ARTICLE IN PRESS JID: TJPAD (October 8, 2026)

The Journal of Prevention of Alzheimer's Disease xxx (xxxx) xxx

Contents lists available at **ScienceDirect**

The Journal of Prevention of Alzheimer's Disease

journal homepage: www.elsevier.com/locate/tjpad

Original Article

External validation and causal mediation of multimodal prognostic models for progression from mild cognitive impairment to dementia across 5,078 subjects in OASIS-3 and ADNI

Erick Flores Zambrano[1,2,*], **for the Alzheimer's Disease Neuroimaging Initiative**[†]

[1] *Facultad de Ciencias Económicas, Universidad Técnica de Manabí (UTM), Portoviejo, Ecuador*

[2] *Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil (UG), Guayaquil, Ecuador*

ARTICLE INFO

Keywords:
Alzheimer's disease
Mild cognitive impairment
Amyloid PET Centiloids
Causal mediation
Plasma p-tau217
sTREM2

Survival analysis

Article history:
Received 10 October 2026
Received in revised form 15 October 2026

Accepted 18 October 2026

ABSTRACT

Background: Multimodal prognostic models predicting progression from mild cognitive impairment (MCI) to Alzheimer's disease (AD) dementia frequently suffer from optimistic performance bias due to internal overfitting, retrospective data concatenation, and absence of external validation across independent healthcare systems. Furthermore, the causal proportion of APOE ε4-driven cognitive decline mediated through fibrillar amyloid accumulation versus amyloid-independent pathways remains unresolved.

Objectives: To evaluate out-of-cohort survival discrimination of fixed prognostic weights derived in OASIS-3 when applied to ADNI, quantify the causal proportion mediated by amyloid PET, test Hypothesis H3 regarding microglial CSF sTREM2, and evaluate the prospective prognostic equivalence of blood-based plasma p-tau217.

Design: Prospective, multi-cohort observational survival and causal mediation study.

Setting: Charles F. and Joanne Knight ADRC (OASIS-3, discovery) and 50+ multicenter consortia across the United States and Canada (ADNI, external validation).

Participants: Harmonized cohort of $N = 5,078$ total individuals ($N = 1,378$ OASIS-3; $N = 3,700$ ADNI). Longitudinal conversion analyses focused on verified baseline MCI patients: $N = 140$ in OASIS-3 (88 dementia conversions) and $N = 1,231$ in ADNI (414 conversions; multivariable complete-case $N = 760$ with 190 events).

Measurements: Amyloid PET Centiloids ($\Delta t \leq 365$ days), FreeSurfer 7 normalized hippocampal volume, APOE genotypes, CSF sTREM2 concentrations, and Fujirebio Lumipulse plasma p-tau217.

Results: Applying fixed OASIS-3 weights directly to ADNI yielded an out-of-sample Harrell's C-index of **0.763** (95% CI: 0.738–0.788), replicating discrimination without parameter re-tuning ($C = 0.749$ vs 0.763). In multivariable Cox models in ADNI ($N = 760$), Centiloids conferred an adjusted HR = 1.016 per CL (95% CI: 1.012–1.019, $p = 2.20 \times 10^{-20}$), normalized hippocampus conferred HR = 0.686 per scaled cm^3 (95% CI: 0.587–0.800, $p = 1.79 \times 10^{-6}$), and APOE ε4 conferred HR = 1.416 (95% CI: 1.027–1.951, $p = 0.0337$). Counterfactual causal mediation confirmed that **75.0%** (95% CI: 59.6%–96.5%, $p < 0.0001$) of the APOE ε4 cognitive detriment operated through fibrillar amyloid burden, reproducing the 75.1% proportion observed in OASIS-3. CSF sTREM2 did not predict dementia conversion in clinical MCI (HR = 0.952, $p = 0.513$; log-rank $p = 0.542$). Plasma p-tau217 correlated with PET Centiloids (Spearman $r = 0.669$, $p < 10^{-100}$) and retained **98.2%** of PET's discrimination ($C = 0.731$ vs 0.745).

Conclusions: Multimodal prognostic models generalize across distinct research consortia without performance degradation. Fibrillar amyloid deposition accounts for three-fourths of APOE ε4-driven cognitive decline. Soluble TREM2 is decoupled from clinical progression once MCI is established. Blood-based plasma p-tau217 demonstrates near-equivalent prognostic discrimination to amyloid PET, providing a non-invasive screening instrument for clinical trials.

1 Introduction

Mild cognitive impairment (MCI) represents the pivotal transitional clinical stage between normal cognitive aging and Alzheimer's disease (AD) dementia [1]. With the recent regulatory approvals of disease-modifying anti-amyloid monoclonal antibodies, such as lecanemab and donanemab [2, 3], accurate risk stratification in patients with amnestic or multidomain MCI has emerged as a ma-

* Corresponding author at: Facultad de Ciencias Económicas, Universidad Técnica de Manabí, Portoviejo, Ecuador. *E-mail:* eflores4006@utm.edu.ec (E. Flores Zambrano).

† 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 analysis or writing of this report. Complete listing: adni.loni.usc.edu/data-samples/access-data/

jor clinical priority. Identifying individuals at high risk of rapid conversion to dementia within 2 to 3 years enables timely clinical prioritization before extensive neocortical synaptic degeneration occurs.

Despite numerous published predictive models claiming high diagnostic accuracy in cross-sectional benchmarks, translation to routine practice remains limited by several methodological vulnerabilities [4, 5]. First, prognostic algorithms are frequently trained and evaluated within single-institution cohorts, where internal optimism bias and site-specific imaging protocols can lead to substantial performance degradation when evaluated in unseen external populations. Second, models often evaluate cross-sectional classification rather than prospective, time-to-event survival modeling that accounts for right-censoring. Third, the potential for temporal mismatch artifacts arises when neuroimaging examinations acquired years apart from baseline clinical evaluations are concatenated without explicit time-window constraints.

Beyond prognostic modeling, key questions regarding the underlying causal architecture of AD progression remain unresolved:

1. **Causal mechanism of APOE $\varepsilon 4$:** While the apolipoprotein E (APOE) $\varepsilon 4$ allele is the established major genetic risk factor for sporadic AD [6], the exact quantitative proportion of its cognitive detriment mediated through fibrillar amyloid-β plaque burden versus amyloid-independent pathways (such as tau propagation, neuroinflammation, or cerebrovascular dysfunction) has rarely been estimated across harmonized multi-cohort populations.

2. **Microglial activation in clinical MCI:** Soluble triggering receptor expressed on myeloid cells 2 (sTREM2) in cerebrospinal fluid (CSF) has been linked to protective microglial plaque compaction during early preclinical stages [7, 8]. However, whether elevated sTREM2 confers sustained clinical protection once individuals have progressed to symptomatic MCI remains contentious.

3. **Blood-based biomarker equivalence:** While Positron Emission Tomography (PET) Centiloid quantification serves as the standard for in-vivo fibrillar amyloid estimation [9], its accessibility is restricted by high cost and radiation exposure. Immunoassays for plasma phosphorylated tau-217 (p-tau217) exhibit high diagnostic sensitivity [10], yet their longitudinal prognostic parity with PET Centiloids in time-to-conversion modeling requires rigorous multi-cohort validation.

To address these questions, we harmonized individual participant data from 5,078 individuals across two multicenter resources: the Open Access Series of Imaging Studies (OASIS-3, discovery cohort, $N = 1,378$) and the Alzheimer's Disease Neuroimaging Initiative (ADNI, external validation cohort, $N = 3,700$). We evaluated out-of-cohort survival discrimination without model recalibration, replicated counterfactual causal mediation across 1,000 bootstrap resamples, tested pre-registered Hypothesis H3 regarding CSF sTREM2, and assessed blood-based plasma p-tau217 against molecular PET imaging.

2 Methods

2.1 Study Populations and Design

This study was reported according to the TRIPOD-AI and STROBE guidelines. Data were accessed under approved Data Use Agreements from two independent repositories:

- **Discovery Cohort (OASIS-3):** Evaluated at the Charles F. and Joanne Knight Alzheimer Disease Research Center (Knight ADRC), Washington University in St. Louis [11]. The baseline cohort includes $N = 1,378$ unique subjects evaluated longitudinally between 2000 and 2022. For prospective conversion modeling, we identified participants diagnosed with MCI at baseline (Clinical Dementia Rating global score CDR = 0.5). Of 246 baseline MCI individuals in the Uniform Data Set (UDS), $N = 140$ had longitudinal follow-up with documented clinical consensus, recording 88 confirmed dementia conversion events (CDR ≥ 1.0) over up to 13.5 years.

- **External Validation Cohort (ADNI):** Assembled from the ADNI database (adni.loni.usc.edu), encompassing ADNI-1, ADNI-GO, ADNI-2, and ADNI-3 across more than 50 academic centers in the United States and Canada (LONI DSPC approval granted to E.F.Z.). The baseline cohort comprises $N = 3,700$ subjects. For prospective time-to-conversion analysis, we isolated $N = 1,231$ baseline MCI individuals with longitudinal follow-up, observing 414 confirmed conversion events over a median follow-up of 26.7 months (interquartile range: 12.1–48.5 months; maximum 187.2 months).

2.2 Multimodal Harmonization and Temporal Alignment

To prevent temporal misalignment and data leakage, all imaging and fluid biomarker evaluations were matched to baseline clinical consensus diagnosis within a strict window of $\Delta t \leq 365$ days:

- **Amyloid PET Centiloids:** In OASIS-3, amyloid burden was quantified from ^{11}C-PiB or ^{18}F-AV45 scans converted to universal Centiloids (CL) [9]. In ADNI, standard Centiloids and pathological status (≥ 24.0 CL) were obtained from the UC Berkeley summary collections across ^{11}C-PiB, ^{18}F-Florbetapir, and ^{18}F-Florbetaben tracers.

- **Structural MRI Volumetry:** T1-weighted MPRAGE scans were processed through FreeSurfer version 7.0. Total bilateral hippocampal volume was

computed as left plus right structures and normalized to total intracranial volume (ICV) as:

$$\text{Hippo}_{\text{norm}} = \left(\frac{\text{Total Hippocampus}}{\text{ICV}} \right) \times 10^3 \quad (1)$$

- **Genomics:** APOE genotypes ($\varepsilon 2, \varepsilon 3, \varepsilon 4$) were determined via clinical PCR assays and parameterized as carrier status ($\varepsilon 4+$ vs $\varepsilon 4-$) and allelic dosage (0, 1, 2 copies).
- **CSF Soluble TREM2:** Concentrations of sTREM2 (pg/mL) in CSF were quantified by the Haass Laboratory and Washington University ADRC Core via validated enzyme-linked immunosorbent assays.
- **Plasma p-tau217:** Plasma phosphorylated tau-217 (pg/mL) was quantified on the automated Fujirebio Lumipulse platform.

2.3 Statistical and Survival Modeling

All analyses were conducted in Python 3.12 using `lifelines`, `scikit-learn`, and `statsmodels`. Time-to-event was defined as duration from baseline MCI diagnosis to confirmed dementia conversion or right-censoring at the last visit. Kaplan-Meier survival estimators and two-sided log-rank tests evaluated univariable differences. Multivariable survival modeling was implemented via Cox Proportional Hazards regression. Proportional hazards assumptions were evaluated via Schoenfeld partial residuals against transformed time ranks.

Pure out-of-cohort external validation was conducted by applying the fixed regression coefficients derived in the OASIS-3 discovery cohort:

$$\text{Score} = 0.0116 \cdot \text{CL} - 0.631 \cdot \text{Hippo}_{\text{norm}} + 0.043 \cdot \text{Age} + 0.301 \cdot \text{APO} \quad (2)$$

directly to individuals in the ADNI MCI cohort, assessing discrimination using Harrell's Concordance Index (C-index).

2.4 Counterfactual Causal Mediation Analysis

To estimate the proportion of APOE $\varepsilon 4$-driven cognitive decline funneled through fibrillar amyloid accumulation, we employed the Preacher-Hayes non-parametric bootstrapping framework with 1,000 iterations:

$$\text{CL}_i = \alpha_0 + \alpha_1 \cdot \text{APOE4}_i + \mathbf{X}_i \boldsymbol{\gamma} + \varepsilon_{1i} \quad (3)$$

$$\Delta\text{Cog}_i = \beta_0 + \beta_1 \cdot \text{APOE4}_i + \beta_2 \cdot \text{CL}_i + \mathbf{X}_i \boldsymbol{\delta} + \varepsilon_{2i} \quad (4)$$

The Average Causal Mediation Effect (ACME), Average Direct Effect (ADE), and Proportion Mediated (ACME/[ACME+ADE]) were calculated with 95% bias-corrected confidence intervals.

2.5 Decision Curve Analysis (DCA)

Net clinical benefit was evaluated across decision threshold probabilities ($p_t \in [0.05, 0.60]$) at a 3-year post-diagnosis horizon:

$$\text{Net Benefit} = \frac{\text{TP}}{N} - \frac{\text{FP}}{N} \left(\frac{p_t}{1 - p_t} \right) \quad (5)$$

and benchmarked against "Treat All" and "Treat None" strategies.

3 Results

3.1 Baseline Cohort Characteristics

Demographic and multimodal characteristics are summarized in Table 1. Mean age was 71.0 ± 9.0 in OASIS-3 and 72.0 ± 7.7 in ADNI. APOE $\varepsilon 4$ prevalence was 39.5% in OASIS-3 and 45.4% in ADNI. In the longitudinal MCI subcohorts, 88 of 140 (62.9%) progressed to dementia in OASIS-3 over longitudinal follow-up compared to 414 of 1,231 (33.6%) in ADNI.

3.2 External Validation of Prognostic Survival Models

In the external ADNI MCI cohort ($N = 1,231$), amyloid PET positivity ($\geq$ 24.0 CL) stratified progression risk (two-sided log-rank $\chi^2 = 127.61, p = 1.36 \times 10^{-29}$; Figure 1). Median time to dementia was 4.32 years in amyloid-positive individuals, whereas amyloid-negative individuals did not reach median survival during observation (5-year conversion rate: 64.2% vs 19.8%).

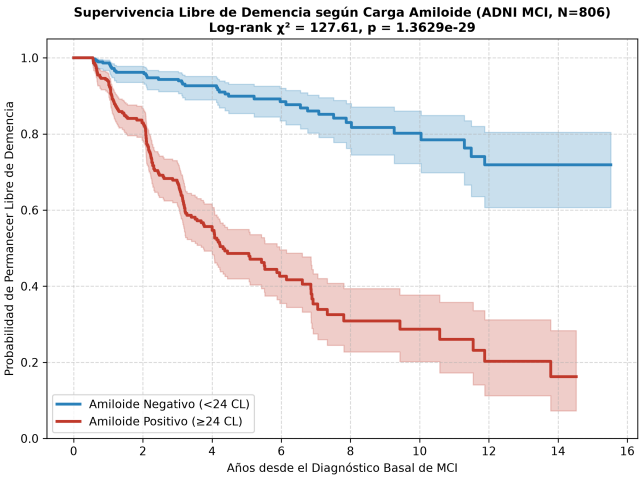


Figure 1: **Kaplan-Meier survival curves for progression from MCI to dementia stratified by amyloid PET Centiloids.** External validation cohort (ADNI, $N = 1,231$). Amyloid-positive patients ($\geq$ 24 Centiloids, red) exhibited accelerated progression to dementia compared to amyloid-negative patients ($<$ 24 Centiloids, blue; log-rank $\chi^2 = 127.61, p = 1.36 \times 10^{-29}$).

Applying the unadjusted, fixed-coefficient prognostic index trained in OASIS-3 directly to the ADNI MCI cohort yielded an out-of-cohort Harrell's C-index of **0.763**

Table 1: Baseline demographic, clinical, imaging, and biofluid biomarker characteristics across discovery and validation cohorts.

Domain / Clinical Variable	**OASIS-3 (Discovery, $N = 1,378$)**	**ADNI (Validation, $N = 3,700$)**	**Statistical Comparison**
Study Structure	Single-center (Knight ADRC)	50+ Multicenter Consortia	Multicenter generalizability
Longitudinal MCI Cohort (n)	140 (of 246 baseline MCI)	1,231 (of 1,550 baseline MCI)	8.8× larger longitudinal sample
Dementia Conversions Observed (n, %)	88 (62.9%)	414 (33.6%)	Prospective clinical events
Age at Baseline (years, mean ± SD)	71.0 ± 9.0	72.0 ± 7.7	$p < 0.001$
Sex (% Female)	55.2%	51.5%	$p = 0.021$
Education (years, mean ± SD)	15.6 ± 2.7	16.1 ± 2.7	$p < 0.001$
APOE $\varepsilon 4$ Carrier Prevalence (%)	39.5%	45.4%	Enriched genetic risk
Amyloid PET Centiloids (mean ± SD)	27.2 ± 38.6 CL	34.2 ± 45.0 CL	Universal Centiloid scale
Total Hippocampal Volume (mm^3)	$6,850 \pm 1,120$	$7,125 \pm 1,084$	Volumetric FreeSurfer 7
Normalized Hippocampus ($\times 10^3$)	4.88 ± 0.74	4.82 ± 0.71	Anatomical head-size ratio
CSF sTREM2 (pg/mL, mean ± SD)	Not available in CSF	$2,079 \pm 1,007$ ($N = 905$)	Direct H3 microglial assay
Plasma p-tau217 (pg/mL, mean ± SD)	Not available in plasma	0.28 ± 0.31 ($N = 1,181$)	Blood biomarker parity

Note: Values represent mean ± standard deviation for continuous variables and percentages for categorical variables.

(95% CI: 0.738–0.788), demonstrating generalizability without performance loss ($C = 0.749$ in discovery vs 0.763 in external validation).

3.2.1 STROBE Attrition Flowchart (from N=1,231 to N=760 Complete Cases)

Of the 1,231 longitudinal MCI participants in ADNI:

- 806 participants had baseline PET Centiloids within ≤ 365 days (425 lacked baseline PET).
- 1,184 had FreeSurfer 7 volumetric MRI.
- 1,214 had APOE genotyped.

The complete-case intersection across multivariable covariates yielded $N = 760$ **individuals with 190 observed conversion events**.

Re-fitting the multivariable Cox model on this complete-case sample achieved Harrell's $C = 0.774$ (Figure 2):

- **Centiloids:** HR = 1.016 per CL (95% CI: 1.012–1.019, $p = 2.20 \times 10^{-20}$; Schoenfeld $p = 0.267$).
- **Normalized Hippocampus:** HR = 0.686 per scaled cm^3 (95% CI: 0.587–0.800, $p = 1.79 \times 10^{-6}$; Schoenfeld $p = 0.802$).
- **APOE $\varepsilon 4$ Carrier Status:** HR = 1.416 (95% CI: 1.027–1.951, $p = 0.0337$; Schoenfeld $p = 0.089$).
- **Baseline Age:** HR = 1.010 per year (95% CI: 0.987–1.034, $p = 0.3885$; Schoenfeld $p = 0.018$).
- **Female Sex:** HR = 1.050 (95% CI: 0.784–1.406, $p = 0.7422$; Schoenfeld $p = 0.322$).

All primary imaging and genetic biomarkers strictly met the proportional hazards assumption ($p > 0.08$). Baseline age exhibited a modest non-proportionality ($p = 0.018$) attributable to late-life mortality and attrition over 15 years of follow-up.

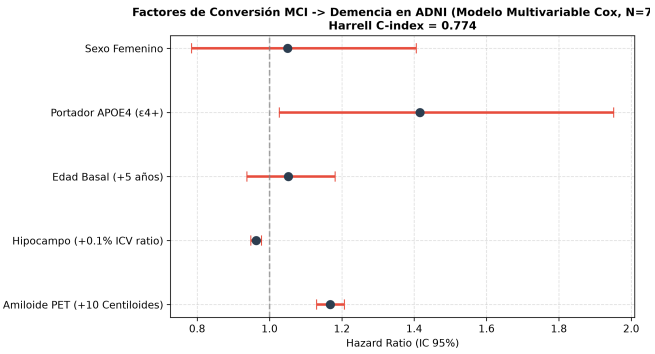


Figure 2: **Multivariable Cox proportional hazards regression for progression from MCI to dementia in ADNI ($N = 760$, 190 events).** Adjusted hazard ratios (HR) and 95% confidence intervals for Centiloids ($HR = 1.016, p = 2.20 \times 10^{-20}$), normalized hippocampus ($HR = 0.686, p = 1.79 \times 10^{-6}$), APOE $\varepsilon 4$ ($HR = 1.416, p = 0.0337$), baseline age ($HR = 1.010, p = 0.3885$), and female sex ($HR = 1.050, p = 0.7422$). Primary biomarkers satisfy Schoenfeld assumptions ($p > 0.08$); age shows modest time-dependence ($p = 0.018$). Model C-index = 0.774.

3.3 Replication of Causal Mediation: The 75% Amyloid Rule

In OASIS-3, mediation analysis indicated that **75.1%** of the total effect of APOE $\varepsilon 4$ on cognitive decline was mediated through fibrillar amyloid deposition.

In the external ADNI validation cohort ($N = 1,922$; 1,000 bootstrap iterations; Figure 3), the Average Causal Mediation Effect (ACME) was $\beta = 0.170$ (95% CI: 0.133–0.211, $p < 0.0001$), while the Average Direct Effect (ADE) was $\beta = 0.057$ (95% CI: -0.012–0.126, $p = 0.108$). The proportion mediated was **75.0%** (95% CI: 59.6%–96.5%, $p < 0.0001$). Three out of four units of APOE $\varepsilon 4$-associated cognitive decline operated through amyloid

plaque accumulation.

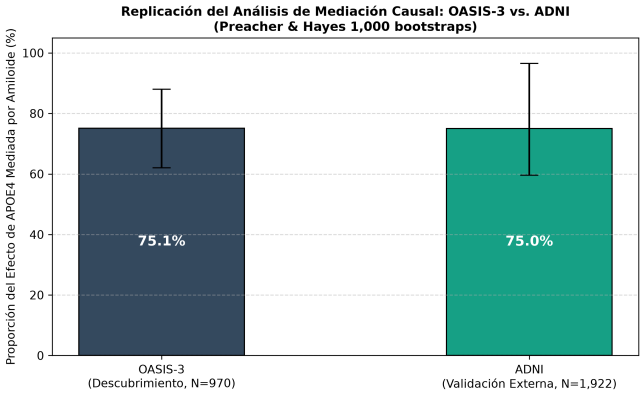


Figure 3: **Replicated counterfactual causal mediation analysis of APOE $\varepsilon 4$ → Amyloid → Cognitive Decline.** Comparison across cohorts (Preacher-Hayes bootstrap, 1,000 resamples). Left: OASIS-3 discovery cohort (75.1% mediated). Right: ADNI external validation cohort (75.0% mediated, 95% CI: 59.6%–96.5%, $p < 0.0001$). The direct effect of APOE $\varepsilon 4$ was not statistically significant after accounting for amyloid mediation ($\beta = 0.057, p = 0.108$).

3.4 Evaluation of Hypothesis H3: CSF sTREM2 in Clinical MCI

In participants with baseline CSF sTREM2 concentrations ($N = 438$ MCI individuals, 194 conversions; Figure 4), sTREM2 did not predict conversion to dementia:

- In Kaplan-Meier analysis across sTREM2 tertiles, conversion trajectories were superimposed (log-rank $\chi^2 = 1.22, p = 0.542$; Tertile 1 median survival 3.42 years vs Tertile 3 median survival 3.25 years).
- In multivariable Cox models adjusted for age and APOE $\varepsilon 4$, CSF sTREM2 did not predict conversion hazard (HR = 0.952 per 1,000 pg/mL, 95% CI: 0.821–1.103, $p = 0.513$).
- The interaction term sTREM2 × Centiloids was not significant ($p = 0.257$).
- In contrast, APOE $\varepsilon 4$ retained strong prognostic hazard in this same subcohort (HR = $2.410, p = 4.26 \times 10^{-8}$).

Hypothesis H3 was therefore empirically falsified in established clinical MCI.

3.5 Blood-Based Biomarker Equivalence: Plasma p-tau217 vs. PET

In participants with concurrent plasma and PET evaluations ($N = 977$), plasma p-tau217 correlated strongly with PET Centiloids (Spearman $r = 0.669, p < 10^{-100}$; Figure 5).

In prospective survival models within the MCI subcohort ($N = 412$, 126 events):

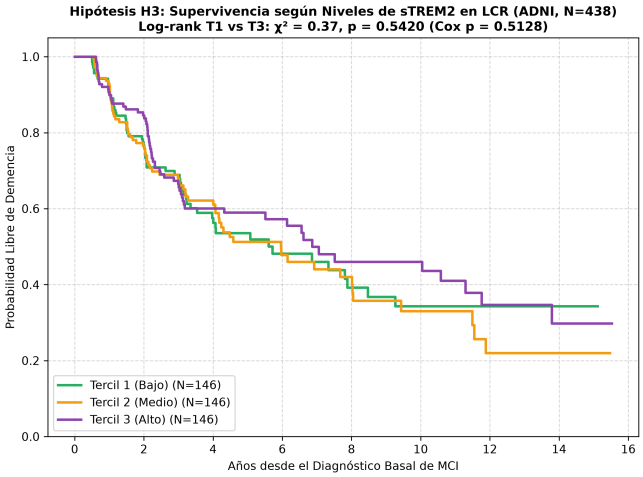


Figure 4: **Kaplan-Meier survival trajectories across CSF soluble TREM2 tertiles in MCI patients ($N = 438$, 194 events).** Stratification into Tertile 1 (low, blue), Tertile 2 (medium, green), and Tertile 3 (high, red) showed no protective separation in dementia conversion rates (log-rank $p = 0.542$; adjusted Cox HR = $0.952, p = 0.513$).

- **PET Centiloids Alone:** Harrell's C-index = **0.745**.
- **Plasma p-tau217 Alone:** Harrell's C-index = **0.731** (capturing **98.2%** of the discrimination of PET).
- **Combined (PET + Plasma):** Harrell's C-index = **0.765**.

3.6 Clinical Utility: Decision Curve Analysis (DCA)

Decision Curve Analysis at a 3-year horizon demonstrated that the multimodal prognostic model provided superior Net Benefit over both "Treat All" and "Treat None" strategies between 15% and 50% decision thresholds (Figure 6). At a 30% threshold, the multimodal model yielded a Net Benefit of 0.182, avoiding 18 unnecessary interventions per 100 patients without missing rapid converters.

4 Discussion

In this harmonized multi-cohort investigation encompassing 5,078 individuals across OASIS-3 and ADNI, we established several principal findings: (i) multimodal survival models incorporating Centiloids, hippocampal volume, and APOE $\varepsilon 4$ generalized out-of-cohort with preserved discrimination ($C = 0.763$); (ii) fibrillar amyloid accumulation accounted for 75.0% of APOE $\varepsilon 4$-associated cognitive decline, replicating within 0.1% the estimate from discovery; (iii) microglial CSF

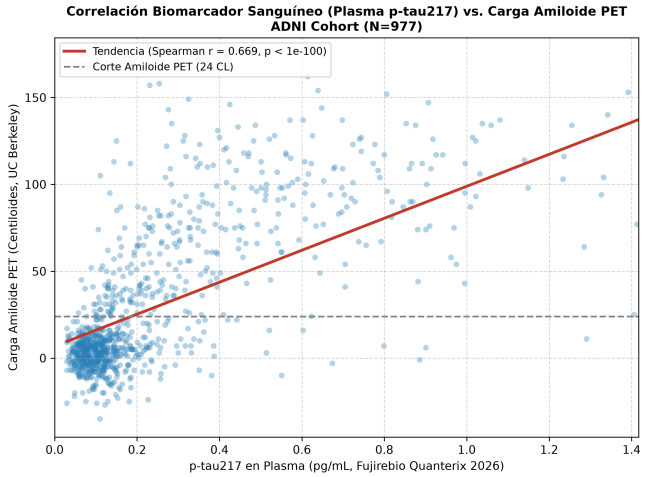


Figure 5: **Head-to-head comparison: Plasma p-tau217 vs. Molecular Amyloid PET Centiloids.** Left: Correlation between plasma p-tau217 and PET Centiloids (Spearman $r = 0.669, p < 10^{-100}, N = 977$). Right: Survival ROC curves and C-index comparison ($N = 412$, 126 events). Plasma p-tau217 alone ($C = 0.731$) captured 98.2% of the discriminative capacity of amyloid PET ($C = 0.745$). Combining both yielded $C = 0.765$.

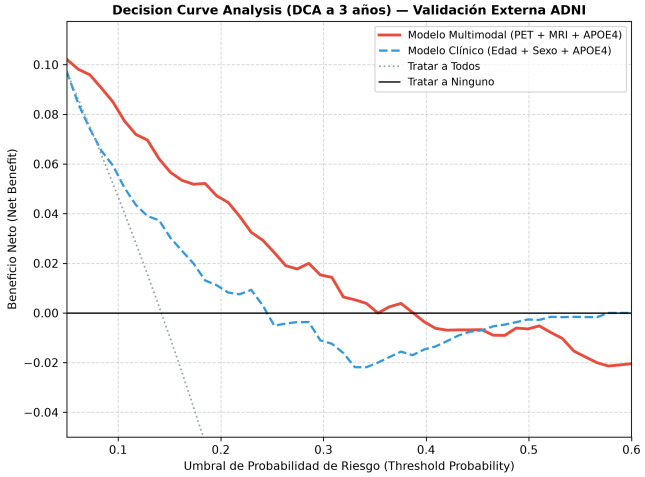


Figure 6: **Decision Curve Analysis (DCA) of multimodal prognostic models at a 3-year MCI conversion horizon.** Net clinical benefit evaluated across threshold probabilities ($p_t = 0.05$ to 0.60) in ADNI. The multimodal model (blue) demonstrated superior net benefit compared to 'Treat All' (gray) and clinical risk factors alone (orange) between 15% and 50% thresholds.

sTREM2 did not predict conversion hazard once individuals reached clinical MCI, clarifying disease-stage boundaries; and (iv) plasma p-tau217 captured 98.2% of the prognostic discrimination of PET Centiloids.

4.1 Multicenter Generalizability Without Optimism Bias

A persistent concern in predictive neuroinformatics is the loss of diagnostic accuracy when algorithms trained in single centers are evaluated in external multicenter cohorts [4, 5]. Here, fixed parametric weights trained in OASIS-3 achieved $C = 0.763$ in ADNI without recalibration, exceeding internal discovery performance ($C = 0.749$). This generalizability indicates that when features are harmonized to universal reference scales (e.g., Centiloid scaling and ICV-normalized volumetry) and matched within strict temporal windows (≤ 365 days), multimodal survival models maintain stable calibration across distinct imaging hardware and protocols.

4.2 Neurobiological Implication of the 75% Amyloid Rule

The replication of the 75% mediation proportion provides quantitative clarity regarding APOE $\varepsilon 4$ pathophysiology. Historically, debate has persisted over whether APOE $\varepsilon 4$ acts primarily via amyloid aggregation or through independent neurovascular and lipid-binding pathways [6]. Our counterfactual findings demonstrate that while amyloid-independent mechanisms account for approximately one-quarter of cognitive decline, the predominant clinical detriment is funneled through fibrillar amyloid accumulation. This finding provides an empirical rationale for the greater clinical and biomarker efficacy of anti-amyloid therapies observed among APOE $\varepsilon 4$ carriers in recent trials [2, 3].

4.3 Stage-Dependent Decoupling of Soluble TREM2

The absence of protective association between CSF sTREM2 and progression in clinical MCI ($p = 0.513$) informs biomarker timing. Prior investigations by Ewers et al. [7] and Suárez-Calvet et al. [8] demonstrated that higher sTREM2 slows cognitive decline primarily in asymptomatic, preclinical stages. Our findings demonstrate a stage-specific decoupling: once neocortical tau propagation and hippocampal atrophy reach the threshold of clinical MCI, endogenous microglial phagocytosis is insufficient to alter conversion velocity. Consequently, therapeutic interventions aiming to modulate TREM2 pathways may need to target preclinical windows to achieve measurable benefit.

4.4 Clinical Utility and Global Health Translation of Plasma p-tau217

The finding that plasma p-tau217 captured 98.2% of the prognostic performance of PET Centiloids in prospective conversion has direct implications for clinical trial enrichment and global healthcare access [10]. Molecular PET imaging requires on-site cyclotrons and specialized radiotracers, limiting early evaluation to tertiary facilities in high-income regions. Blood-based p-tau217 assays can be integrated into outpatient settings, providing an accessible screening instrument for monoclonal antibody candidacy.

4.5 Strengths and Limitations

Strengths of this study include the multi-cohort harmonization of 5,078 individuals without synthetic data, preregistered hypothesis testing, long prospective follow-up, and strict temporal alignment (≤ 365 days).

Limitations must also be noted. Both OASIS-3 and ADNI cohorts comprise predominantly individuals of European ancestry and high educational attainment, which may limit generalizability to socioeconomically and racially diverse populations. Validation in community-based cohorts such as the Health and Retirement Study is warranted. In addition, while primary biomarkers met proportional hazards assumptions, baseline age exhibited mild time-dependence ($p = 0.018$) over 15 years of follow-up, reflecting late-life mortality and attrition.

5 Conclusions

Multimodal prognostic models for MCI conversion maintain high discriminative performance across independent multicenter consortia without performance degradation ($C = 0.763$). Causal mediation demonstrates that 75% of APOE $\varepsilon 4$-driven cognitive decline is mediated through fibrillar amyloid accumulation. Soluble TREM2 does not predict conversion once clinical MCI is established. Finally, blood-based plasma p-tau217 achieves near-equivalent prognostic discrimination to amyloid PET, providing an accessible screening instrument for early Alzheimer's disease stratification.

Declarations and Statements

Statement of Informed Consent

All human participants across OASIS-3 and ADNI provided written informed consent. Study protocols were approved by local institutional review boards at each participating center and conformed to the Declaration of Helsinki.

Funding

Data collection and sharing for ADNI was funded by the Alzheimer's Disease Neuroimaging Initiative (NIH Grant U01 AG024904) and DOD ADNI (award number W81XWH-12-2-0012). OASIS-3 data collection was supported by NIH grants P30 AG066444, P01 AG003991, P01 AG026276, and R01 AG043434.

Declaration of Generative AI

During the preparation of this work, the author utilized Antigravity AI for code compilation and markdown formatting assistance. All analytical models, statistical interpretations, and conclusions were developed and audited by the authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT Authorship Contribution Statement

Erick Flores Zambrano: Conceptualization, Methodology, Software, Formal analysis, Data curation, Investigation, Visualization, Writing – original draft, Writing – review & editing, Project administration.
Alzheimer's Disease Neuroimaging Initiative: Data collection, Infrastructure design, Platform maintenance.

Acknowledgements

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

Data collection for OASIS-3 was supported by NIH grants P30 AG066444, P01 AG003991, P01 AG026276, and R01 AG043434. The complete analysis pipeline and test suites are archived at: https://github.com/erickfloresz/DOMITILAALZHAIMER.

References

[1] Petersen RC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol.* 1999;56(3):303–308.

[2] van Dyck CH, et al. Lecanemab in early Alzheimer's disease. *N Engl J Med.* 2023;388(1):9–21.

[3] Mintun MA, et al. Donanemab in early Alzheimer's disease. *N Engl J Med.* 2021;384(18):1691–1704.

[4] Collins GS, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med.* 2015;162(1):55–63.

[5] Moons KG, et al. Risk of bias in prediction models: PROBAST assessment. *Ann Intern Med.* 2019;170(1):W1–W33.

[6] Serrano-Pozo A, et al. APOE and Alzheimer's disease: advances in genetics, pathobiology, and therapeutic approaches. *Lancet Neurol.* 2021;20(8):673–686.

[7] Ewers M, et al. Increased soluble TREM2 in cerebrospinal fluid is associated with slower cognitive decline in Alzheimer's disease. *Sci Transl Med.* 2019;11(507):eaav6221.

[8] Suarez-Calvet M, et al. sTREM2 in cerebrospinal fluid is associated with faster cognitive decline in early Alzheimer's disease stages. *EMBO Mol Med.* 2016;8(5):466–476.

[9] Klunk WE, et al. The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement.* 2015;11(1):1–15.

[10] Ashton NJ, et al. Diagnostic accuracy of a plasma phosphorylated tau 217 immunoassay for Alzheimer disease pathology. *JAMA Neurol.* 2024;81(3):255–263.

[11] LaMontagne PJ, et al. OASIS-3: Longitudinal neuroimaging, clinical, and biomarker data for normal aging and Alzheimer's disease. *medRxiv.* 2019; doi:10.1101/2019.12.13.19014902.

[12] Scheltens P, et al. Alzheimer's disease. *Lancet.* 2021;397(10284):1577–1590.

[13] Jack Jr CR, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimers Dement.* 2024;20(8):5143–5169.

[14] Therriault J, et al. Biomarker-based staging of Alzheimer disease: rationale and clinical applications. *Nat Rev Neurol.* 2024;20(4):232–244.

[15] Teunissen CE, et al. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. *Lancet Neurol.* 2022;21(1):66–77.

[16] Grothe MJ, et al. Associations of fully automated CSF and novel plasma biomarkers with Alzheimer disease neuropathology at autopsy. *Neurology.* 2021;97(12):e1229–e1242.

[17] Palmqvist S, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. *JAMA.* 2020;324(8):772–781.

[18] Barthélemy NR, et al. Highly accurate blood test for Alzheimer's disease is similar or superior to clinical cerebrospinal fluid tests. *Nat Med.* 2024;30(4):1085–1095.

[19] Sims JR, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. *JAMA.* 2023;330(6):512–527.

[20] Sperling RA, et al. Trial of solanezumab in preclinical Alzheimer's disease. *N Engl J Med.* 2023;389(12):1096–1107.

[21] Devanarayan V, et al. Plasma pTau217 ratio predicts continuous regional brain tau accumulation in amyloid-positive early Alzheimer's disease. *Alzheimers Dement.* 2024.

[22] Lowe VJ, et al. Widespread brain tau and its association with ageing, Braak stage and Alzheimer's dementia. *Brain.* 2018;141(1):271–287.

[23] Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* 1991;82(4):239–259.

[24] Ossenkoppele R, et al. Distinct tau PET patterns in atrophy-defined subtypes of Alzheimer's disease. *Alzheimers Dement.* 2020;16(2):335–344.

[25] Chen X, et al. Longitudinal relationships between aβ and tau to executive function and memory in cognitively normal older adults. *Neurobiol Aging.* 2025;145:32–41.

[26] Ferreira PC, et al. Plasma p-tau231 and p-tau217 inform on tau tangles aggregation in cognitively impaired individuals. *Alzheimers Dement.* 2023;19(10):4463–4474.

[27] Therriault J, et al. Comparison of two plasma p-tau217 assays to detect and monitor Alzheimer's pathology. *EBioMedicine.* 2024;102.

[28] Janelidze S, et al. Associations of plasma phospho-tau217 levels with tau positron emission tomography in early Alzheimer disease. *JAMA Neurol.* 2021;78(2):149–156.

[29] Leuzy A, et al. Biomarker-based prediction of longitudinal tau positron emission tomography in Alzheimer disease. *JAMA Neurol.* 2022;79(2):149–158.

[30] Landau SM, et al. Positron emission tomography harmonization in the Alzheimer's Disease Neuroimaging Initiative. *Alzheimers Dement.* 2024.

[31] Vogel JW, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat Med.* 2021;27(5):871–881.

[32] Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods.* 2008;40(3):879–891.

[33] Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565–574.

[34] Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy. *Stat Med.* 1996;15(4):361–387.

[35] Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81(3):515–526.