HapMap 3 (The International HapMap 3 Consortium *et al.*, 2010) and native Americans from Reich *et al.* (2012). For this study, only participants with predicted central European ancestry of 80% or more were retained. Next, using the imputed and merged data genetic relatedness between central European participants was computed. First, the SNP content was restricted to SNPs with MAF ≥5% and LD-pruning was carried out in PLINK v1.9 (–indep-pairwise 1000 50 0.1). The genetic relatedness matrix was computed using the remaining autosomal SNPs and the dataset was trimmed to remove subjects with relatedness >0.1 (–rel-cutoff 0.1).

Definition of genetic risk

In this study, we focused on two sources of genetic risk: (1) the risk conferred through the *APOE* gene based on the genetic markers for APOE-ε2 and APOE-ε4 and (2) the polygenic risk conferred by the remaining genome. As described previously (Altmann *et al.*, 2020), PRSs were computed using the software PRSice v2.1.9 (Euesden *et al.*, 2015). As base GWAS the stage 1 results of the Alzheimer's disease GWAS featuring a clinically defined Alzheimer's disease phenotype was used (Kunkle *et al.*, 2019). For PRS computation, SNPs with MAF ≥5% were considered and SNPs were selected using LD-clumping (1000 KB, R2 of 0.1 and P-value threshold of 1.0), missing SNPs were simply ignored on subject level (using the setting–missing SET_ZERO) and the *APOE* region was excluded (hg19 coordinates chr19 from 44,400,000 to 46,500,000). For this study, we used only the P-value cutoff of 1.0e-8 to build the PRS (**Table S1**). PRSs were computed for all ADNI participants with genome-wide genotyping data. The N=441 ADNI-1 participants who contributed to the AD GWAS (Kunkle *et al.*, 2019) were excluded from the analysis to ensure independence between training and application dataset for PRS.

Definition of amyloid status

For this project we relied on two modalities to define amyloid status: Amyloid-β PET using the ^{18}F-florbetapir (FBP) and the ^{18}F-florbetaben (FBB) PET tracer, and CSF measures of amyloid-β (1-42) using the Roche Elecsys® immunoassay. We used processed data by ADNI for both modalities. Detailed information on the PET processing is available elsewhere (http://adni-info.org)(Landau *et al.*, 2013); and information on CSF amyloid-β(1-42) processing is detailed in (Bittner *et al.*, 2016; Hansson *et al.*, 2018). For CSF amyloid-β(1-42) we used the cutoff of 880 pg/mL (Hansson *et al.*, 2018) and for amyloid-β PET we used the tracer-specific SUVR (whole cerebellum reference) cutoffs of 1.11 and 1.08 for FBP and FBB (Royse *et al.*, 2021), respectively.

A participant's visit was labelled as A+ if either the PET result or the CSF result indicated a positive amyloid finding (i.e., in cases where the results were discordant, the visit would be labelled as A+). Visits with only negative amyloid findings were labeled as A- and visits without any information on amyloid (i.e., neither a PET nor a CSF result) were labelled as 'Amyloid missing'.

Definition of tau status

In keeping with our definition of amyloid positivity we used available data from CSF and PET imaging to define the tau status. More precisely, we used the Phospho-Tau(181P) CSF Roche Elecsys® immunoassay and PET imaging using the ^{18}F-Flortaucipir tracer. Details on the processing are available elsewhere (http://adni-info.org) (Bittner *et al.*, 2016; Maass *et al.*, 2017; Hansson *et al.*, 2018). For CSF we used phosphorylated 181P tau (pTau) with a cutoff of 34.61 pg/mL and for tau PET we used the cutoff of 1.42 in the meta temporal ROI comprising amygdala, entorhinal cortex, fusiform gyrus, inferior and middle temporal gyri (Jack Jr *et al.*, 2017) when normalized to the inferior cerebellar gray matter (Diedrichsen, 2006). Both cut-offs were data-driven: (1) the tau PET cutoff corresponds to a z-score of 2.0 in the cognitively normal participants in ADNI (N=506) and is close to the 'high tau' cutoff of 1.43 defined in by Jack Jr *et al.* (2023); (2) the pTau cutoff was set to maximize the Youden's index between CSF pTau and tau PET (at the 1.42 cutoff) in (N=502) ADNI participants with concurrent CSF and PET measurements. The same labeling scheme as for amyloid was applied: a visit was labelled as T+ if either the CSF or the PET indicated a positive finding, T-

if there were only negative tau findings and 'Tau missing' if neither data on CSF pTau nor on tau PET were available.

Statistical modeling

We used cox-proportional hazards models to investigate the genetic contribution of progressing from (1) A-T- to A+T- and from (2) A+T- to A+T+. For (1) we included every eligible participant with genetic data and who was A-T- based on their biomarker results as described above. This earliest A-T- visit was considered the 'start' visit (i.e., the status of previous visits was ignored). A subject was considered a converter when the available biomarkers indicated A+T- at a later visit. The time of the first A+T- biomarker finding after their initial A-T- visit was used as the conversion time. For non-converters we recorded the last visit where both amyloid and tau biomarker information was available to define the maximum follow-up time. Similarly, for (2) we included every participant with genetic data and who was A+T-. This first A+T- visit was considered the 'start' visit. A subject was considered a converter when the available biomarkers indicated A+T+ at a later visit. The time of the first A+T+ biomarker finding after their initial A+T- visit was used as the conversion time. As before, for non-converters we recorded the last visit where amyloid and tau biomarker information was available as their last point of contact.

For the analyses, the cox-proportional hazards model included age at the inclusion visit (i.e., either the A-T- or the A+T- visit), sex, and education. As variables of interest, we further included genetic variables for PRS as well as allele counts of APOE-e2 and APOE-e4. The proportional hazards assumption was tested for each covariate and for the global model using statistical tests and graphical diagnostics based on scaled Schoenfeld residuals. Statistical tests were carried out in R (v 4.1.0) using the `survival` (v 3.5.5) and `survminer` (v 0.4.9) packages.

Sensitivity analyses

In addition to the two main analyses, we conducted a series of sensitivity analyses addressing the conversion definition, biomarker cutoffs and biomarker source.

Conversion definition

To maximize the available data, we relaxed the requirement for both amyloid and tau biomarker to be available at the same visit to define conversion. As before, both biomarkers at the same time were required to define the 'start' visit as either A-T- or A+T-. However, for defining progression from A-T- or A+T-, a single A+ or T+ visit was sufficient, respectively.

Biomarker cutoffs

We varied the tau PET cutoff from 1.0 to 3.0 standard deviations above the mean in the cognitively normal ADNI participants. Notably, the lowest cutoff resulted in 1.31 and was close to the neuropathologically defined cutoff of 1.29 by Lowe *et al.* (2019). The pTau cutoffs were adjusted accordingly to maximize Youden's index between CSF pTau and tau PET.

Biomarker source

To maximize data and follow-up time, the main analyses combined data from two biomarker sources: CSF and PET. Additional sensitivity analyses relied exclusively on either CSF biomarkers or PET biomarkers. For this analysis pTau cutoffs were varied in the range from 22 to 31, covering the values of 24.25 and 29.19, which were found to indicate tau PET positivity in Braak III/IV and Braak V/VI regions, respectively (Leuzy *et al.*, 2021). Tau PET cutoffs were varied, as before, from 1.0 to 3.0 standard deviations above the mean in cognitively normal participants (i.e., in the same range as the main analysis).

Results

Out of 16,401 visits recorded in ADNI, 3789 (23.1%) visits had both amyloid and tau biomarker data available. For both survival models we identified more than 400 subjects in the ADNI database with both biomarkers available (**Table 1**): 425 individuals were A-T- of whom 68 converted to A+T-. The mean age of the A-T- group was 71.3 (6.95) years and there was an almost equal number of males and females (48.0% females). N=402 individuals were A+T- at any stage, of whom 46 converted to A+T+. The mean age of the A+T- group was 73.4 (7.1), significantly older than the A-T- group (unpaired t-test: t=3.56, p<0.001). The fraction of females in that cohort was comparable to the A-T- group (46.8% females). The distribution of APOE-e4 alleles (chi-square test; chi-square= 104.01; df=2; P=2.5e-23) and APOE-e2 alleles (fisher's exact test; P= 0.013) differed significantly between A-T- and A+T. There were more APOE-e2 carriers and fewer APOE-e4 carriers in the A-T- group than in the A+T- group. The cohort with the relaxed conversion criterion showed comparable characteristics (**Table S2**).

APOE-e4 influences progression from A-T- to A+T-

The survival analysis showed a significant contribution by APOE-e4 allele count (HR=2.71; 1.65-4.45) but no significant contribution by PRS (HR=1.06; 0.82-1.4) (**Figure 1**). The APOE-e2 allele count directionally favoured a protective effect, but this was not significant (HR=0.56; 0.22-1.4). The cox proportional hazards assumption held for this model (Global Schoenfeld test P=0.26); only sex marginally violated the proportional hazards assumption (Schoenfeld individual test; P=0.029). (**Figure S1**). This pattern of associations of *APOE* and PRS with progression was largely independent of the tau PET and pTau thresholds (**Table S3**). Furthermore, using the more relaxed conversion criteria (i.e., confirmed A+ status was sufficient instead of a confirmation of A+T-) yielded more conversions (88 instead of 68) but qualitatively the same result (**Figure S2**), i.e. a significant contribution by APOE-e4 allele count (HR=3.13; 2.05-4.79) but not by PRS (HR=1.08; 0.86-1.3).

Polygenic risk affects the progression from A+T- to A+T+

The survival analysis showed a marginally significant contribution by APOE-e4 burden (HR=1.65; 1.05-2.59), which was mainly driven by APOE-e4 homozygotes (HR=2.81; 1.14-6.91) rather than APOE-e4 heterozygotes (HR=1.5; 0.75-2.99) (**Figure 2**). Furthermore, there

was a significant contribution by PRS (HR=1.67; 1.22-2.28). The APOE-e2 allele count again was directionally consistent with a protective effect, but this was not significant (HR=0.54; 0.13-2.29). The cox proportional hazards assumption held for this model (**Figure S3**). Moreover, the association pattern of *APOE* and PRS with progression was largely independent of the tau PET and pTau cutoffs (**Table S3**). In addition, education showed a marginally protective association with conversion to A+T+ (HR=0.90; 0.82-0.99). Applying the more relaxed conversion criteria (i.e., confirmed T+ status was sufficient instead of a confirmation of A+T+) yielded more conversions (52 instead of 46) but qualitatively the same result (**Figure S4**): a significant contribution by PRS (HR=1.53; 1.15-2.00) but no significant contribution by APOE-e4 allele burden (HR=1.47; 0.97-2.23).

Results are independent of biomarker source

The main analysis combined different biomarker sources to maximize the available data and the observation time for the conversion analysis. Just relying on PET biomarkers alone, only 7.5% of visits (1237 of 16,401) had both biomarkers and it resulted in shorter observation times for A-T- to A+T- conversion (3.92 (sd=1.37) years) and A+T- to A+T+ conversion (3.58 (1.47) years) compared to the main analysis. Relying solely on CSF biomarkers, 19.2% of visits (3155 of 16,401) had both biomarkers. This led to a comparable observation time with respect to the main analysis for A-T- to A+T- conversion (6.66 (sd=3.41) years) and A+T- to A+T+ conversions (5.24 (sd=3.18) years). Furthermore, relying on a single source of biomarkers led to reduced sample sizes (from > 400 in the main analyses to 100-250 in the sensitivity analyses) and observed conversions (**Table S4** and **Table S5**). However, despite the reduced statistical power these sensitivity analyses confirmed the pattern observed in the main analysis: APOE-e4 contributed to the A-T- to A+T- conversion, while PRS contributed to the A+T- to A+T+ conversion **(Figure 3,Table S4, Table S5).**

Discussion

This longitudinal survival analysis demonstrated that APOE-e4 plays an important role for the progression from A-T- to A+T-, but *APOE* is only of marginal importance in A+T- participants who progress to A+T+. Conversely, polygenic risk for AD exhibited the inverse pattern: there was no contribution to the progression from A-T- to A+T-, but a significant contribution to faster progression from A+T- to A+T+. Notably, when assessing covariates, a differential effect of years of education was observed: higher education had no effect (**Figure 1**) or was marginally harmful (**Figure S2**) for the progression from A-T- to A+T-, but was protective for the conversion from A+T- to A+T+ (**Figure 2**). The bisection of the genetic risk by disease stage was largely independent of the applied biomarker cutoffs. Moreover, relying on only a single source of biomarkers for defining stage and conversion confirmed the findings of the main analysis despite reduced sample size and observation time.

These findings from this longitudinal analysis are in line with previous reports in the ADNI cohort of *APOE* and polygenic risk on amyloid and tau. For instance, cross-sectional amyloid biomarkers in the CSF and in the brain were mainly driven by *APOE*, while cross-sectional CSF tTau and pTau measurements were associated with PRS beyond the *APOE* locus (Altmann *et al.*, 2020). Moreover, *APOE* was found to predict amyloid status while polygenic risk for AD improved predictions of diagnosis as well as clinical progression from MCI to AD above *APOE* alone (Leonenko *et al.*, 2019). These observations extend to plasma markers of tau pathology: PRS (that excluded the *APOE* region) were found to be only associated with plasma p-tau181 in A+ participants (Zettergren *et al.*, 2021). This association between polygenic risk (beyond *APOE)* and CSF tau biomarkers rather than with amyloid and neurodegeneration was also observed outside of the ADNI study (Kumar *et al.*, 2022). Polygenic risk (beyond *APOE*) was associated with non-amyloid endophenotypes in a large cohort of people with MCI. This suggests that these variants are more closely linked with neuronal degeneration than with AD-related amyloid pathology (Louwersheimer *et al.*, 2016). All these previous studies made the connection between existing amyloid pathology and correlations between polygenic risk and tau pathology using cross-sectional study designs. In a recent longitudinal study of tau PET in the ADNI cohort, higher polygenic risk for AD was associated with accelerated increase in tau signal in brain and this effect was modulated by amyloid pathology: A+ participants showed a stronger effect of PRS on tau accumulation (Rubinski *et al.*, 2023). Our longitudinal analysis, which combined CSF and PET

data to maximize the sample size, confirms these observations and indicates that polygenic AD risk (outside the *APOE* region) contributes to tau pathology in A+ participants, but has no meaningful contribution in A- participants. Moreover, recent data from animal models in Alzheimer's disease suggest a state-dependent effect of genetic risk factors related to microglia. For instance, deletion of *Trem2* in mouse models exacerbated tau accumulation and spreading leading to brain atrophy, but only in the presence of existing β-amyloid pathology (Lee *et al.*, 2021). Along the same line, physical contact between microglia and plaques as well as a functioning *TREM2* gene are necessary for the appropriate microglia response to amyloid pathology (Wood *et al.*, 2022). Thus, defects in *TREM2* can only contribute to neurodegeneration once amyloid pathology has been established. Our longitudinal analysis presented here is the first to support such a state-dependent genetic risk model in humans.

The partition of AD genetic risk into *APOE*-related and polygenic risk beyond *APOE* is a simplification in this analysis. Recent works have linked established AD risk loci outside *APOE*, such as *CR1*, to amyloid biomarker levels (Jansen *et al.*, 2022). Conversely, studies of biomarker levels of tau repeatedly highlight the *APOE* locus (Deming *et al.*, 2017; Jansen *et al.*, 2022). However, if being A+ were a pre-requisite to exhibit pathological accumulation of tau, then the strong genetic association with *APOE* in these GWASs would merely reflect the necessary condition rather than a genuine direct molecular process that affects tau levels. The strong dependency of tau levels on established amyloid pathology is supported by mediation analyses in recent cross-sectional (Kumar *et al.*, 2022) and longitudinal (Rubinski *et al.*, 2023) studies. Moreover, the known AD genetic risk variants are likely to contribute differently to molecular pathways (Sims *et al.*, 2020), where each pathway in turn will exercise differential effects on the AD biomarkers including markers for vascular pathology (Lorenzini *et al.*, 2023).

In addition to the cascading effect of genetic risk in AD, we also observed a stage-dependent effect of non-genetic risk factors. Education has been shown to have a protective effect against dementia (Livingston *et al.*, 2020): here we show that higher rates of education do not influence transition to amyloid positivity, but do slow progression from A+T- to A+T+. Consequently, other non-genetic risk factors may show a similarly state-dependent effect on the pathological pathway from A-T- to A+T+ and further neurodegeneration.

The study has several limitations. Firstly, the current analysis was limited to just two biomarkers in Alzheimer's disease: β-amyloid and tau. It would be desirable to include neurodegeneration (N; of the ATN framework) or potentially more fine-grade staging from advanced data-driven disease progression modeling (Young *et al.*, 2014). However, at this point adding further stages would reduce the available sample size. Secondly, we partitioned the genetic risk in Alzheimer's disease into two components: *APOE* and other top variants combined into a single polygenic risk score. Further work should explore a more fine-grained partition of the polygenic risk into individual SNPs or into pathway-PRS (Lorenzini *et al.*, 2023). Finally, although the ADNI cohort is a large cohort, the number of subjects who were eligible for our analysis was reduced due to the requirement of concordant and longitudinal recordings of multiple biomarkers as well as genetics. Thus, analyses in further longitudinal cohorts are required to confirm the observation of stage-dependent genetic vulnerability in Alzheimer's disease.

In this work we demonstrated, in a simplified setting, that genetic risk for late onset Alzheimer's disease unfolds in a disease stage-dependent fashion. A better understanding of the interplay between disease stage and genetic risk can lead to a better understanding of the molecular processes leading to AD as well as opening therapeutic windows for targeted interventions and personalized approaches to dementia prevention.

Data Availability Statement

Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu) and are freely available after registration.

Acknowledgements

This study was supported by the Early Detection of Alzheimer's Disease Subtypes (E-DADS) project, an EU Joint Programme - Neurodegenerative Disease Research (JPND) project (see www.jpnd.eu). The project is supported under the aegis of JPND through the following funding organizations: United Kingdom, Medical Research Council (MR/T046422/1); Netherlands, ZonMW (733051106); France, Agence Nationale de la Recherche (ANR-19-JPW2–000); Italy, Italian Ministry of Health (MoH); Australia, National Health & Medical Research Council (1191535); Hungary, National Research, Development and Innovation Office (2019–2.1.7-ERA-NET-2020–00008). N.P.O. is a UKRI Future Leaders Fellow (MR/S03546X/1). JMS acknowledges the support of the National Institute for Health Research University College London Hospitals Biomedical Research Centre, Wolfson Foundation, Alzheimer's Research UK, Brain Research UK, Weston Brain Institute, Medical Research Council, British Heart Foundation, and Alzheimer's Association. This research was funded in whole, or in part, by the Wellcome Trust [227341/Z/23/Z]. For the purpose of open access, the author has applied a CC BY-ND public copyright license to any Author Accepted Manuscript version arising from this submission.

References

Altmann A, Scelsi MA, Shoai M, de Silva E, Aksman LM, Cash DM, et al. A comprehensive analysis of methods for assessing polygenic burden on Alzheimer's disease pathology and risk beyond APOE. Brain Commun 2020; 2: fcz047.

Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012; 367: 795–804.

Bellenguez C, Küçükali F, Jansen IE, Kleineidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat Genet 2022; 54: 412–436.

Belloy ME, Napolioni V, Greicius MD. A Quarter Century of APOE and Alzheimer's Disease: Progress to Date and the Path Forward. Neuron 2019; 101: 820–838.

Bittner T, Zetterberg H, Teunissen CE, Ostlund Jr RE, Militello M, Andreasson U, et al. Technical performance of a novel, fully automated electrochemiluminescence immunoassay for the quantitation of $β$-amyloid (1--42) in human cerebrospinal fluid. Alzheimer's \& Dement 2016; 12: 517–526.

Braak H, de Vos RAI, Jansen ENH, Bratzke H, Braak E. Neuropathological hallmarks of Alzheimer's and Parkinson's diseases. Prog Brain Res 1998; 117: 267–285.

Chen C-Y, Pollack S, Hunter DJ, Hirschhorn JN, Kraft P, Price AL. Improved ancestry inference using weights from external reference panels. Bioinformatics 2013; 29: 1399–1406.

Darst BF, Koscik RL, Racine AM, Oh JM, Krause RA, Carlsson CM, et al. Pathway-specific polygenic risk scores as predictors of amyloid-$β$ deposition and cognitive function in a sample at increased risk for Alzheimer's disease. J Alzheimer's Dis 2017; 55: 473–484.

Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 2013; 10: 5–6.

Deming Y, Li Z, Kapoor M, Harari O, Del-Aguila JL, Black K, et al. Genome-wide association study identifies four novel loci associated with Alzheimer's endophenotypes and disease modifiers. Acta Neuropathol 2017; 133: 839–856.

Diedrichsen J. A spatially unbiased atlas template of the human cerebellum. Neuroimage

2006; 33: 127–138.

Donohue MC, Jacqmin-Gadda H, Le Goff M, Thomas RG, Raman R, Gamst AC, et al. Estimating long-term multivariate progression from short-term data. Alzheimer's & Dement J Alzheimer's Assoc 2014; 10: S400-10.

Durbin R. Efficient haplotype matching and storage using the positional Burrows-Wheeler transform (PBWT). Bioinformatics 2014; 30: 1266–1272.

Ebenau JL, Timmers T, Wesselman LMP, Verberk IMW, Verfaillie SCJ, Slot RER, et al. ATN classification and clinical progression in subjective cognitive decline: The SCIENCe project. Neurology 2020; 95: e46--e58.

Escott-Price V, Myers AJ, Huentelman M, Hardy J. Polygenic risk score analysis of pathologically confirmed alzheimer disease [Internet]. Ann Neurol 2017Available from: http://onlinelibrary.wiley.com/doi/10.1002/ana.24999/abstract

Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain 2015; 138: 3673–84.

Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2015; 31: 1466–8.

van der Flier WM, Scheltens P. The ATN framework—moving preclinical Alzheimer disease to clinical relevance. JAMA Neurol 2022; 79: 968–970.

Fonteijn HM, Modat M, Clarkson MJ, Barnes J, Lehmann M, Hobbs NZ, et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. Neuroimage 2012; 60: 1880–1889.

Hansson O, Seibyl J, Stomrud E, Zetterberg H, Trojanowski JQ, Bittner T, et al. CSF biomarkers of Alzheimer's disease concord with amyloid-$β$ PET and predict clinical progression: a study of fully automated immunoassays in BioFINDER and ADNI cohorts. Alzheimer's \& Dement 2018; 14: 1470–1481.

Hibar DP, Adams HHH, Jahanshad N, Chauhan G, Stein JL, Hofer E, et al. Novel genetic loci associated with hippocampal volume. Nat Commun 2017; 8: 13624.

Hollingworth P, Harold D, Sims R, Gerrish A, Lambert J-C, Carrasquillo MM, et al. Common

variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nat Genet 2011; 43: 429–435.

Ingala S, De Boer C, Masselink LA, Vergari I, Lorenzini L, Blennow K, et al. Application of the ATN classification scheme in a population without dementia: Findings from the EPAD cohort. Alzheimer's \& Dement 2021; 17: 1189–1204.

Jack CRJ, Bennett DA, Blennow K, Carrillo MC, Dunn B. NIA-AA research framework: towards a biological definition of Alzheimer's disease [Internet]. 2017Available from: https://www.alz.org/aaic/_downloads/nia-aa-draft-11-27-2017.pdf

Jack Jr CR, Knopman DS, Jagust WJ, Petersen RC, Weiner MW, Aisen PS, et al. Update on hypothetical model of Alzheimer's disease biomarkers. Lancet Neurol 2013; 12: 207.

Jack Jr CR, Wiste HJ, Algeciras-Schimnich A, Figdore DJ, Schwarz CG, Lowe VJ, et al. Predicting amyloid PET and tau PET stages with plasma biomarkers. Brain 2023; 146: 2029–2044.

Jack Jr CR, Wiste HJ, Weigand SD, Therneau TM, Lowe VJ, Knopman DS, et al. Defining imaging biomarker cut points for brain aging and Alzheimer's disease. Alzheimer's \& Dement 2017; 13: 205–216.

Jansen IE, van der Lee SJ, Gomez-Fonseca D, de Rojas I, Dalmasso MC, Grenier-Boley B, et al. Genome-wide meta-analysis for Alzheimer's disease cerebrospinal fluid biomarkers. Acta Neuropathol 2022; 144: 821–842.

Jansen IE, Savage JE, Watanabe K, Bryois J, Williams DM, Steinberg S, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. Nat Genet 2019; 51: 404–413.

Jedynak BM, Lang A, Liu B, Katz E, Zhang Y, Wyman BT, et al. A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease neuroimaging initiative cohort. Neuroimage 2012; 63: 1478–1486.

Kumar A, Janelidze S, Stomrud E, Palmqvist S, Hansson O, Mattsson-Carlgren N, et al. $β$-amyloid--dependent and--independent genetic pathways regulating CSF tau biomarkers in Alzheimer disease. Neurology 2022; 99: e476--e487.

Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and

lipid processing. Nat Genet 2019; 51: 414.

Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45: 1452.

Lambert J-C, Ramirez A, Grenier-Boley B, Bellenguez C. Step by step: towards a better understanding of the genetic architecture of Alzheimer's disease. Mol Psychiatry 2023: 1–12.

Landau SM, Breault C, Joshi AD, Pontecorvo M, Mathis CA, Jagust WJ, et al. Amyloid-β Imaging with Pittsburgh Compound B and Florbetapir: Comparing Radiotracers and Quantification Methods. J Nucl Med 2013; 54: 70–77.

Lee S-H, Meilandt WJ, Xie L, Gandham VD, Ngu H, Barck KH, et al. Trem2 restrains the enhancement of tau accumulation and neurodegeneration by $β$-amyloid pathology. Neuron 2021; 109: 1283–1301.

Leonenko G, Shoai M, Bellou E, Sims R, Williams J, Hardy J, et al. Genetic risk for Alzheimer's disease is distinct from genetic risk for amyloid deposition. Ann Neurol 2019

Leuzy A, Janelidze S, Mattsson-Carlgren N, Palmqvist S, Jacobs D, Cicognola C, et al. Comparing the clinical utility and diagnostic performance of CSF P-Tau181, P-Tau217, and P-Tau231 assays. Neurology 2021; 97: e1681--e1694.

Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020; 396: 413–446.

Lorenzi M, Filippone M, Frisoni GB, Alexander DC, Ourselin S, Initiative ADN, et al. Probabilistic disease progression modeling to characterize diagnostic uncertainty: application to staging and prediction in Alzheimer's disease. Neuroimage 2019; 190: 56–68.

Lorenzini L, Collij LE, Tesi N, Ingala S, Sudre CH, Wolz R, et al. APOE-dependent and-independent polygenic pathways determining early Alzheimer's Disease pathological changes in CSF and MRI. In: Alzheimer's Association International Conference. 2023

Louwersheimer E, Wolfsgruber S, Espinosa A, Lacour A, Heilmann-Heimbach S, Alegret M, et al. Alzheimer's disease risk variants modulate endophenotypes in mild cognitive

impairment. Alzheimer's \& Dement 2016; 12: 872–881.

Lowe VJ, Lundt ES, Albertson SM, Min H-K, Fang P, Przybelski SA, et al. Tau-positron emission tomography correlates with neuropathology findings. Alzheimer's \& Dement 2019

Maass A, Landau S, Baker SL, Horng A, Lockhart SN, La Joie R, et al. Comparison of multiple tau-PET measures as biomarkers in aging and Alzheimer's disease. Neuroimage 2017; 157: 448–463.

McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 2016; 48: 1279–1283.

Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. Neurology 2016; 87: 481–488.

Ossenkoppele R, Schonhaut DR, Schöll M, Lockhart SN, Ayakta N, Baker SL, et al. Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. Brain 2016; 139: 1551–1567.

Ramanan VK, Wang XK, Przybelski SA, Pillai SKR, Kosel ML, Batzler AK, et al. Genome-wide association study of tau-PET: Association of variant near NTNG2 with resistance to tau deposition: Neuroimaging: Imaging genetics. Alzheimer's \& Dement 2020; 16: e044321.

Reich D, Patterson N, Campbell D, Tandon A, Mazieres S, Ray N, et al. Reconstructing Native American population history. Nature 2012; 488: 370–374.

Royse SK, Minhas DS, Lopresti BJ, Murphy A, Ward T, Koeppe RA, et al. Validation of amyloid PET positivity thresholds in centiloids: a multisite PET study approach. Alzheimer's Res \& Ther 2021; 13: 1–10.

Rubinski A, Frerich S, Malik R, Franzmeier N, Ramirez A, Dichgans M, et al. Polygenic Effect on Tau Pathology Progression in Alzheimer's Disease. Ann Neurol 2023; 93: 819–829.

Sabuncu MR, Buckner RL, Smoller JW, Lee PH, Fischl B, Sperling RA, et al. The association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. Cereb cortex (New York, NY 1991) 2012; 22: 2653–2661.

Saykin AJ, Shen L, Foroud TM, Potkin SG, Swaminathan S, Kim S, et al. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims,

progress, and plans. Alzheimer's Dement 2010; 6: 265–73.

Scelsi MA, Khan RR, Lorenzi M, Christopher L, Greicius MD, Schott JM, et al. Genetic study of multimodal imaging Alzheimer's disease progression score implicates novel loci. Brain 2018; 141: 2167–2180.

Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25 years. EMBO Mol Med 2016; 8: 595–608.

Sims R, Hill M, Williams J. The multiplex model of the genetics of Alzheimer's disease. Nat Neurosci 2020; 23: 311–322.

The International HapMap 3 Consortium, Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, et al. Integrating common and rare genetic variation in diverse human populations. Nature 2010; 467: 52–58.

Wightman DP, Jansen IE, Savage JE, Shadrin AA, Bahrami S, Holland D, et al. A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. Nat Genet 2021; 53: 1276–1282.

Wood JI, Wong E, Joghee R, Balbaa A, Vitanova KS, Stringer KM, et al. Plaque contact and unimpaired Trem2 is required for the microglial response to amyloid pathology. Cell Rep 2022; 41

Yan Q, Nho K, Del-Aguila JL, Wang X, Risacher SL, Fan K-H, et al. Genome-wide association study of brain amyloid deposition as measured by Pittsburgh Compound-B (PiB)-PET imaging. Mol Psychiatry 2021; 26: 309–321.

Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, et al. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 2014; 137: 2564–2577.

Zettergren A, Lord J, Ashton NJ, Benedet AL, Karikari TK, Lantero Rodriguez J, et al. Association between polygenic risk score of Alzheimer's disease and plasma phosphorylated tau in individuals from the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's Res \& Ther 2021; 13: 1–10.

Figures

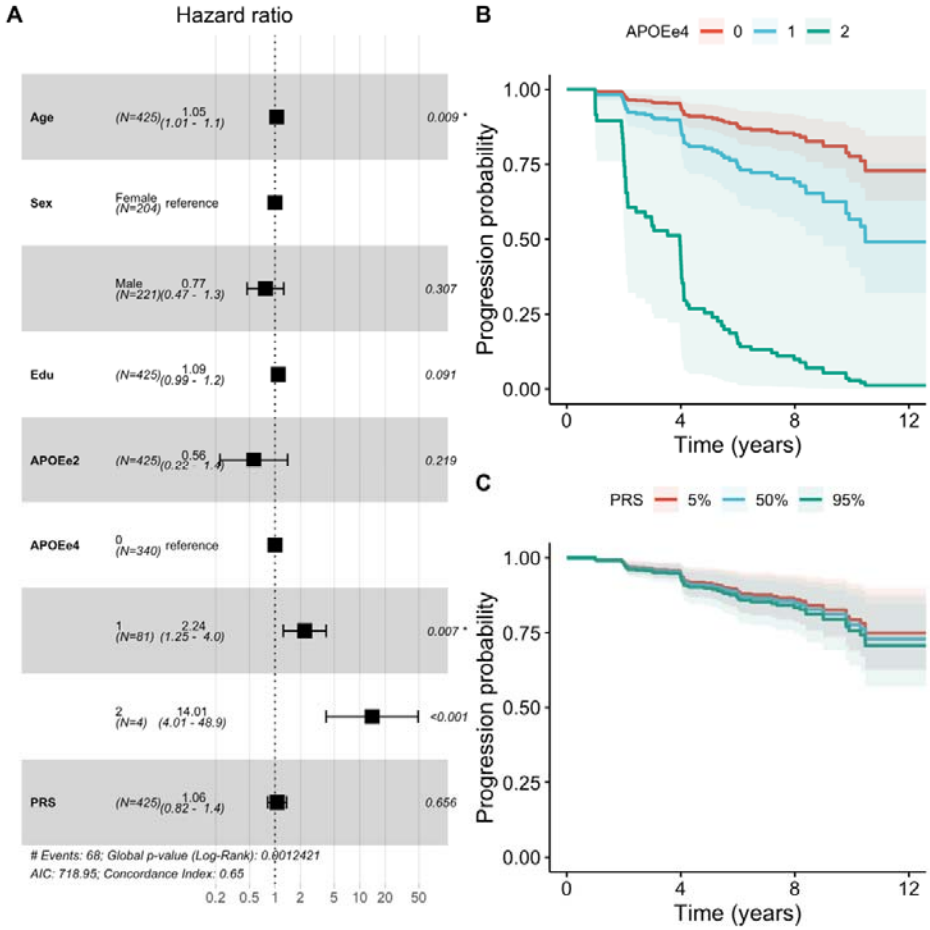

Figure 1: Hazard ratios for the conversion from A-T- to A+T-. (A) Forest plot depicting the Hazards Ratios (HR) for all covariates in the model. (B) Estimated survival curves stratified by APOEe4 genotype. (C) Estimated survival curves stratified by PRS percentile (5%, 50%, 95%). Edu = Years of Educations; APOEe2 = number of APOE2 alleles, APOEe4 = number of APOE4 alleles; PRS = polygenic risk score, scaled to zero mean and unit standard deviation.

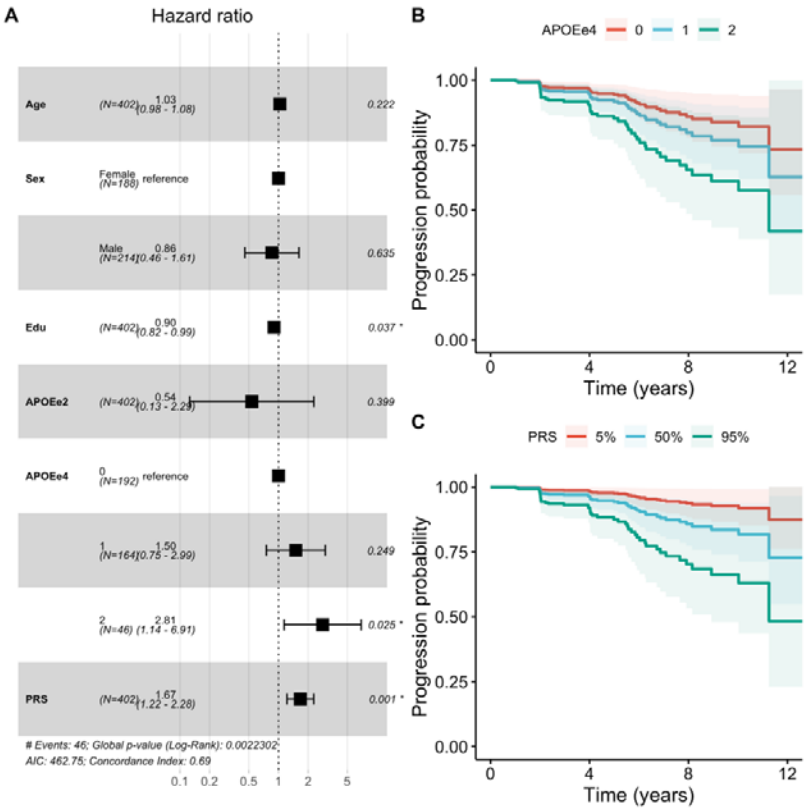

Figure 2: Hazard ratios for the conversion from A+T- to A+T+. (A) Forest plot depicting the Hazards Ratios (HR) for all covariates in the model. (B) Estimated survival curves stratified by APOEe4 genotype. (C) Estimated survival curves stratified by PRS percentile (5%, 50%, 95%). Edu = Years of Educations; APOEe2 = number of APOE2 alleles, APOEe4 = number of APOE4 alleles; PRS = polygenic risk score, scaled to zero mean and unit standard deviation.

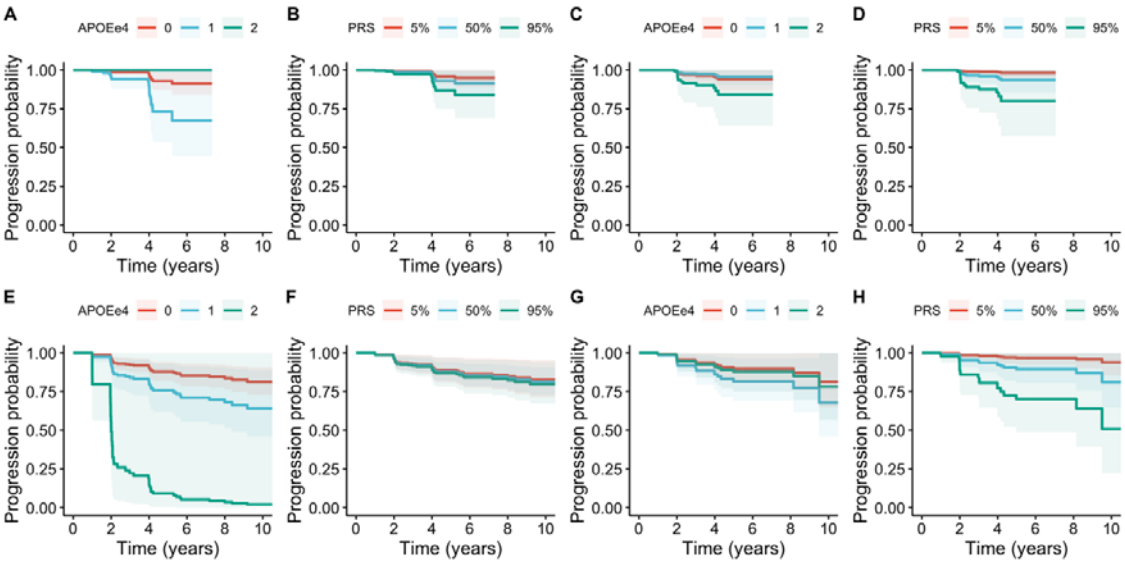

Figure 3: Survival curves for PET only and CSF only analyses. The top row (A-D) is based on results from AT(N) definitions based exclusively on PET biomarkers (amyloid and tau) with a cutoff of 1.45 for tau PET; the bottom row (E-H) relied on CSF biomarkers (ABETA42 and pTAU) with a cutoff of 26 for pTau. The two left columns (A, B, E, F) display the progression probability from A-T- to A+T- stratified by APOEe4 genotype and PRS quantile (5%, 50%, 95%). The two right columns (C, D, G, H) display the progression probability from A+T- to A+T+.

Tables

Table 1: Demographics

	A-T-			A+T-		
	Total	**Stable**	**Converter**	**Total**	**Stable**	**Converter**
N	425	357	68	402	356	46
Age, y (sd)	71.3 (6.95)	71.0 (6.81)	72.2 (7.30)	73.4 (7.1)	73.4 (7.2)	73.1 (6.8)
Sex (%female)	48.0	47.0	52.9	46.8	46.6	47.8
DX (CN/MCI/AD)	240/179/6	194/158/5	46/21/1	172/205/25	154/179/26	18/26/2
Education, y (sd)	16.7 (2.4)	16.6 (2.5)	17.1 (2.3)	16.5 (2.6)	16.6 (2.6)	15.5 (2.3)
Years Follow-up, y (sd)	6.1 (3.3)	5.7 (3.1)	8.1 (3.5)	5.1 (3.0)	4.8 (2.9)	7.2 (2.8)
Time to event, y (sd)	n/a	n/a	4.5 (2.7)	n/a	n/a	4.5 (2.5)
APOE-e4 (0/1/2)	340/81/4	293/63/1	47/18/3	192/164/46	176/143/37	16/21/9
APOE-e2 (0/1/2)	370/54/1	307/49/1	63/5/0	372/30/0	328/28/0	44/2/0
PRS (sd)	0.013 (0.012)	0.013 (0.012)	0.013 (0.011)	0.014 (0.013)	0.013 (0.013)	0.019 (0.012)

DX = diagnosis (cognitively normal [CN]; mild cognitive impairment [MCI]; Alzheimer's Disease (AD); PRS = Polygenic Risk Score