# Multimodal Integration of Plasma, MRI, and Genetic Risk for Cerebral Amyloid Prediction

Yichen Wang[1,2,3], Hao-Jie Chen[1,2,3], Yuxin Cheng[4], Yuyan Cheng[5], Shiyun Zhao[5], Kexin Wang[1,2,3], Yidong Jiang[1], Tianyu Bai[1,2,3], Yanxi Huo[1,2,3], Mingkai Zhang[6], Chengxiang Yi[1], Guozheng Feng[8], Weijie Huang[7], Han Ying[6,9,10,11,12,13]*, Ni Shu[1,2,3]*

[1]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

[2]BABRI Centre, Beijing Normal University, Beijing, China

[3]Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing Normal University, Beijing, China

[4]State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Southwest Petroleum University, Chengdu, 610500, Sichuan, P. R. China

[5]School of Mathematical Sciences, Beijing Normal University, Beijing, China

[6]Department of Neurology, XuanWu Hospital of Capital Medical University, Beijing, China

[7]College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, MOE Key Laboratory of Brain Computer Intelligence Technology, Nanjing 211106, China

[8]Tri-Institutional Center for Translational Research in Neuroimaging and Data Science 12 (TReNDS), Atlanta, GA, USA

[9]School of Biomedical Engineering, Hainan University, Haikou 570228, China 11

[10]Center of Alzheimer's Disease, Beijing Institute for Brain Disorders, Beijing 100053, China

[11]National Clinical Research Center for Geriatric Diseases, Beijing 100053, China

[12]The Central Hospital of Karamay, Xinjiang 834000, China

[13]Institute of Biomedical Engineering, Shenzhen Bay Laboratory, Shenzhen 518132, China

**Corresponding author**

*Ni Shu, Email: nshu@bnu.edu.cn, No. 19, Xinjiekouwai St, Haidian District, Beijing, 100875, P. R. China

*Han Ying, Email: hanying@xwh.ccmu.edu.cn

**ORCID**

Yichen Wang: https://orcid.org/0009-0006-4283-7986

Hao-Jie Chen: https://orcid.org/0000-0002-4362-1056

Tianyu Bai: https://orcid.org/0009-0002-9357-5497

Weijie Huang: https://orcid.org/0000-0002-2481-1188

Guozheng Feng: https://orcid.org/0000-0002-6937-8592

Ni Shu: https://orcid.org/ 0000-0003-2420-2910

# Abstract

Accurate estimation of cerebral amyloid-β (Aβ) burden is critical for early detection and risk stratification in Alzheimer's disease (AD). While Aβ positron emission tomography (PET) remains the gold standard, its high cost, invasive nature and limited accessibility hinder broad clinical application. Blood-based biomarkers offer a non-invasive and cost-effective alternative, but their standalone predictive accuracy remains limited due to biological heterogeneity and limited reflection of central nervous system pathology. Here, we present a high-precision, multimodal prediction machine learning model that integrates plasma biomarkers, brain structural magnetic resonance imaging (sMRI) features, diffusion tensor imaging (DTI)-derived structural connectomes, and genetic risk profiles. The model was trained on 150 participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and externally validated on 111 participants from the SILCODE cohort. Multimodal integration substantially improved Aβ prediction, with $R^2$ increasing from 0.515 using plasma biomarkers alone to 0.637 when adding imaging and genetic features. These results highlight the potential of this multimodal machine learning approach as a scalable, non-invasive, and economically viable alternative to PET for estimating Aβ burden.

Keywords: Alzheimer's disease, amyloid-β, plasma biomarkers, structural MRI, diffusion MRI, structural connectome, polygenic risk score, APOE, multimodal prediction, machine learning

# 1 Background

Alzheimer's disease (AD) is the most prevalent neurodegenerative disorder, characterized by the accumulation of amyloid-β (Aβ) plaques and neurofibrillary tangles, both of which are closely linked to cognitive decline[1,2]. Aβ positron emission tomography (Aβ-PET) remains the gold standard for in vivo quantification of cerebral Aβ burden, providing region-specific estimates of amyloid deposition[3]. However, its high cost, invasive nature, and limited accessibility hinder its applicability for large-scale or routine individual screening.

Plasma biomarkers—including phosphorylated tau at threonine 217 (p-tau217) and p-tau181, both reflecting tau hyperphosphorylation; β-amyloid 42/40 ratio (Aβ42/40), reflecting amyloid burden; neurofilament light chain (NfL), marking axonal injury; and glial fibrillary acidic protein (GFAP), indicating astrocytic activation—have emerged as promising, non-invasive tools owing to their accessibility and cost-effectiveness[4]. Although these biomarkers show moderate correlations with PET-derived Aβ burden[5,6], their standalone predictive performance remains suboptimal—particularly in accurately quantifying individual-level deposition—due to biological heterogeneity and their limited capacity to directly reflect central nervous system pathology. Structural MRI (sMRI) and diffusion tensor imaging (DTI) offer non-invasive insights into Alzheimer's disease–related brain changes[7]. sMRI characterizes regional atrophy, whereas DTI supports the construction of structural connectomes that represent inter-regional white matter connectivity[8–10]. Genetic factors also influence the predictive capacity of plasma

biomarkers. The apolipoprotein E (APOE) ε4 allele[11], the most established genetic risk factor for AD, is easily genotyped and widely used in clinical settings. Polygenic risk scores (PRS) [12], although theoretically more comprehensive by aggregating genome-wide variants, are often limited by their cost and complexity[13]. These features may improve Aβ prediction by providing complementary information not captured by plasma biomarkers alone[13,14].

In this study, we developed a high-precision machine learning model to predict cerebral Aβ burden by integrating plasma biomarkers, sMRI features, DTI-derived structural connectomes, and genetic information[15]. To systematically evaluate the contribution of each modality, we performed three analyses: (1) developed a multimodal prediction model and evaluated the incremental predictive value of structural MRI, DTI-based structural connectomes, and genetic features beyond plasma biomarkers for estimating continuous Aβ burden in the ADNI cohort[16]; (2) tested whether the addition of imaging and genetic features improved prediction performance in the independent SILCODE cohort to assess the generalizability of modality-specific effects; and (3) compared APOE and PRS within this framework to assess the incremental predictive value of each and to evaluate whether APOE genotyping could serve as a cost-effective proxy for genome-wide risk profiling[15].

# 2 Methods

## 2.1 Participants and Study Design

This study conducted a secondary analysis based on publicly available data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), which aims to identify early diagnostic biomarkers for AD through multimodal data collection. Participants were included if they were aged between 55 and 90 years, generally in good health, fluent in English or Spanish, and had a Geriatric Depression Scale score below 6. Clinical diagnoses (cognitively unimpaired [CU], mild cognitive impairment [MCI], or AD) were determined using standardized ADNI criteria based on subjective complaints, objective neuropsychological assessments, and clinician-rated global functioning[17]. To ensure multimodal data availability, participants with sMRI, DTI, genetic data, plasma biomarkers, and Aβ-PET imaging acquired within a one-year interval were selected from the ADNI-GO, ADNI-2, and ADNI-3 phases.

To assess the generalizability of the findings, external validation was performed using the SILCODE cohort, a prospective multicenter study of AD and cognitive impairment in the mainland Chinese population. SILCODE participants were included if they were aged between 45 and 90 years, right-handed, native Mandarin speakers, and had complete datasets for sMRI, DTI, plasma biomarkers, APOE genotyping, and Aβ-PET within a one-year window. Demographic characteristics of both cohorts, stratified by diagnostic group, are summarized in **Table 1**.

| | ADNI | | | SILCODE | | |
|---|---|---|---|---|---|---|
| **Diagnose** | AD | CU | MCI | AD | CU | MCI |
| **N** | 10 | 97 | 43 | 20 | 65 | 26 |
| **Age** | 72.27±6.57 | 71.36±6.02 | 74.84±8.62 | 71.40±10.98 | 68.32±5.72 | 70.42±6.46 |
| **Female** | 7(70%) | 36(37%) | 26(60%) | 12(60%) | 40(62%) | 11(42%) |

**Table 1 | Demographic characteristics of participants in the ADNI and SILCODE cohorts.**
Sample size (N), mean age (± standard deviation), and number and proportion of female participants are reported for each diagnostic group: Alzheimer's disease (AD), mild cognitive impairment (MCI), and cognitively unimpaired (CU). Age values are presented as mean ± SD. Female values represent count (percentage).

## 2.2 MRI Data Acquisition

sMRI and DTI data were obtained from both the ADNI and SILCODE cohorts. In ADNI, MRI scans were acquired using 3.0 Tesla (3T) scanners manufactured by Philips, Siemens, or GE Healthcare across the ADNI-GO, ADNI-2, and ADNI-3 phases. For sMRI, a standardized protocol was used with typical parameters including a repetition time (TR) of 6.7 ms, echo time (TE) of 3.1 ms, flip angle of 9°, field of view (FOV) of 256 × 256 × 170 $mm^3$, slice thickness of 1.2 mm, and voxel size ranging from 1.0 to 1.11 $mm^3$, covering 170 sagittal slices. DTI scans in ADNI-GO and ADNI-2 were acquired using GE 3T scanners with TR = 9000 ms, TE = 60–70 ms, and voxel size = 1.37 × 1.37 × 2.7 $mm^3$, comprising 30 diffusion-weighted directions (b = 1000 s/$mm^2$) and five b = 0 images. In ADNI-3, DTI data were acquired with an updated protocol using a resolution of 2 × 2 × 2 $mm^3$ (TR = 7200 ms, TE = 56 ms), 35 axial slices, and a total scan time of approximately 7.5 minutes.

In the SILCODE cohort, imaging was performed using a 3T integrated PET/MR scanner (SIGNA PET/MR, GE Healthcare) at Xuanwu Hospital, Capital Medical University. Structural MRI was acquired using a gradient-recalled echo sequence with TR = 2300 ms, inversion time (TI) = 900 ms, TE = 2.26 ms, flip angle = 8°, FOV = 256 × 256 $mm^2$, slice thickness = 1 mm, and isotropic voxel size of 1 $mm^3$. DTI data were acquired using a spin-echo echo-planar imaging (EPI) sequence with TR = 4500 ms, TE = 65 ms, voxel size = 2 × 2 × 2 $mm^3$, and 64 slices. Diffusion gradients were applied in 128 directions, including 64 directions at b = 1000 s/$mm^2$ and 64 at b = 2000 s/$mm^2$, along with one non-diffusion-weighted (b = 0) image.

## 2.3 MRI Data Preprocessing and Structural Connectome Construction

All imaging data were preprocessed using standardized protocols adapted from the UK Biobank pipeline[18], implemented using *FSL*[19], *FreeSurfer*[20], and *MRtrix3*[21]. Raw DICOM images were converted to NIfTI format via *dcm2niix*[22]. For diffusion imaging, b-value distributions were inspected prior to preprocessing. In the ADNI cohort, only volumes with b = 0 and b = 1000 s/$mm^2$ were retained for analysis. In contrast, the SILCODE cohort included full multi-shell data (b = 0, b = 1000, and b = 2000 s/$mm^2$) for downstream processing. sMRI preprocessing for both cohorts included skull stripping and bias field correction using *FSL*, followed by cortical surface reconstruction via *FreeSurfer's recon-all* pipeline. DTI data were corrected for head motion, eddy currents, and EPI distortions using FSL's eddy[23]. Fieldmap estimation differed across cohorts: ADNI employed Synb0-DISCO[24] to generate synthetic fieldmaps from paired reverse

phase-encoded images, while SILCODE used fieldmaps derived from phase and magnitude images.

Structural connectomes (SCs) were constructed based on tractography models tailored to the diffusion acquisition scheme. In ADNI, fiber orientation distributions (FODs) were modeled using single-shell constrained spherical deconvolution (CSD)[25–27], followed by probabilistic tractography. For SILCODE, which acquired multi-shell data, FODs were estimated using multi-shell multi-tissue CSD (MSMT-CSD) [28]. In both cohorts, tractography was seeded from the gray–white matter interface using the *5ttgmwmi* algorithm[29]. Brain network nodes were defined using a hybrid atlas comprising 100 cortical regions from the Schaefer atlas[30] and 16 subcortical regions from the Melbourne Subcortex Atlas (MSA) [31]. The resulting SC matrices were used for downstream feature extraction and predictive modeling.

## 2.4 Quantification and Harmonization of Amyloid PET Measures

In ADNI cohort, Aβ burden was quantified using Centiloid (CL) values [32,33], processed uniformly by the University of California, Berkeley, and publicly released alongside the official dataset [UC Berkeley Amyloid PET Processing Methods, 2023]. CL values were derived from cortex-wide SUMMARY_SUVR measures, normalized to the whole cerebellum, and converted using tracer-specific formulas:

For [$^{18}$F]-florbetapir (FBP):

$$CL = 188.22 \times SUMMARY_SUVR - 189.16$$

For [$^{18}$F]-florbetaben (FBB):

$$CL = 157.15 \times SUMMARY_SUVR - 151.87$$

SUMMARY_SUVR values were based on sMRI-guided segmentation of four cortical regions (frontal, cingulate, parietal, and temporal), with normalization to the whole cerebellum; regional definitions are detailed in the Supplementary Materials(table .1). In this study, we directly utilized the standardized CL values provided by ADNI rather than recalculating SUVRs. This decision was based on two considerations: (i) the CL framework enables cross-tracer and cross-scanner harmonization, improving comparability and robustness; and (ii) the official CL pipeline has been extensively validated and widely adopted, ensuring high reproducibility.

In the SILCODE cohort, Aβ PET imaging was conducted using the radiotracer [18F]-florbetapir ([18F]-AV-45). Participants received an intravenous injection of 7 to 10 mCi of [18F]-florbetapir, followed by a rest period of approximately 40 min, after which a 20-minute static PET scan was performed. PET data acquisition employed a time-of-flight ordered subset expectation maximization (TOF-OSEM) algorithm[34,35] with 8 iterations, a field of view of 350 × 350 mm2, subset matrices of 32 at 192 × 192, and a full width at half maximum (FWHM) of 3 mm. In this study, Aβ burden was quantified using SUMMARY_SUVR values, calculated with the same cortical and reference regions as in the ADNI processing pipeline.

## 2.5 Plasma Biomarker Data

In the ADNI cohort, plasma concentrations of Aβ42, Aβ40, p-tau217,NfL, and GFAP were centrally measured by the Biomarker Core Laboratory at the University of Pennsylvania[4,5]. Aβ42, Aβ40, and p-tau217 were quantified using the Lumipulse G1200 platform (Fujirebio) [4], while NfL and GFAP were measured using the Simoa HD-X platform (Quanterix). All plasma assays followed standardized protocols, underwent centralized quality control, and were reported in picograms per milliliter (pg/mL), ensuring cross-sample comparability and analytical readiness[5].

For the SILCODE cohort, plasma levels of Aβ42, Aβ40, p-tau181, NfL, and GFAP were measured using the Simoa HD-X platform in accordance with manufacturer protocols from the Neurology 4-Plex E Advantage Kit and the p-tau-181 Advantage V2 Kit. All samples were assayed in duplicate, with each plate including internal calibrators and quality control standards. The lower limits of detection were: Aβ42 (0.136 pg/mL), Aβ40 (0.384 pg/mL), NfL (0.090 pg/mL), GFAP (0.441 pg/mL), and p-tau181 (0.028 pg/mL). For all participants, plasma sampling and Aβ-PET acquisition occurred within a 30-day interval to ensure temporal alignment.

## 2.6 Genetic Risk Assessment

In ADNI cohort, Genotyping was performed using either the Illumina HumanOmniExpress array or the Infinium Global Screening Array v2. To ensure consistency across batches, each dataset underwent separate preprocessing, imputation,

and quality control (QC). Preprocessing included standardizing sample IDs, recoding chromosome labels, removing non-autosomal variants, eliminating duplicate SNPs, and correcting strand orientations using *PLINK*[36]and *Bcftools*[37]. Phasing was conducted with the 1000 Genomes Phase 3 reference panel[38], and imputation was performed using *Impute5*[39]. Variants with low imputation quality (INFO score < 0.8) were excluded. After merging batches, a unified quality control (QC) procedure was applied. Single nucleotide polymorphisms (SNPs) with a minor allele frequency (MAF) < 0.01, genotype missingness > 1%, or significant deviation from Hardy–Weinberg equilibrium ($p < 1\times10^{-6}$) were removed. Individuals were excluded if they showed high genotype missingness, excessive heterozygosity (F-statistic > ±3 standard deviations), or cryptic relatedness (pairwise identity-by-descent > 0.125). Discrepancies between reported and genetically inferred sex were also checked and removed.

PRS were constructed using the clumping and thresholding (C+T) method[40], based on Alzheimer's disease GWAS summary statistics[41]. We tested multiple p-value inclusion thresholds ($p < 10^{-6}$, LD-based clumping) combined with LD-based clumping (window size = 250k kb, $r^2 < 0.01$). For each individual *j*, the PRS was calculated as:

$$PRS_j = \frac{\sum_i^N S_i * G_{ij}}{P * M_j}$$

where: $S_i$ is the effect size (β) of $SNP_i$ from the external GWAS; $G_{ij}$ is the number of effect alleles observed for $SNP_i$ in individual *j*; *P* is the ploidy; *N* is the total number of

SNPs included in the PRS; $M_j$ is the number of non-missing SNPs in individual $j$.The resulting scores were standardized (Z-scores) before downstream analyses. The top five genetic principal components from a principal component analysis (PCA) were included as covariates to control for population stratification.

The SILCODE cohort did not undergo genome-wide genotyping. Only APOE genotypes (rs429358 and rs7412) were available, obtained using standardized laboratory protocols and included as categorical covariates (ε2, ε3, ε4) in the analyses.

## 2.7 Multimodal Feature Integration and Predictive Modeling

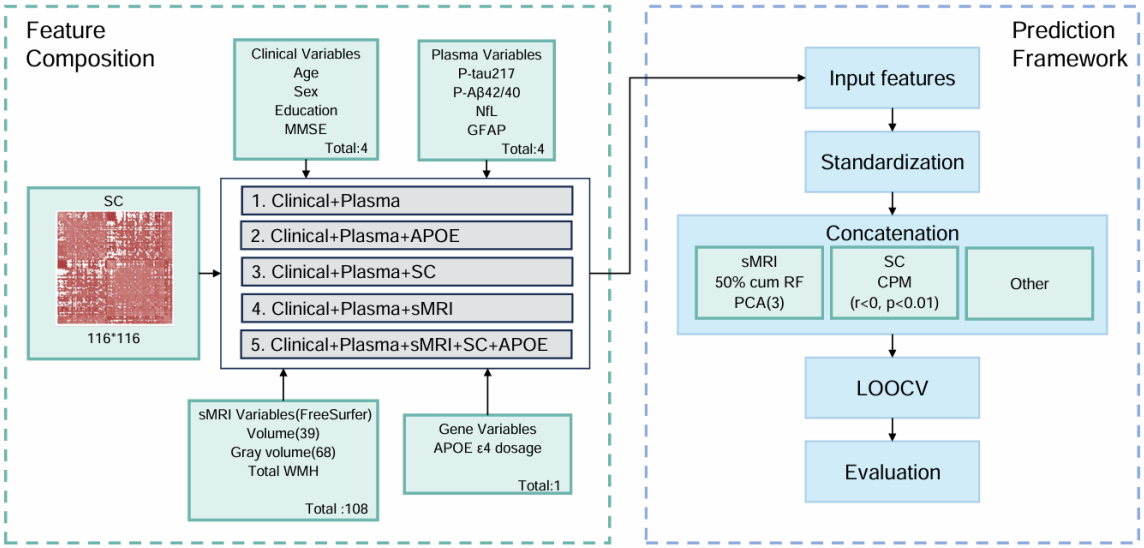


Figure 1 | Multimodal machine learning framework for predicting cerebral amyloid-β burden.
An integrative pipeline was constructed to combine clinical, plasma, genetic, sMRI, and SC features. Five feature combinations were tested under a LOOCV framework. Imaging and connectomic features were dimensionally reduced or selected based on established criteria. Models were trained using both linear and ensemble regressors.
Abbreviations: Aβ, amyloid-β; APOE, apolipoprotein E4 dosage; CPM, connectome-based predictive modeling; DTI, diffusion tensor imaging; LOOCV, leave-one-out cross-validation; sMRI, magnetic resonance imaging; PCA, principal

component analysis; PRS, polygenic risk score; SC, structural connectome; sMRI, structural MRI; MMSE, Mini-Mental State Examination.

To build an accurate prediction model and assess the added predictive value of sMRI, SC, and genetic features beyond plasma biomarkers, we implemented a modular machine learning framework enabling multimodal integration and systematic comparison across feature combinations[13,14,42]. Based on plasma features as the baseline, five feature combinations were defined and compared: Clinical + Plasma, Clinical + Plasma + APOE, Clinical + Plasma + SC, Clinical + Plasma + MRI, and Clinical + Plasma + MRI + SC + APOE. Two regression algorithms, linear regression and random forest(RF) regression, were used for model benchmarking (Fig. 1).

All numerical variables, including plasma biomarkers and sMRI-derived features, were z-score standardized within the training folds of LOOCV to prevent data leakage[43]. A total of 108 sMRI features were included, encompassing cortical volumes, surface areas, and white matter hyperintensities (WMH), all normalized by intracranial volume to control for inter-individual differences in brain size[44]. Within each training set, features were ranked by importance using RF regression, with the top 50% retained and further reduced to three orthogonal components via principal PCA[45]. SC features were extracted using the CPM approach[46,47,47]. For each training fold, edges negatively correlated with CL values (Pearson's $r < 0$, $p < 0.01$) were identified, and the mean strength of these edges (SC_NegAvg) was computed as a single summary feature per subject.

Model training and validation were performed using LOOCV. All preprocessing, feature selection, and dimensionality reduction steps were strictly confined to the training data in each fold. The resulting model was then applied to the held-out subject to obtain unbiased performance estimates.

## 2.8 Interpretability Analysis

To further elucidate the contributions of individual features to model predictions and enhance interpretability, we employed SHapley Additive exPlanations (SHAP) [48] to interpret the machine learning models. SHAP is a game theory–based feature attribution method proposed by Lundberg and colleagues. It quantifies the marginal contribution of each input feature to a given prediction while accounting for feature interactions.

Specifically, SHAP constructs a local linear model to approximate and interpret complex nonlinean models $f(x)$, formulated as:

$$g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i \, x'_i$$

where $x'$ represents a simplified binany version of the input features (indicating feature presence or absence), $\phi_0$ denotes the expected output when all features are absent, and $\phi_i$ represents the Shapley value of feature $i$, reflecting its marginal contribution to the prediction.

In this study, SHAP values were computed across the entire sample using the best-performing model. A feature importance summary plot was generated to identify the key variables contributing most strongly to Aβ burden prediction, both before and after the inclusion of SC features.

## 2.9 Sensitivity Analysis

To evaluate the robustness of the proposed framework, we conducted sensitivity analyses across three methodological dimensions: (1) polygenic risk score (PRS) construction, (2) structural connectome (SC) definition, and (3) machine learning algorithm.

First, PRSs were reconstructed using clumping and thresholding with a range of p-value cutoffs ($p < 10^{-5}$, $10^{-7}$, $10^{-8}$, $10^{-9}$, and $10^{-10}$), in addition to the primary threshold ($p < 10^{-9}$). Second, to assess the influence of brain parcellation, SC features were derived using an alternative atlas combining aparc and MSA-16 (totaling 84 regions), in place of the default Schaefer-100 + MSA-16 template. Third, we replaced the primary learning algorithms (linear regression and random forest) with CatBoost to examine the dependency of model performance on algorithm selection.

# 3 Results

## **3.1** Associations between multimodal features and cerebral amyloid-β burden

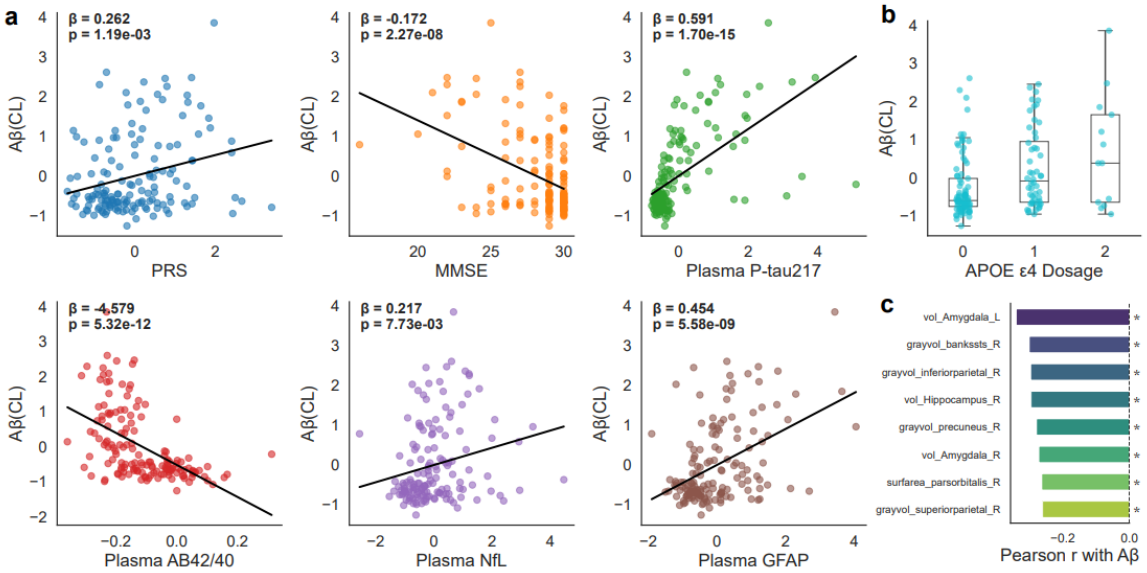


Figure 2 | Associations between plasma biomarkers, cognitive function, genetic risk, and brain structural features with standardized Aβ burden.

a, Scatter plots showing the linear associations between Z-scored CL values and PRS, MMSE, p-tau217, Aβ42/40 ratio, NfL, and GFAP. Standardized regression coefficients (β) and corresponding p-values are annotated in each panel.

b, Boxplots illustrating the distribution of Z-scored Aβ burden stratified by APOE ε4 allele dosage (0, 1, or 2 copies).

c, The top eight ICV-normalized sMRI features (volume or surface area) most strongly correlated with Aβ burden, ranked by Pearson correlation coefficients. Asterisks (*) indicate significant associations ($p < 0.05$, uncorrected).

Abbreviations: Aβ, amyloid-β; APOE, apolipoprotein E; CL, Centiloid; GFAP, glial fibrillary acidic protein; MMSE, Mini-Mental State Examination; NfL, neurofilament light chain; PRS, polygenic risk score.

Among plasma biomarkers, p-tau217 exhibited the strongest positive association with Aβ burden (standardized regression coefficient $\beta = 0.591$, $p = 1.70 \times 10^{-15}$), followed by GFAP ($\beta = 0.454$, $p = 5.58 \times 10^{-9}$). The plasma Aβ 42/40 ratio was significantly negatively associated with Aβ burden ($\beta = -4.579$, $p = 5.32 \times 10^{-12}$), consistent with

prior evidence that increased brain amyloid accumulation is accompanied by reduced peripheral Aβ 42/40 levels[49]. NfL showed a moderate positive correlation, though weaker than those observed for p-tau217 and GFAP. MMSE scores were weakly negatively associated with Aβ burden ($\beta = -0.172$, $p = 2.27 \times 10^{-8}$), indicating that cognitive decline is modestly related to amyloid pathology, whereas plasma biomarkers demonstrated greater predictive sensitivity(Fig. 2a).

Genetic features showed comparatively lower predictive value: PRS($p<1\times10^{-6}$) was positively but weakly associated with Aβ burden ($\beta = 0.262$, $p = 1.19 \times 10^{-3}$), while APOE ε4 status exhibited a clear dose-dependent relationship with Aβ burden, with higher allele dosage associated with greater accumulation (Fig. 2b). sMRI analyses identified several ICV-normalized volume and surface area features significantly associated with Aβ burden (Fig. 2c). The top eight features, ranked by correlation strength, were predominantly located in the medial temporal lobe, parietal cortex, and subcortical regions, including the amygdala and hippocampus. These findings align with early anatomical vulnerability in Alzheimer's disease and reinforce the association between structural brain alterations and amyloid deposition.

## 3.2 Multimodal feature integration enhances Aβ burden prediction performance

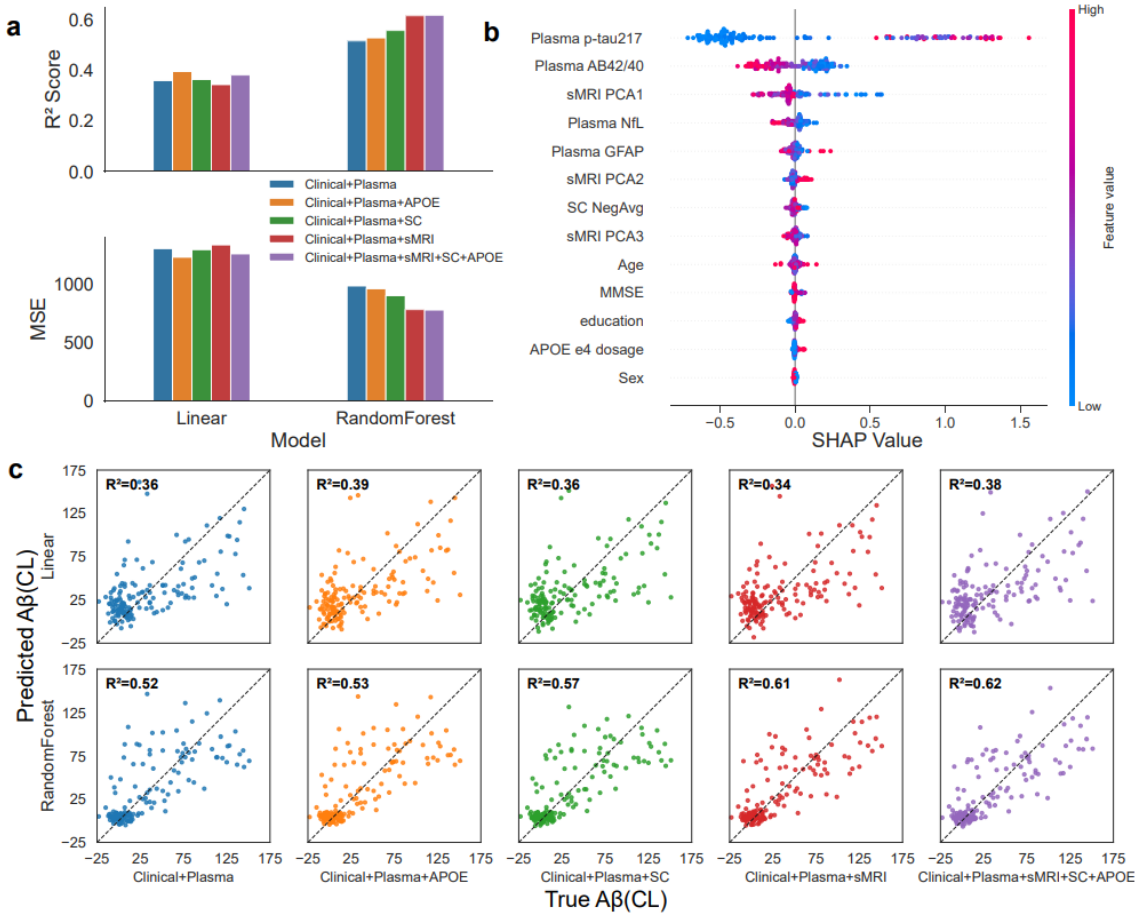


Figure 3 | Model performance and SHAP-based feature importance for Aβ burden prediction.

a, Performance comparison of different feature combinations using LR and RF models under a LOOCV framework. The upper panel shows $R^2$ scores and the lower panel shows MSE values for each combination.

b, SHAP summary plot for the best-performing RF model (Clinical + Plasma + sMRI + SC + APOE), ranking features by their mean absolute SHAP values. The x-axis indicates each feature's directional contribution to the predicted Aβ burden.

c, Scatter plots of observed versus predicted Aβ burden for both LR and RF models across all feature combinations. $R^2$ scores are annotated in each subplot.

Abbreviations: Aβ, amyloid-β; APOE, apolipoprotein E; GFAP, glial fibrillary acidic protein; MMSE, Mini-Mental State Examination; NfL, neurofilament light chain; PCA, principal component analysis; SC, structural connectivity;

SHAP, SHapley Additive exPlanations; RF, random forest; LR, linear regression; MSE, mean squared error; Clinical, Clinical, demographic and cognitive measures.

We systematically evaluated the predictive performance of different feature combinations under a LOOCV framework. Overall, linear regression exhibited limited predictive capacity across all feature sets (e.g., $R^2 = 0.36$ for the Clinical + Plasma model), suggesting that it fails to capture the complex nonlinear relationships inherent to multimodal features. In contrast, the RF model consistently outperformed linear regression, with prediction accuracy improving as additional modalities were incorporated.

Specifically, the RF model using clinical and plasma features achieved an $R^2$ of 0.515. Incorporating APOE ε4 genotype modestly improved performance ($R^2 = 0.527$), and the addition of structural connectivity (SC) features yielded a further increase ($R^2 = 0.556$). Including T1-weighted structural MRI (sMRI) features substantially enhanced predictive power ($R^2 = 0.615$; MSE = 781.5). The best overall performance was achieved when all modalities—clinical, plasma, sMRI, SC, and APOE—were integrated, yielding an $R^2$ of 0.617 and the lowest MSE (770.1). Detailed $R^2$ and MSE values for all feature combinations are reported in Supplementary Materials(table .2).

To interpret the model outputs, SHAP analysis was applied to the best-performing RF model. Plasma biomarkers, particularly p-tau217 and the Aβ42/40 ratio, contributed most strongly to the model's predictions, followed by principal components derived from sMRI

features and SC-based connectivity metrics. Age, MMSE scores, and APOE ε4 dosage also contributed to predictive performance. The feature importance distribution suggests that while plasma biomarkers provide the dominant predictive signal, image and genetic features offer meaningful complementary information.

Finally, scatter plot comparisons of predicted versus observed Aβ burden further confirmed the advantage of multimodal integration (Fig. 3c). The Clinical + Plasma + sMRI + SC + APOE model achieved the highest model fit ($R^2 = 0.617$), outperforming all other combinations. These findings highlight the added value of combining plasma, structural, connectome, and genetic data for improving non-invasive prediction of cerebral amyloid burden.

## 3.3 Distributed Network Disruption Associated with Aβ Burden

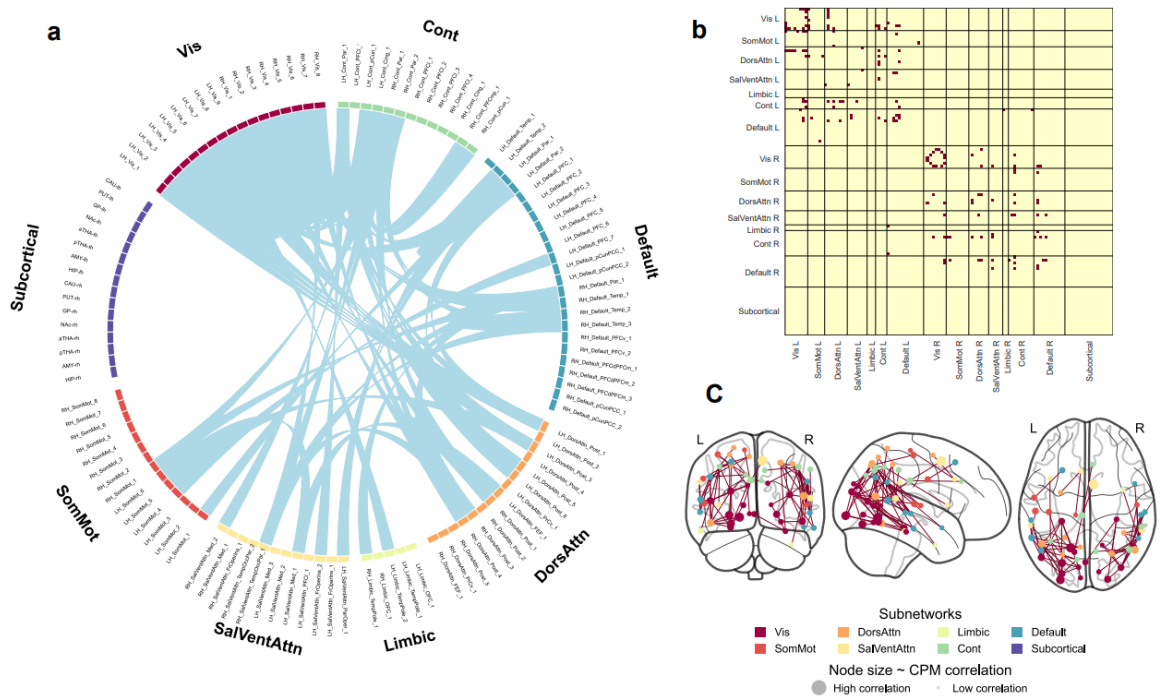


Figure 4 | Stable structural connections identified through leave-one-out cross-validation and their visualization.

a, Chord diagram illustrating the intersection of significantly positive connections identified by CPM across all training sets under the LOOCV framework. Nodes are grouped by functional subnetworks, and line width reflects the strength of correlation.

b, Inter-subnetwork connectivity matrix displaying the distribution of stable connections among major functional modules. Red dots indicate the presence of consistently selected edges; rows and columns represent different subnetworks.

c, Cortical surface projection of stable structural connections. Node size indicates CPM correlation strength, and both nodes and edges are colour-coded by functional subnetwork.

Abbreviations: CPM, connectome-based predictive modeling; LOOCV, leave-one-out cross-validation; Vis, visual network; SomMot, somatomotor network; DorsAttn, dorsal attention network; SalVentAttn, salience/ventral attention network; Limbic, limbic network; Default, default mode network; Cont, control network; Subcortical, subcortical structures.

To identify structural connections consistently weakened by Aβ burden, we applied CPM under a LOOCV framework. In each fold, connections negatively associated with CL values ($p < 0.01$) were selected, and those recurring across all folds were retained, yielding 82 stable disconnections. As shown in Fig. 4, the visual network exhibited the highest number of inter-network disconnections, particularly with the control, somatomotor, dorsal attention, salience, and limbic networks. The most frequently affected pairs included visual – control ($n = 5$), visual – somatomotor ($n = 4$), and visual – default ($n = 4$). By contrast, default and subcortical networks were largely spared. These patterns suggest selective vulnerability of perceptual and attentional integration pathways during early amyloid accumulation.

Spatial mapping revealed consistent involvement of the primary visual cortex, precentral gyrus, lateral parietal cortex, and medial prefrontal cortex—regions involved in sensorimotor processing, attentional control, and executive function. Notably, the medial prefrontal cortex and lateral parietal regions have been implicated as early sites of Aβ accumulation, particularly within the default mode and frontoparietal control networks[50,51] These findings align with prior studies showing that Aβ pathology alters large-scale network organization beyond memory circuits[52], and can emerge in asymptomatic individuals[53,54]. Collectively, the observed disconnection patterns support a network-level view of AD, wherein early amyloid deposition perturbs integrative brain systems. Structural connectome features identified through robust CPM may thus serve as sensitive and interpretable biomarkers for early disease characterization and risk stratification.

## 3.4 External validation in the SILCODE cohort

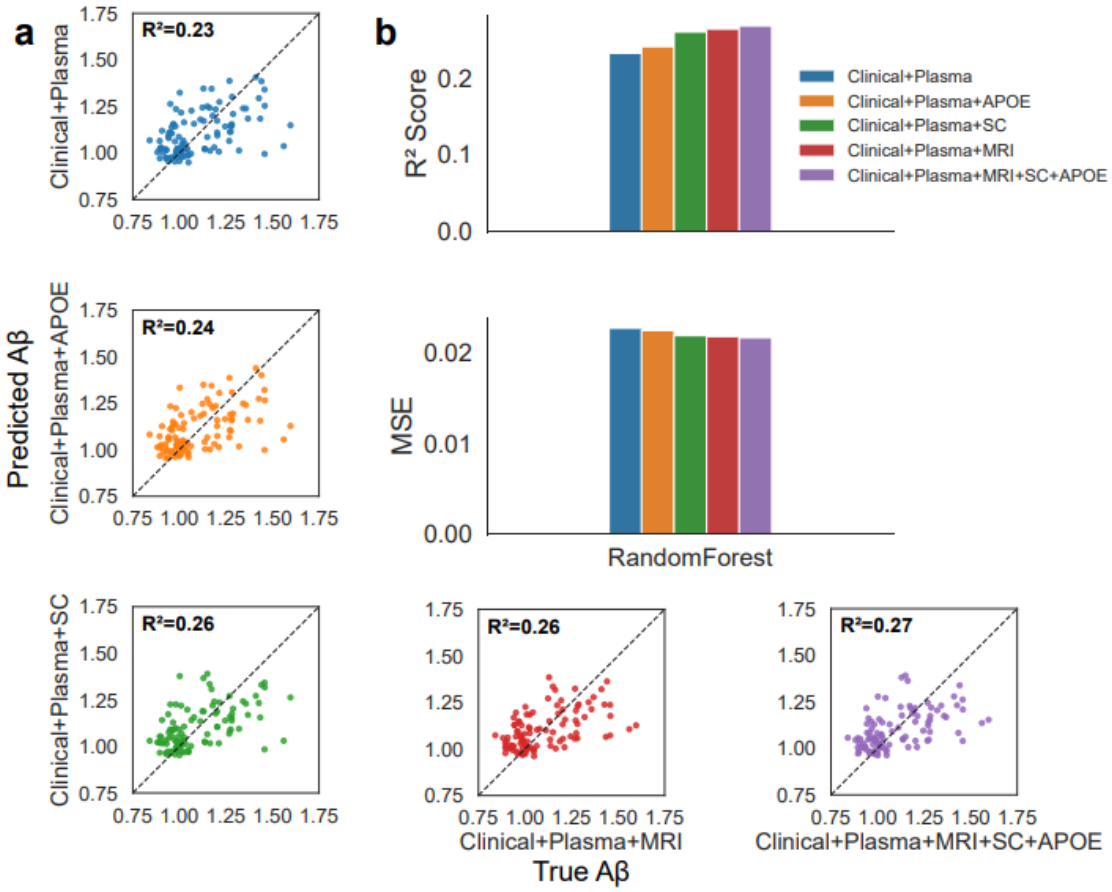


Figure 5 | Model performance for Aβ burden prediction in the SILCODE validation cohort.

(A) Scatter plots depict observed versus predicted Aβ SUVR values across five feature combinations using random forest (RF) models. $R^2$ scores are annotated in each panel to indicate model fit. The x-axis denotes observed SUVR values; the y-axis shows model-predicted values.

(B) Bar plots summarize model performance under a leave-one-out cross-validation (LOOCV) framework. The upper panel shows $R^2$ values, and the lower panel displays mean squared error (MSE) across feature combinations.

Abbreviations: Aβ, amyloid-β; APOE, apolipoprotein E; GFAP, glial fibrillary acidic protein; MMSE, Mini-Mental State Examination; NfL, neurofilament light chain; PCA, principal component analysis; SC, structural connectivity; RF, random forest; MSE, mean squared error. Clinical, Clinical, demographic and cognitive measures.

To evaluate the generalizability of our predictive framework, we conducted external validation in the independent SILCODE cohort using five feature combinations with RF models to predict Aβ PET SUVR values (Fig. 5). Compared to the baseline model (Clinical+Plasma, R=0.232), the addition of APOE genotype ($\Delta R^2 = +0.009$), structural connectivity (SC; $\Delta R^2 = +0.027$), and structural MRI (sMRI) features ($\Delta R^2 = +0.032$) each incrementally improved performance, with the full model (Clinical+Plasma+MRI+SC+APOE) achieving the highest $R^2$ (0.267) $\Delta R^2=0.035$. While this performance was lower than that observed in the ADNI cohort, the pattern of incremental gains from each modality remained consistent across datasets, underscoring the robustness of multimodal contributions. A detailed comparison of $R^2$ and mean squared error (MSE) across all models is provided in Supplementary Materials(table .3).

The lower predictive accuracy in SILCODE is mainly attributable to differences in plasma biomarker sensitivity. ADNI employed plasma p-tau217, which has demonstrated superior accuracy in identifying Aβ pathology, whereas SILCODE utilized p-tau181, a more widely available but less sensitive biomarker. Prior studies have shown that p-tau217 better distinguishes amyloid-positive individuals, suggesting that the assay difference may partly explain the performance gap[6]. Despite differences in population demographics, imaging protocols, and biomarker platforms, the consistent predictive value of sMRI, SC, and APOE genotype across cohorts supports the generalizability and translational utility of the proposed framework.

## 3.5 Evaluating APOE and polygenic risk in model performance

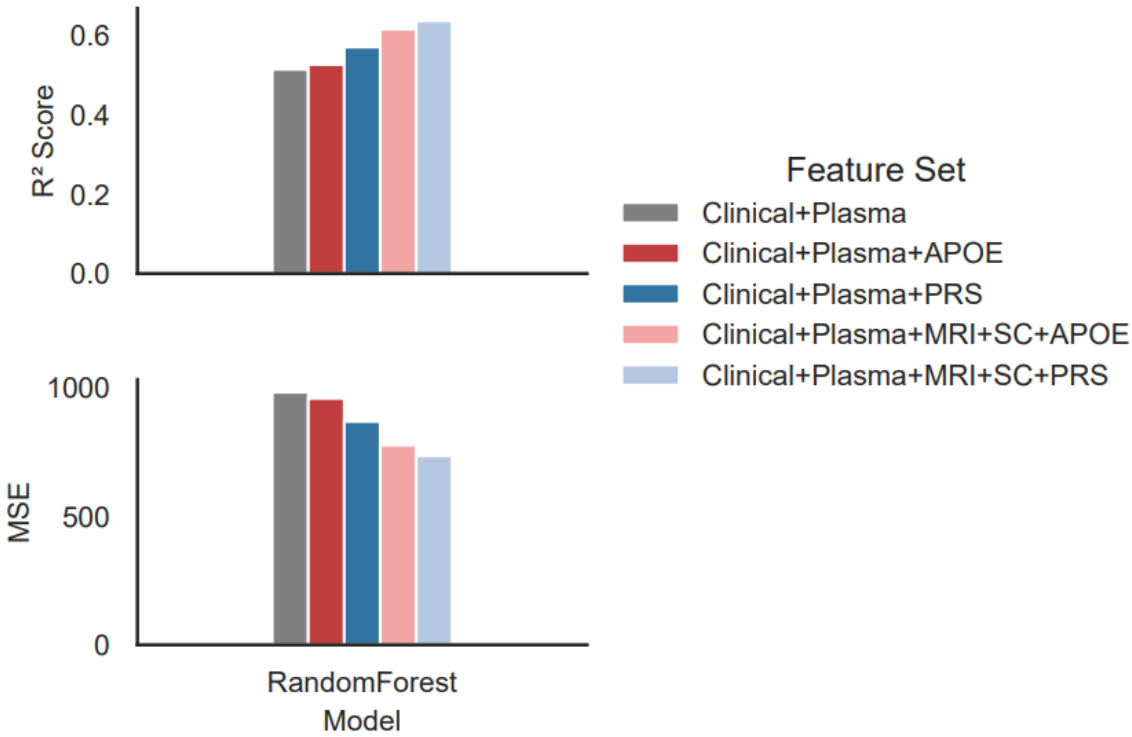


Figure 6 | Model performance with APOE or PRS for predicting Aβ burden in the ADNI cohort. Bar plots show $R^2$ (top) and MSE (bottom) for models using different combinations of clinical, plasma, genetic, and imaging features. Models were trained using RF.

To evaluate the additive predictive value of genetic information, we compared the performance of APOE genotype and PRS in predicting cerebral Aβ burden within the ADNI cohort (Fig. 6). In the baseline model comprising clinical and plasma biomarkers, inclusion of APOE modestly improved the prediction ($R^2$ from 0.515 to 0.527). In contrast, replacing APOE with PRS resulted in a greater increase ($R^2$ = 0.571; $\Delta R^2$ = +0.056 vs. +0.012), along with a substantial reduction in mean squared error (MSE: from 983.1 to 869.2). A detailed comparison is provided in Supplementary Materials(table .4).

This advantage of PRS remained consistent in more complex models integrating sMRI and SC features. Although the performance gap narrowed in the full multimodal model ($R^2$ = 0.637 for PRS vs. 0.617 for APOE), PRS consistently outperformed APOE across feature combinations, particularly in lower-dimensional settings. These findings suggest that PRS captures a broader spectrum of genetic risk beyond the APOE locus, offering enhanced predictive utility for Aβ burden despite the additional cost of genome-wide genotyping.

## 3.6 Sensitivity Analysis

To assess the robustness of the proposed framework, we performed sensitivity analyses across three methodological dimensions: PRS computation, SC parcellation, and machine learning algorithm.

First, we tested a range of p-value thresholds ($p < 10^{-5}$ to $10^{-10}$) for PRS construction using the clumping-and-thresholding approach. All thresholds except $p < 10^{-5}$ improved prediction relative to the plasma-only baseline, with the best performance observed at $p < 10^{-6}$. These results highlight the importance of appropriate thresholding in leveraging genome-wide polygenic information for improved prediction. Second, we examined the influence of parcellation schemes on SC-based prediction. Using alternative atlases with 84 and 116 regions, both configurations yielded improved performance over the baseline, though minor differences were observed. These findings indicate that the predictive contribution of SC features is generally stable across different anatomical resolutions, and

that higher-resolution schemes may help delineate more disease-relevant connectivity patterns. Third, we evaluated algorithmic sensitivity by comparing CatBoost with the primary random forest models. Across all feature combinations, CatBoost consistently improved performance over the baseline, confirming the benefit of multimodal integration. While the magnitude of improvement associated with specific modalities varied slightly between algorithms, the overall trend remained stable, supporting the generalizability of our findings across learning frameworks. Together, these analyses demonstrate that the predictive gains achieved through integration of imaging and genetic features are robust to key modeling and design choices. Full results are provided in Supplementary Materials(table .5-7)

# 4 Disscusion

This study presents a modular, multimodal machine learning framework for non-invasive prediction of cerebral Aβ burden, developed using ADNI data and externally validated in the independent SILCODE cohort. In the ADNI cohort, the baseline model including clinical and plasma features achieved an $R^2$ of 0.515. Under the full multimodal configuration (Clinical + Plasma + MRI + SC + APOE), model performance improved to $R^2 = 0.617$. Replacing APOE genotype with genome-wide PRS in the same feature combination further increased performance to $R^2 = 0.637$. External validation in SILCODE confirmed consistent performance gains across modalities, highlighting the

additive predictive value of sMRI, SC, and genetic features, and suggesting that PRS may offer a more informative alternative to APOE genotype alone in multimodal models.

sMRI and SC features provided complementary structural information: sMRI captured global atrophy patterns via data-driven dimensionality reduction, while SC reflected Aβ-related disruptions in long-range white matter connectivity, particularly across the visual, control, and SalVentAttn networks. Genetic features further improved predictive performance: APOE ε4 conferred consistent gains across cohorts, whereas PRS yielded greater incremental value in ADNI, underscoring the enhanced sensitivity of genome-wide risk profiling. Nevertheless, PRS application remains limited by cost and standardization challenges, and although APOE genotyping is more feasible for clinical use, it may not fully match the predictive utility of PRS. Collectively, these multimodal features help compensate for the limited neuroanatomical specificity of plasma biomarkers and reinforce the value of integrative approaches for improving non-invasive Aβ burden estimation.

To ensure methodological rigor, all modeling steps—including feature selection, dimensionality reduction, and prediction—were nested within each fold of LOOCV, and SC features were extracted using CPM in training data only. Consistent improvements across ADNI and SILCODE support the robustness and generalizability of the framework.

Several limitations warrant consideration. First, performance discrepancies between cohorts reflects differences in plasma assays (p-tau217 in ADNI vs. p-tau181 in

SILCODE) [6]. Second, modest sample sizes (ADNI: n = 150; SILCODE: n = 111) may limit statistical power and generalizability. Third, SC features were derived using univariate connectome-based predictive modeling (CPM), which may not capture higher-order topological interactions and remains sensitive to noise. In addition, the use of moderate-resolution atlases (Schaefer-100 + MSA-16) facilitated interpretability but may have overlooked finer-grained network effects. The predictive framework relied on linear and random forest models, which, although interpretable, may underutilize the complexity of multimodal features; future work could incorporate deep learning architectures to improve performance. Future directions should also explore individualized parcellation schemes, graph neural networks[55] for SC feature extraction, and longitudinal modeling to track Aβ accumulation trajectories and cognitive decline.

In summary, this study introduces a biologically interpretable and generalizable framework for Aβ burden prediction using multimodal, non-invasive data. These findings provide a methodological foundation for early AD risk stratification and motivate future development of scalable screening tools across populations.

## Data availability

The data used in this study are publicly available from the Alzheimer's Disease Neuroimaging Initiative (ADNI; http://adni.loni.usc.edu) and the SILCODE study (upon reasonable request from the corresponding author.

## Code availability

The code used for data preprocessing, feature extraction, and model implementation is available at:https://github.com/bluedog-1/Non-invasive-prediction-of-amyloid-burden-in-Alzheimer-s-disease-via-multimodal-machine-learning.


1. The neuropathological diagnosis of Alzheimer's disease | Molecular Neurodegeneration | Full Text. https://molecularneurodegeneration.biomedcentral.com/articles/10.1186/s13024-019-0333-5.
2. Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med* **8**, 595–608 (2016).
3. Jack, C. R. *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* **14**, 535–562 (2018).

4. Nakamura, A. *et al.* High performance plasma amyloid-β biomarkers for Alzheimer's disease. *Nature* **554**, 249–254 (2018).

5. Nabizadeh, F. *et al.* Plasma neurofilament light levels correlate with white matter damage prior to Alzheimer's disease: results from ADNI. *Aging Clin Exp Res* **34**, 2363–2372 (2022).

6. Palmqvist, S. *et al.* Discriminative Accuracy of Plasma Phospho-tau217 for Alzheimer Disease vs Other Neurodegenerative Disorders. *JAMA* **324**, 772–781 (2020).

7. Zhang, M. *et al.* Relationship between topological efficiency of white matter structural connectome and plasma biomarkers across the Alzheimer's disease continuum. *Human Brain Mapping* **45**, e26566 (2024).

8. Jack, C. R. *et al.* Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol* **9**, 119–128 (2010).

9. Racine, A. M. *et al.* Associations between white matter microstructure and amyloid burden in preclinical Alzheimer's disease: A multimodal imaging investigation. *NeuroImage: Clinical* **4**, 604–614 (2014).

10. Huang 黄伟杰, W. *et al.* Individual Variability in the Structural Connectivity Architecture of the Human Brain. *J Neurosci* **45**, e2139232024 (2025).

11. Liu, C.-C., Liu, C.-C., Kanekiyo, T., Xu, H. & Bu, G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat Rev Neurol* **9**, 106–118 (2013).

12. Leonenko, G. *et al.* Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. *Nat Commun* **12**, 4506 (2021).

13. Ramanan, V. K. *et al.* Genetic risk scores enhance the diagnostic value of plasma biomarkers of brain amyloidosis. *Brain* **146**, 4508–4519 (2023).

14. Karlsson, L. *et al.* Machine learning prediction of tau-PET in Alzheimer's disease using plasma, MRI, and clinical data.

15. Rajbanshi, B. *et al.* Clinical and neuropathological associations of plasma Aβ42/Aβ40, p-tau217 and neurofilament light in sporadic frontotemporal dementia spectrum disorders. *Alzheimers Dement (Amst)* **17**, e70078 (2025).

16. Qiu, S. *et al.* Multimodal deep learning for Alzheimer's disease dementia assessment. *Nat Commun* **13**, 3404 (2022).

17. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization - PubMed. https://pubmed.ncbi.nlm.nih.gov/20042704/.

18. Alfaro-Almagro, F. *et al.* Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage* **166**, 400–424 (2018).

19. Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).

20. FreeSurfer - ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S1053811912000389?via%3Dihub.

21. Tournier, J.-D. *et al.* MRtrix3: A fast, flexible and open software framework for medical image processing and visualization. *NeuroImage* **202**, 116137 (2019).

22. Li, X., Morgan, P. S., Ashburner, J., Smith, J. & Rorden, C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods* **264**, 47–56 (2016).

23. Andersson, J. L. R. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* **125**, 1063–1078 (2016).

24. Schilling, K. G. *et al.* Synthesized b0 for diffusion distortion correction (Synb0-DisCo). *Magnetic Resonance Imaging* **64**, 62–70 (2020).

25. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging - PubMed. https://pubmed.ncbi.nlm.nih.gov/24038308/.

26. (ISMRM 2016) A novel iterative approach to reap the benefits of multi-tissue CSD from just single-shell (+b=0) diffusion MRI data. https://archive.ismrm.org/2016/3010.html.

27. Tournier, J.-D., Calamante, F. & Connelly, A. Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage* **35**, 1459–1472 (2007).

28. Jeurissen, B. & Szczepankiewicz, F. Multi-tissue spherical deconvolution of tensor-valued diffusion MRI. *NeuroImage* **245**, 118717 (2021).

29. Smith, R. E., Tournier, J.-D., Calamante, F. & Connelly, A. Anatomically-constrained tractography: Improved diffusion MRI streamlines tractography through effective use of anatomical information. *NeuroImage* **62**, 1924–1938 (2012).

30. Schaefer, A. *et al.* Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cerebral Cortex* **28**, 3095–3114 (2018).

31. Tian, Y., Margulies, D. S., Breakspear, M. & Zalesky, A. Hierarchical organization of the human subcortex unveiled with functional connectivity gradients. *Nature Neuroscience* **23**, 1421–1432 (2020).

32. Landau, S. M. *et al.* Amyloid deposition, hypometabolism, and longitudinal cognitive decline. *Annals of Neurology* **72**, 578–586 (2012).

33. Klunk, W. E. *et al.* The Centiloid Project: standardizing quantitative amyloid plaque estimation by PET. *Alzheimers Dement* **11**, 1-15.e1–4 (2015).

34. Chen, H.-J. *et al.* Amyloid pathology related to aberrant structure-function coupling of brain networks in Alzheimer's disease: insights from [18F]-florbetapir PET imaging. *Eur J Nucl Med Mol Imaging* (2025) doi:10.1007/s00259-025-07172-8.

35. Clark, C. M. *et al.* Use of florbetapir-PET for imaging beta-amyloid pathology. *JAMA* **305**, 275–283 (2011).

36. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575 (2007).

37. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

38. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

39. Rubinacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat Genet* **53**, 120–126 (2021).

40. PRSice-2: Polygenic Risk Score software for biobank-scale data | GigaScience | Oxford Academic. https://academic.oup.com/gigascience/article/8/7/giz082/5532407.

41. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet* **51**, 404–413 (2019).

42. Prakash, R. S. *et al.* A whole-brain functional connectivity model of Alzheimer's disease pathology. *Alzheimers Dement* **21**, e14349 (2025).

43. Varoquaux, G. Cross-validation failure: Small sample sizes lead to large error bars. *Neuroimage* **180**, 68–77 (2018).

44. Fjell, A. M. *et al.* One-year brain atrophy evident in healthy aging. *J Neurosci* **29**, 15223–15231 (2009).

45. A review of feature reduction techniques in neuroimaging - PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC4040248/.

46. Finn, E. S. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *nature NEUROSCIENCE* **18**, (2015).

47. Shen, X. *et al.* Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nat Protoc* **12**, 506–518 (2017).

48. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. in *Advances in Neural Information Processing Systems* vol. 30 4765–4774 (2017).

49. Wang, J. *et al.* Diagnostic accuracy of plasma p-tau217/Aβ42 for Alzheimer's disease in clinical and community cohorts. *Alzheimers Dement* **21**, e70038 (2025).

50. Phases of Aβ-deposition in the human brain and its relevance for the development of AD | Neurology. https://www.neurology.org/doi/10.1212/WNL.58.12.1791.

51. Earliest accumulation of β-amyloid occurs within the default-mode network and concurrently affects brain connectivity | Nature Communications. https://www.nature.com/articles/s41467-017-01150-x.

52. Amyloid β-based therapy for Alzheimer's disease: challenges, successes and future | Signal Transduction and Targeted Therapy. https://www.nature.com/articles/s41392-023-01484-7.

53. In Vivo Tau, Amyloid, and Gray Matter Profiles in the Aging Brain - PubMed. https://pubmed.ncbi.nlm.nih.gov/27413148/.

54. Grothe, M. J., Teipel, S. J., & Alzheimer's Disease Neuroimaging Initiative. Spatial patterns of atrophy, hypometabolism, and amyloid deposition in Alzheimer's disease correspond to dissociable functional brain networks. *Hum Brain Mapp* **37**, 35–53 (2016).

55. Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage* **169**, 431–442 (2018).