**Title:** Neuropathologically validated MRI to tau PET synthesis via Covariate-modulated attention networks.

**Authors:** Marcell Borhi[1], Rita Qiuran Lyu[1], William J Jagust[2], Jingshen Wang[1]*, Joseph Giorgio[1,2,3]*, for the Alzheimer's Disease Neuroimaging Initiative[†]

* Contributed equally as senior authors

Affiliations:

[1] Division of Biostatistics, School of Public Health, University of California, Berkeley, California, USA

[2] Department of Neuroscience, University of California, Berkeley, California, USA, 94720

[3] School of Psychological Sciences, College of Engineering, Science and the Environment, University of Newcastle, Newcastle, New South Wales, Australia, 2308

Correspondence: Dr. Joseph Giorgio, jgiorgio@berkeley.edu

[†] 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. A complete listing of ADNI investigators can be found at:

http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

# Abstract

Tau PET is a powerful tool to assess tau pathology in vivo; however, in comparison to MRI, its development is more recent, is rarely available at scale, and substantially more difficult to acquire. Here, we present Covariate-Modulated Attention UNet (CoMA-UNet) to synthesize subject-specific 3D tau PET from T1 MRI while incorporating in the synthesis procedure readily available covariates. Across six external validation datasets, CoMA-UNet reproduced regional patterns of tau PET uptake showing strong agreement with true PET that was generalizable across tracers. Next, we submitted the synthetic tau PET to a series of downstream clinically relevant tasks. First, MMSE associations between the synthetic tau PET were statistically indistinguishable from true PET. Second, the synthetic tau PET achieved out-of-sample diagnostic classification of dementia with an AUROC=0.99. Third, out-of-sample synthetic tau PET tracked longitudinal progression with subject-level slopes closely matching true PET. Fourth, in two independent autopsy cohorts, voxel wise synthetic tau PET images closely followed neuropathologically defined Braak-stages. These findings demonstrate that the novel CoMA-UNet MRI-based synthesis augmented with covariate information can approximate tau PET with sufficient accuracy for downstream scientific and clinical applications.

# Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide[1,2] characterized by progressive loss of cognitive and functional abilities[3]**.** The neuropathological hallmarks of AD are extracellular β-amyloid (Aβ) plaques and intracellular neurofibrillary tangles composed of hyperphosphorylated tau, with the latter showing the tightest coupling to symptom severity, atrophy, and clinical progression[4–6]. To capture these neuropathological hallmarks PET imaging agents have been developed to measure both Aβ and

tau in-vivo[7]. Despite the uptake of these imaging technologies in recent years they are relatively new in the research field, with tau-PET imaging only available for about a decade. Accordingly, to leverage the decades of AD biobanking and clinical trials pre-dating tau-PET, novel methods are required to infer the presence and severity of tau pathology in legacy data.

In typical sporadic late onset AD, tau aggregates emerge and spread in a stereotyped sequence captured by the Braak staging framework, beginning in transentorhinal/entorhinal cortex (I/II), advancing to limbic structures (III/IV), and culminating in neocortical association areas (V/VI)[8]. PET-based adaptations of Braak staging recapitulate this topography in vivo, enabling quantitative modeling of disease severity and progression from living patients[9]. Modern biological definitions of AD are established by amyloid (A), tau (T) and neurodegeneration (N) in an ATN framework[3]. Among tau biomarkers, PET imaging with radiotracers such as [$^{18}$F]-flortaucipir (AV-1451), [$^{18}$F]-MK-6240, and [$^{18}$F]-PI-2620 provide unique whole-brain maps of regional tau burden, resolving voxel-level spatial patterns aligned with Braak stages and with strong links to cognition[10–12]. While CSF and plasma phosphorylated tau (e.g., p-tau217) are increasingly accurate for detecting abnormal Aβ and tau and for forecasting decline[13,14], fluid biomarkers lack spatial specificity; by contrast, tau PET yields topographic information critical for image-based analyses, staging, and anatomically informed treatment monitoring[9]. Despite these strengths, tau PET remains costly and logistically complex due to challenges related to radiochemistry and limited scanner access. As such, tau-PET is sparsely available compared to MRI, thus limiting routine clinical use and leaving tau unmeasured for many individuals and many historical datasets.

One potential way to infer underlying tau-PET levels when it is not available is through the use of tau-PET surrogates. Existing surrogate methodologies typically predict regional summary values rather than full voxel-wise volumes[15–17]; limiting voxel-wise analyses, aggregation with existing

volumetric tau-PET data, and downstream applications that depend on spatial fidelity. Recently, volumetric synthesis of tau PET using MRI has been attempted[18] however these previous approaches have disregarded critical disease related covariate information in the synthesis problem. To address these previous limitations, novel approaches are required to seamlessly integrate predictive covariate information in the synthesis of tau-PET volumetric data from MRI. Furthermore, given the variable access to covariate information (e.g. nascent plasma markers for tau pathology or genetics) this novel formulation must flexibly handle covariate sets with varied missingness.

Here, we present CoMA-UNet, a covariate-conditioned deep learning model for MRI-to-tau synthesis that yields anatomically consistent 3D tau PET volumes in native space. In contrast to regional or tabular surrogates, this voxel-wise formulation aims to preserve disease-relevant topology while remaining compatible with real-world clinical and legacy data constraints, thus maintaining flexibility in downstream tasks. We evaluate the fidelity of the synthetic tau-PET across diverse datasets and tracers. In particular, we highlight the fidelity of the image based on ground truth neuropathological staging and examine image utility in a range of downstream tasks including re-stratification of clinical trials based on anticipated levels of baseline tau pathology. Together, we present CoMA-UNet as an effective method to assess in-vivo tau pathology when tau PET is missing.

# Results

## Participants and Datasets

Seven distinct datasets taken from three independent cohorts were used in the training and out-of-sample validation of the CoMA-UNet framework. Training and internal cross-validation cohorts were drawn from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Longitudinal

Evaluation of Amyloid Risk and Neurodegeneration (LEARN) Study subset of the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) Study (ADNI-A4 training; n=1250 MRI-tau PET image pairs; 684 clinically unimpaired, 228 clinically impaired). These participants had paired T1-weighted MRI and [$^{18}$F]flortaucipir (FTP) tau PET imaging as well as thirteen covariates grouped into demographic (age, sex, education, marital status, race, height, weight), vital sign (systolic and diastolic blood pressure, pulse, respiratory rate, temperature), genetic (APOE4 allele count) features, and where available, plasma biomarkers (p-tau217, NfL, GFAP, A$\beta$40, and A$\beta$42). Validation within the ADNI-A4 training set was performed using five-fold cross-validation, with all presented results from the held-out test folds. Among external validation datasets, varying degrees of overlapping MRI and co-variate or ground truth tau PET data was available. The external validation datasets included (a) 100 ADNI and 100 A4 held out participants (cross-sectional validation; n=200; 81.5% with plasma ptau-217 available), (b) Held out ADNI participants with longitudinal tau-PET (Longitudinal validation; n=126 subjects, n=298 observations, with 4.0% plasma ptau-217 available), (c) The National Alzheimer's Coordinating Center (NACC) Standardized Centralized Alzheimer's & Related Dementias Neuroimaging (SCAN) dataset (NACC-SCAN; n=1291 subjects, 1409 observations, with 0% plasma ptau-217 available), (d) Treatment and placebo arms from the A4 clinical trial excluding the participants from the A4-LEARN present in the ADNI-A4 training data (A4 validation; n = 795, with 91.2% plasma ptau-217 available), and two datasets with paired neuropathology assessment, (e) ADNI with neuropathology (ADNI-NP validation; n=99, with 99% available ptau-217), and (f) NACC participants with non-SCAN-compliant MRI imaging (NACC non-SCAN; n=475, with 0% plasma ptau-217) (Table 1). All results reported in the external validation datasets are generated from the fixed model trained on the complete ADNI-A4 training dataset.

*Table 1* Cohort-level demographics, clinical variables, and imaging availability for all datasets included in this study. Values represent means (±SD) for continuous variables and counts (percent) for categorical variables.

| **Variable** | **Training** | **External Validation** |
|---|---|---|

| Cohort | ADNI-A4 training | Cross-sectional validation | Longitudinal validation | NACC-SCAN | NACC non-SCAN | A4 validation | ADNI-NP validation |
|---|---|---|---|---|---|---|---|
| n | 1250 | 200 | 126 | 1409 | 475 | 795 | 99 |
| Age mean (std) | 72.93 (7.14) | 72.24 (6.22) | 77.1 (7.4) | 72.33 (8.13) | 79.50 (12.65) | 71.9 (4.9) | 77.47 (7.84) |
| Sex n (Male%) | Male: 652 (52.2%) | Male: 85 (42.5%) | Male: 60 (47.6%) | Male: 470 (48.8%) | Male: 252 (53.1%) | Male: 315 (39.6%) | Male: 70 (73.7%) |
| Education mean (std) | 16.40 (2.59) | 16.43 (2.63) | 16.4 (2.6) | 16.33 (3.45) | 15.65 (8.40) | 15.7 (2.8) | 15.73 (4.06) |
| Aβ Status (Positive %) | 636 (51%) | 47 (47.5%) | 55 (52.4%) | N/A | N/A | N/A | N/A |
| APOE4 count | zero: 479 (58.8%), one: 299 (36.7%), two: 37 (4.5%) | zero: 82 (56.6%), one: 56 (38.6%), two: 7 (4.8%) | zero: 72 (62.1%), one: 41 (35.3%), two: 3 (2.6%) | zero: 517 (60.1%), one: 285 (33.1%), two: 58 (6.7%) | zero: 251 (55.2%), one: 169 (37.1%), two: 35 (7.7%) | zero: 322 (40.5%), one: 473 (59.5%), two: 0 (0%) | N/A |
| MMSE mean (std) | 28.21 (2.55) | 28.45 (1.62) | 28.0 (3.03) | 29.09 (2.00) | N/A | 28.9 (3.5) | N/A |
| pT217_F mean (std) | 0.29 (0.25) | 0.26 (0.18) | 0.272 (0.147) | N/A | N/A | 0.279 (0.16) | 0.775 (0.714) |
| Cognitively Impaired (%) | 228 (25%) | 25 (15.3%) | 42 (39.3%) | 427 (64.7%) | 342 (72%) | 0 (0%) | 0 (0%) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Braak Stages | N/A | N/A | N/A | N/A | 0: 30, I: 26, II: 53, III: 62, IV: 63, V: 69, VI: 172 | N/A | 0: 1, I: 7, II: 11, III: 4, IV: 6, V: 48, VI: 22 |
| FTP tracer n (%) | 1250 (100%) | 200 (100%) | 298 (100%) | 958 (74.2%) | 475 (100%) | N/A | 99 (100%) |

## Cross-validation and Cross-sectional Performance of Covariate-modulated Attention UNet

CoMA-UNet was built to incorporate AD related covariates into the voxel-to-voxel Attention-UNet in three stages. First, covariates are used in a series of CatBoost models estimating regional tau burden, temporal meta-ROI tau (MetaTempTau), Aβ positivity, and MMSE. In the second stage, the deep learning model integrates MetaTempTau, Aβ status, and MMSE estimates through a conditional decoder. In the final stage, regional predicted tau estimates are used to modulate the decoded image from the conditioned attention-UNet output. In doing so, disease related covariate information is used to conditionally up sample based on predicted AD severity (i.e. MetaTempTau, Aβ, and cognitive impairment), while also embedding regionally specific covariate related signal thorough the spatially constrained modulation of the up sampled image (Figure 1C). Critically, through an initial imputation model prior to the CatBoost prediction, CoMA-UNet can seamlessly handle variable missingness in the covariate feature set (i.e. when plasma markers are unavailable).

Synthetic tau PET images were generated using the CoMA-UNet framework trained and internally validated across five-folds of the ADNI-A4 dataset (n=1250). Performance metrics include the average correlation (Pearson's) coefficient ($Corr_{AVG}$) of predicted and ground truth tau in selected

regions of interest (ROIs), average voxel-wise mean absolute percentage error (MAPE) and mean absolute error (MAE), and the structural similarity index (SSIM). Across the five folds, the converged models achieved an average held out prediction performance of $Corr_{AVG}$=0.5180 ($\pm$ 0.034), MAPE =10.88% ($\pm$ 1.23), MAE =0.143 ($\pm$ 0.021), and a SSIM=0.823 ($\pm$ 0.017) (Figure 1A). Fidelity was assessed across all held-out cross-validation folds in the ADNI-A4 training dataset by correlating synthetic and actual tau PET SUVRs in key ROIs (Supp Figure 1). In key AD related regions, a high cross-validated correlation was achieved in inferior temporal (r = 0.63), middle temporal (r = 0.63), as well as key Parietal lobe regions, including the inferior parietal (r = 0.64) and precuneus (r = 0.63). The trained model was then applied to the fully held out cross-sectional validation dataset (n = 200). CoMA-UNet attained consistent performance, with $Corr_{AVG}$ = 0.462, MAPE = 10.87%, MAE = 0.131, and SSIM=0.79 (Figure 1B).

We next contrasted the goodness-of-fit results of CoMA-UNet to alternate model architectures, including Attention UNet, UNet, UNETR, and UNETR variant with attention gates in the Decoder (Supp Table 2). The CoMA-UNet produced substantial gains over all alternative architectures, which typically achieved $Corr_{AVG}$ < 0.15 across the five folds. The Attention UNETR, and Attention UNET models had the next lowest average regional MAPEs with 12.96% ($\pm 2.21$), and 13.05% ($\pm 1.43$) in the cross-validation, respectively. The SSIM of all architectures were relatively high, with the highest average SSIM being 0.7506 achieved by the Attention UNet.

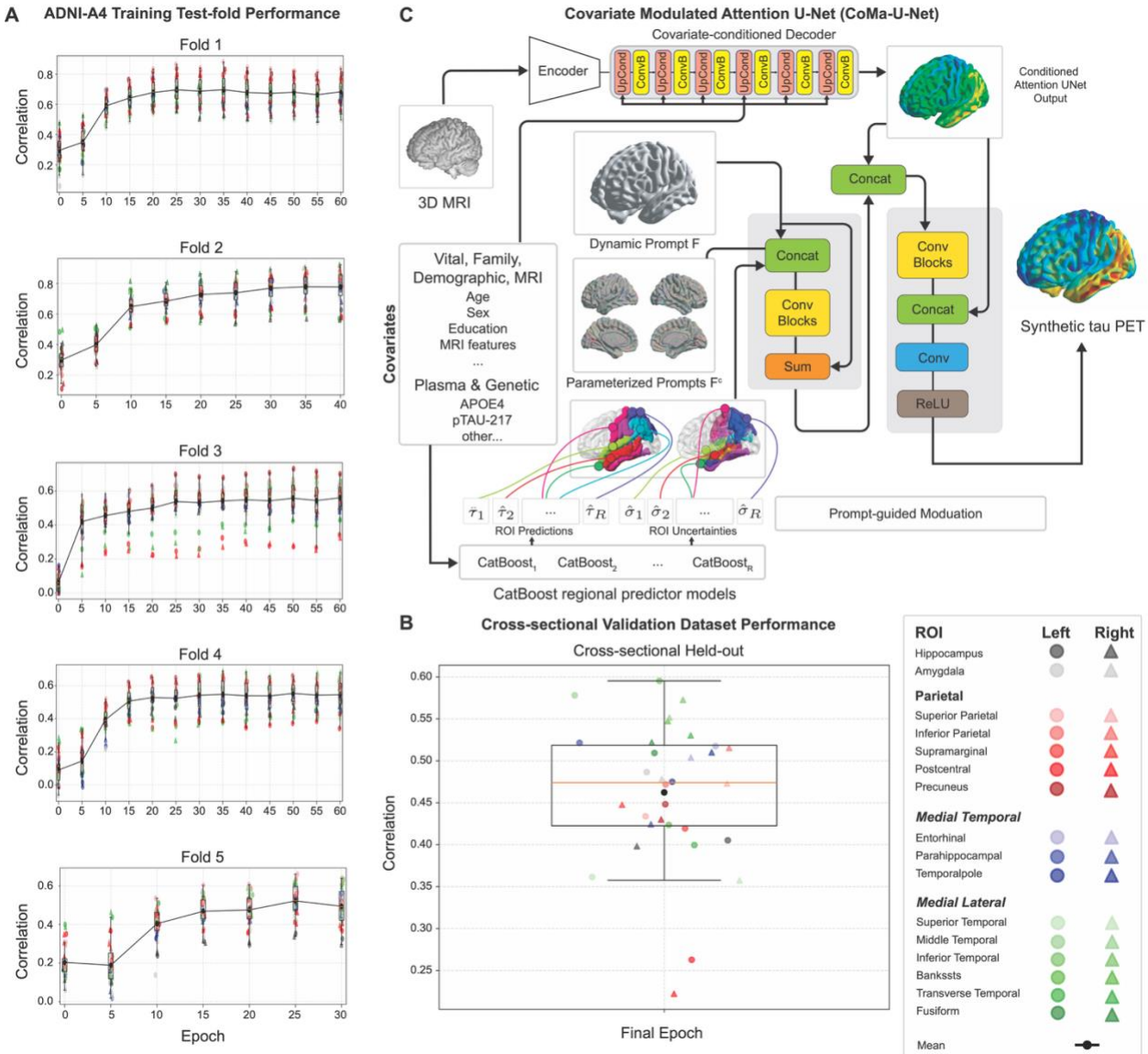


*Figure 1* (A) Training curves for all five folds of cross-validation within the ADNI-A4 training dataset, showing correlation between predicted and ground-truth regional FTP SUVR values in the held-out test partition at each test epoch. Colored points correspond to individual regions of interest (ROIs), while the black line indicates the average correlation across ROIs (CorrAVG). Across all folds, model performance steadily improves with training, with both CorrAVG and ROI-specific correlations showing consistent convergence. (B) Cross-sectional validation dataset performance for the final trained model, showing ROI-level correlations between predicted and observed regional FTP SUVR in the independent test set. Higher correlation reflects more accurate regional tau estimation, with temporal and parietal regions showing the strongest predictive agreement, recapitulating cross-validation findings. (C) Overview of the Covariate-Modulated Attention U-Net (CoMA-UNet) architecture. A 3D MRI is encoded through a convolutional encoder, while subject-level covariates (vital, demographic, MRI-derived features, genetic variables, and plasma biomarkers) are transformed into dynamic, parameterized prompts. These prompts condition the decoder via concatenation and covariate-modulated attention blocks, enabling voxel-wise tau prediction that adapts to individual covariate profiles. The model outputs a full synthetic FTP tau PET volume.

a

Using the CoMA-UNet model trained on the ADNI-A4 training dataset, we evaluated the performance of the model to track longitudinal changes in tau PET on the Longitudinal validation dataset consisting of held-out ADNI subjects with longitudinal ground truth tau PET images (n = 126 subjects, with total 298 observations). Strong regional correlations were observed, with the model attaining a $Corr_{AVG}$=0.576 among all samples (Supp figure 2).

Next, we examined how closely $\widehat{MetaTempTau}$ tracked longitudinal changes using a linear mixed-effects framework. There was a strong correspondence between synthetic and true measures ($\Delta\chi^2 = 20.67, df = 2, p = 5.4 \times 10^{-6}$), as illustrated by close alignment between the line of best fit and the unity line (Supp Figure 3A). The fixed intercept and slope were $0.53 \pm 0.11$ and $0.47 \pm 0.09$ ($p < 0.001$ for both), respectively, confirming that within-subject increases in synthetic tau were closely coupled to within-subject increases in true tau with a slightly attenuated slope relative to the identity line. Subject-specific trajectories further illustrated this, with the population-average fit (red line) closely paralleling individual relationships (Supp Figure 3B). Together, these results demonstrate that $\widehat{MetaTempTau}$ captures both baseline burden and subtle progression at the subject level, opening the possibility of use as a quantitative surrogate to infer longitudinal tau accumulation when ground truth images are not available.

## Performance on SCAN-compliant MRIs from the NACC-SCAN dataset & generalizability to different radiotracers

We next used the NACC-SCAN dataset to assess the model's ability to synthesize tau PET without the inclusion of plasma ptau-217 as a covariate. This dataset consisted of entirely held out subjects from the NACC cohort with SCAN compliant tau PET (n=1409), and included a mix of tau PET tracers (n=958 FTP tau PET; n=451 MK-6240 tau PET), allowing us to also assess if

the model could capture variance in tau PET images acquired with an unseen radiotracer. Since CoMA-UNet was trained to synthesize FTP tau PET images that operate on a different scale from MK tau PET, we assessed predictive performance using Pearson correlation, excluding other metrics that are sensitive to scale differences.

We observed strong performance in both Pearson correlation and $R^2$ with both the FTP and MK-6240 tau PET (Supp Figure 4). In particular, the CoMA-UNet model achieved correlations of r=0.471 in the entorhinal, and an overall average of r=0.486 in the Temporal regions. Across both tracers, the model attained a $\text{Corr}_{\text{AVG}}$ of r=0.498. Despite not being trained on images with the MK-6240 radiotracer, CoMA-UNet maintained similar performance, indicating that it does not merely learn contrasts, but rather an underlying association. Notably, the model achieves these results in the absence of plasma biomarkers, underpinning the predictive capabilities of the model under different patient populations, even with few available covariates.

## Classification of CN, MCI and Dementia using Synthetic tau PET

We next evaluated whether the $\widehat{MetaTempTau}$ generated by CoMA-UNet could match the diagnostic utility of $MetaTempTau$ extracted from actual tau PETs. As all samples drawn from the A4 datasets had clinical diagnoses of CN, we limited this analysis to ADNI samples (n = 602; 7% with plasma ptau-217), avoiding any biases from population differences.

Within the five-folds of the ADNI-A4 training dataset, we evaluated diagnostic classification using only completely held-out (ADNI) subjects from each fold. For each fold, a logistic-regression classifier was trained on synthetic $\widehat{MetaTempTau}$ values from the remaining four folds and evaluated on that fold's held-out participants. Across the five folds, CN vs Dem classification achieved ROC-AUC values exceeding 0.80 in four of the five held-out subsets (Figure 2B).

Interestingly, classifiers using $\widehat{MetaTempTau}$ consistently achieved higher mean AUC scores than those using $MetaTempTau$ in all diagnostic comparisons across the five folds, but these improvements did not reach statistical significance (paired two-tailed t-tests on AUCs across five folds: $-2.71 \leq t \leq -1.53$, $0.054 \leq p \leq 0.19$; Supp Figure 5, 6). This indicates that synthetic tau PET can statistically match or even exceed classification capabilities of actual tau PETs.

We further evaluated the discriminative utility of the synthetic tau PETs in the completely held-out cross-sectional validation samples. To protect against bias resulting from population differences, we again only considered classification in the ADNI subjects (n= 63; 100% with plasma ptau-217). The logistic regression classifier using $\widehat{MetaTempTau}$ showed similar or substantially better performance in comparison to classifiers using $MetaTempTau$. In particular, the use of $\widehat{MetaTempTau}$ resulted in an AUROC=0.99 in CN vs Dem classification, compared to the 0.82 AUC score achieved by using $MetaTempTau$ (Figure 2D).

Finally, we evaluated the discriminative utility of the synthetic tau PET on the voxel level running voxel-wise two sample t-tests to assess mean difference between levels of clinical impairment. As before we excluded A4-LEARN subjects using the held-out ADNI synthetic tau PET volumes generated from the five-fold cross-validation procedure. We computed unthresholded voxel-wise Cohen's $d$ in the MNI-space normalized synthetic tau PETs for both the held out ADNI training dataset samples (Figure 2A) and cross-sectional validation samples (Figure 2C). Across both cohorts, we observed particularly large effect sizes ($d < -2$) in the Temporal and Parietal lobes for the CN vs Dem comparison, with more modest effect sizes for CN vs MCI and MCI vs DEM comparisons.

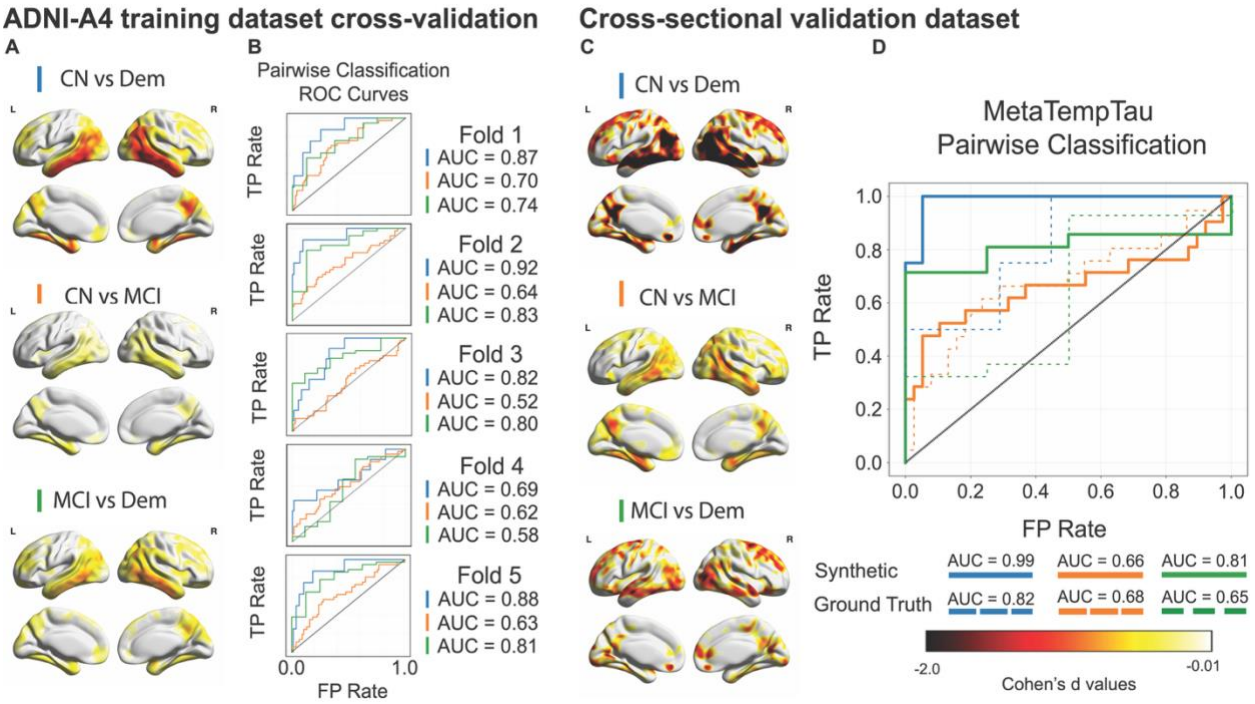


*Figure 2* (A) Voxel-wise Cohen's $d$ values within all held-out cross-validation samples in the ADNI-A4 training dataset show large effect sizes ($d < -0.8$) in the temporal regions for all comparisons. (B) ROC Curves for logistic regression classifiers using $MetaTempTau$ across each of the held-out samples across the five folds of the ADNI-A4 training dataset, showing consistent discriminative performance. (C) Voxel-wise Cohen's $d$ values for each pairwise classification within all held-out cross-sectional samples show very large effect sizes ($d < -2$) in the Temporal and Parietal regions for all comparisons. (D) A logistic regression classifier using synthetic $MetaTempTau$ (solid lines) achieves higher ROC-AUC scores than the classifier using the ground truth $MetaTempTau$ (dashed lines) in the CN vs Dem, and MCI vs Dem pairwise classification tasks.

## Cross sectional MMSE prediction

We next examined the regional association between synthetic tau PET and MMSE, and MRI thickness and MMSE in the cross-sectional validation dataset (n = 198 with available MMSE, 63.1% with plasma ptau-217) to assess how well synthetic tau PETs capture subject-wise levels of cognition. We then contrasted the differences in these associations between imaging measures and MMSE using Steiger's Z tests. This revealed that synthetic tau PET values had a stronger association with cognition in the temporal and parietal lobe regions than the MRI used to generate the synthetic tau PET images. When analyses were stratified by Aβ PET status (Aβ positivity defined as global cortical florbetapir SUVR $\geq 1.11$ using whole cerebellum as reference), the synthetic tau PETs within the Aβ positive subgroup showed significantly stronger association with

subject baseline MMSE scores than MRI volume (Benjamini–Hochberg correction, $q < 0.05$, $p_{FDR} = 0.0239$). This was observed in tau-PET AD related regions such as the banks of the superior temporal sulcus ($Z = -2.581$, $p_{uncorrected} = 0.0049 < 0.0239$), inferior parietal ($Z = -2.730$, $p_{uncorrected} = 0.0031 < 0.0239$), middle temporal ($Z = -2.823$, $p_{uncorrected} = 0.0023 < 0.0239$), and precuneus (-2.443, $p_{uncorrected} = 0.0072 < 0.0239$) (Figure 3A). At the voxel-level in spatially normalized MNI space, we observed strong correlations in the temporoparietal regions ($r < -0.5$), with weaker correlation in the frontal regions (Figure 3B) in Aβ positive individuals. Within the Aβ negative subgroup however, there were no differences in the association between MRI and Cognition and Synthetic tau and cognition, suggesting that the synthetic tau captured AD related variance in cognition (i.e. dependent on Aβ status). Steiger's *Z* values were not statistically significant after Benjamini & Hochberg Family-wise Error correction.

Finally, we found that within the cross-sectional validation dataset, the association between the $\widehat{MetaTempTau}$ and MMSE ($r = -0.442$, df = 196, 95% CI [-0.55, -0.32], $p < 0.001$; Figure 3D) did not differ significantly ($Z = 0.974$, df = 195, $p = 0.1650$) from the ground truth association between $MetaTempTau$ and MMSE ($r = -0.491$, df = 196, 95% CI [-0.59 , -0.38], $p < 0.001$; Figure 3C).

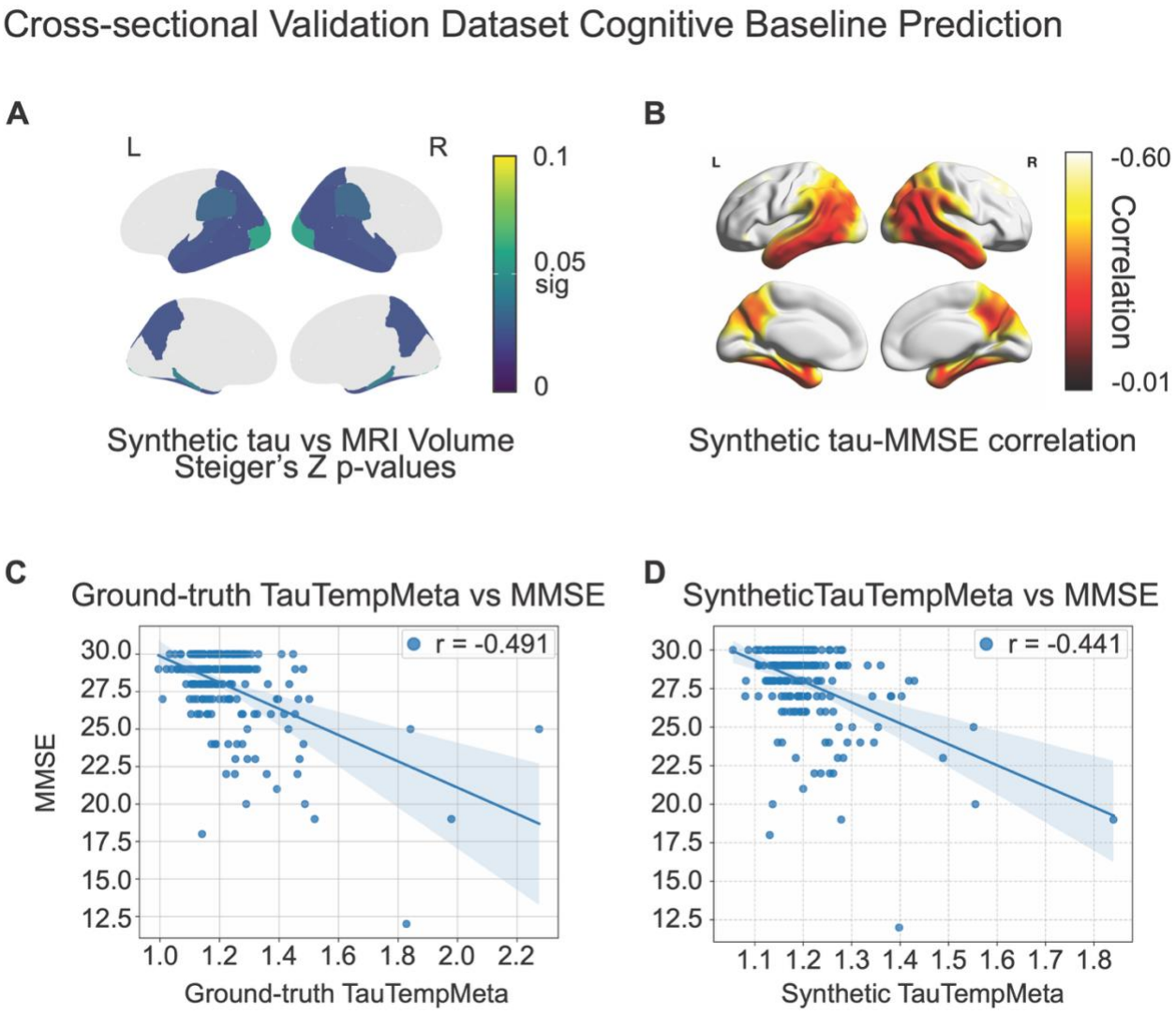


*Figure 3* (A) Benjamini & Hochberg Family-wise Error corrected p-values for Steiger's Z values between synthetic tau PET values and MRI volumetric measures within amyloid positive subpopulation of the cross-sectional validation dataset subjects. (B) Voxel-wise correlation between synthetic tau PET and MMSE among amyloid positive subjects. (C, D) Correlation between the ground-truth $MetaTempTau$ and MMSE vs $Meta\widehat{TempTau}$ and MMSE.

## Performance on A4 samples with only MRI observed

To evaluate the utility of the model in settings where MRI scans lack paired tau PET data, we re-stratified participants in the A4 validation dataset (n = 795) into three groups according to tertile thresholds of $Meta\widehat{TempTau}$ values derived from synthetic tau PET scans. Using these tertile assignments, we computed longitudinal PACC trajectories within the treatment and placebo arm using natural cubic splines consistent with the modeling approach in the A4 study[19]. We repeated

this same procedure using synthetic tau PETs generated without plasma ptau-217 as a covariate input to CoMA-UNet, yielding corresponding trajectories for the treatment and placebo groups, thereby probing the robustness of the model to missing covariate information.

Across all conditions, participants in the highest MetaTempTau tertile (T3) showed steeper cognitive decline relative to those in the lowest tertile (T1). In the treatment arm, differences in PACC trajectories were significant when using plasma biomarkers in the input (T3–T1 = –3.11 ± 0.64, t = –4.82, $p = 1.5 \times 10^{-6}$; T3–T2 = –2.42 ± 0.64, t = –3.79, $p = 1.5 \times 10^{-4}$; Figure 4B). These differences remained significant without plasma biomarker inputs between T3 and T1 ($t$ = -4.82, $p < 0.0001$), as well as between T2 and T3 ($t$ = -3.29, p-value = 0.001; Figure 4A). Similarly, in the placebo arm, participants in T3 also exhibited greater cognitive decline both when using plasma inputs (T3–T1 = –2.64 ± 0.47, t = –5.57, $p = 2.7 \times 10^{-8}$; T3–T2 = –2.37 ± 0.47, t = –5.05, $p = 4.6 \times 10^{-7}$; Figure 4D), and without plasma (T3–T1 = –1.91 ± 0.47, t = –4.09, $p = 4.4 \times 10^{-5}$; T3–T2 = –1.74 ± 0.49, t = –3.58, $p = 3.4 \times 10^{-4}$; Figure 4C).

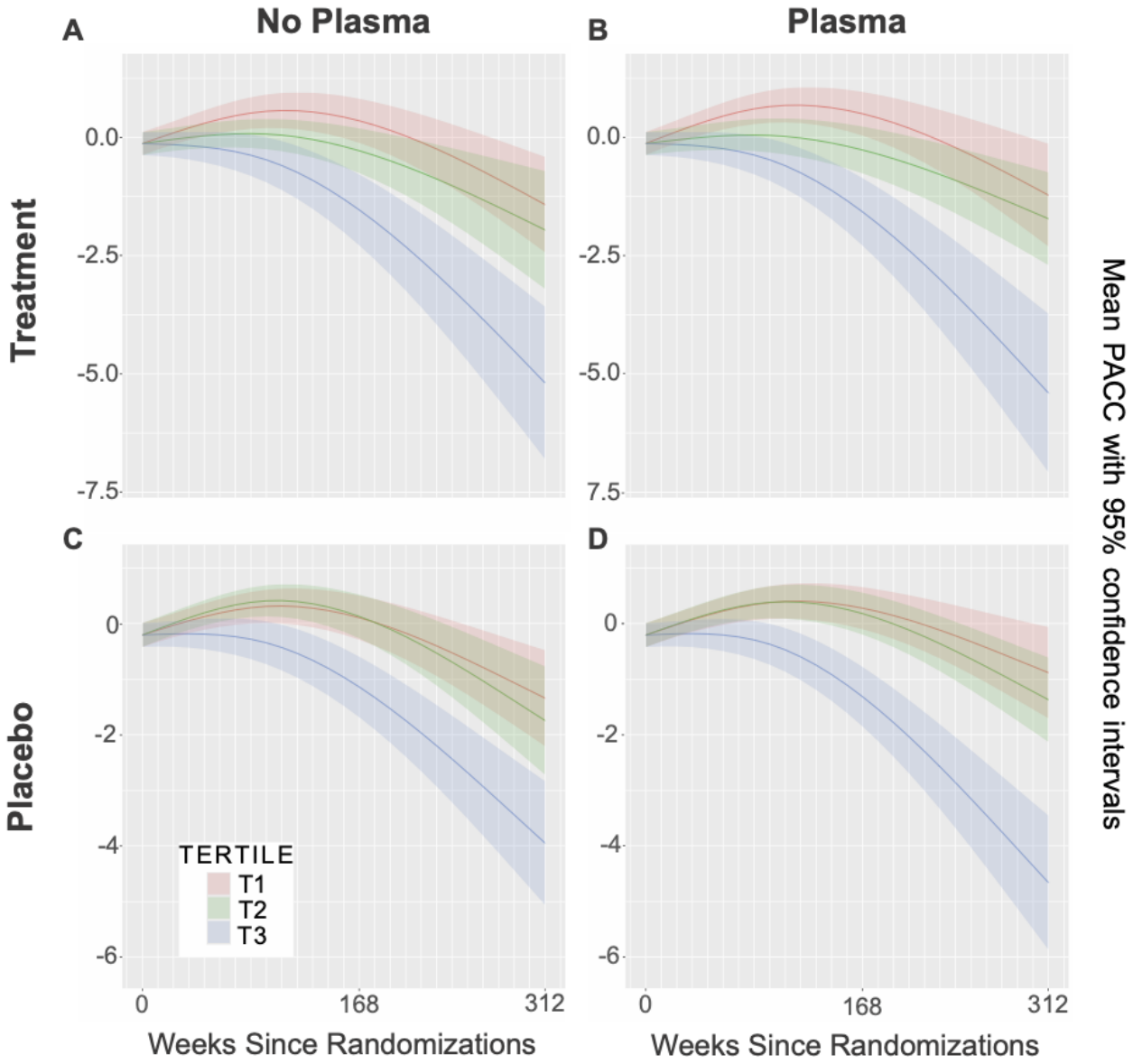


*Figure 4* Held-out A4 subjects stratified according to tertiles of average temporal meta ROI SUVR derived from synthetic tau PETs generated by the CoMA-UNet including plasma ptau-217 as a covariate (panels B and D), and in the absence of plasma ptau-217 as a covariate (panels A and C). Panels A and B display PACC score trajectories for each tertile among subjects that received treatment. Panels C and D display PACC score trajectories for each tertile among control (placebo) subjects.

## Synthetic tau PET association with Braak Stage neuropathology

We next assessed how well the synthetic tau PET images derived using the CoMA-UNet recapitulated underlying Braak stage pathology as assessed by post-mortem neuropathology assessment. Using the ADNI-NP validation dataset (n=99; 99% with plasma ptau-217 biomarkers), which consists of subjects with autopsy-confirmed Braak stage data, we evaluated

how faithfully the synthetic tau PET derived ante-mortem could track the gold standard neuropathological progression of tau. We grouped subjects by adjacent Braak stages: Braak Stage group I/II, III/IV, and V/VI and contrasted the voxel-wise uptake difference between adjacent tau stages. The analysis revealed a “spread” of tau from medial temporal lobe/hippocampus (stage I/II) to other temporal regions, posterior cingulate/retrosplenial cortex, parietal & frontal regions (stage III/IV), and finally higher neocortical regions (stage V/VI) recapitulating the underlying stages of tau neuropathology (Figure 5A, B). Ordinal regression of $\widehat{MetaTempTau}$ signal against Braak group confirmed these voxel-level findings, showing a monotonic increase in the predicted probability of higher Braak stage groups with increasing temporal tau burden ($\beta = 10.69, 95\%\ CI[7.25,14.12], z = 6.10, p < 0.0001$) (Figure 5C**;** Supp Figure 7, 8).

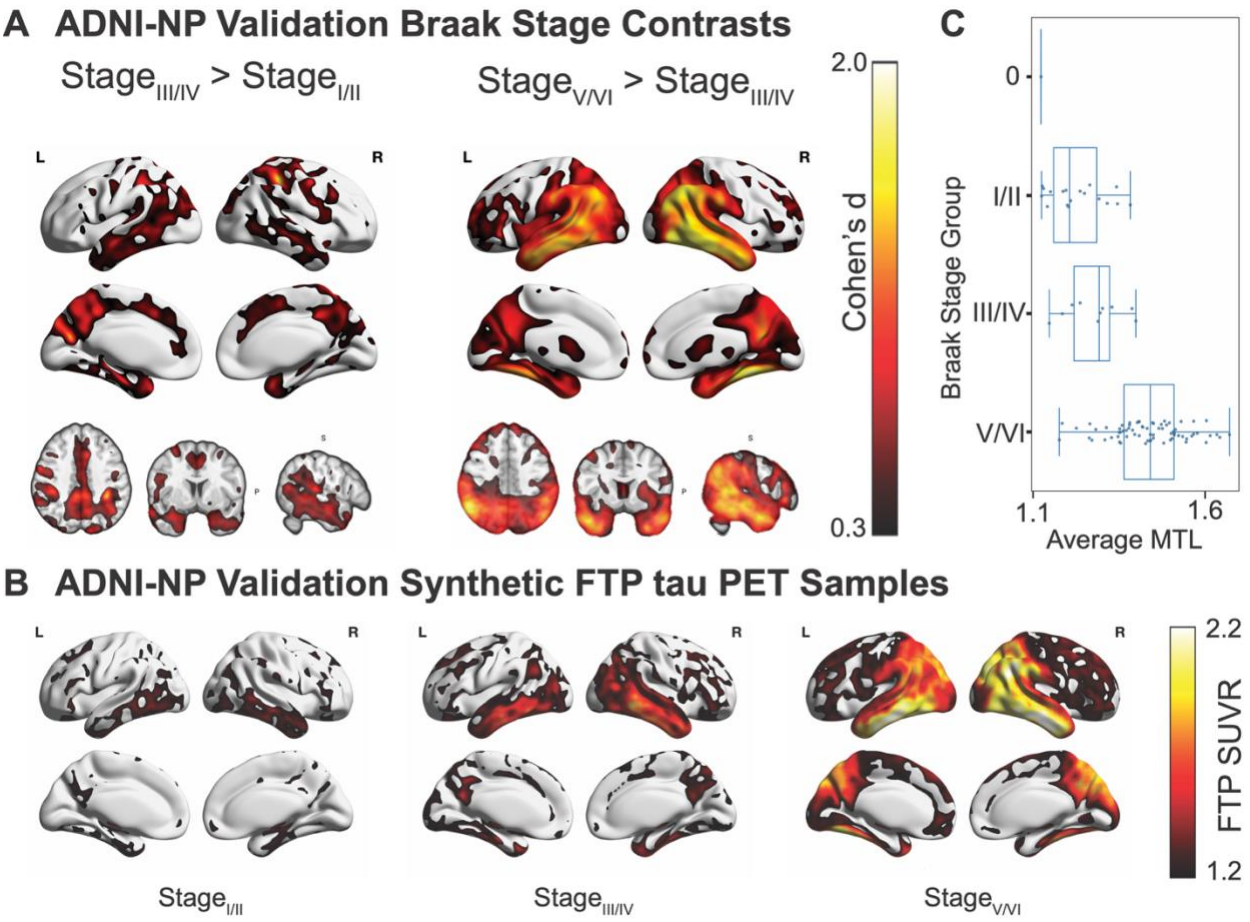


*Figure 5* Braak-stage effects in NACC-NP participants using synthetic (FTP) tau PET (A) Voxel-wise Cohen's $d$ maps comparing synthetic tau deposition between Braak stage I/II vs III/IV (left) and III/IV vs V/VI (right). The contrast between I/II and III/IV shows focal effects in medial and inferior temporal regions, whereas the contrast between III/IV and V/VI shows larger clusters across lateral temporal, parietal, and occipital cortex. (B) Representative synthetic FTP tau PET images from Braak Stage I/II, III/IV, and V/VI (left to right), showing low medial-temporal uptake in I/II, higher and more diffuse temporal uptake in III/IV, and the highest SUVR levels with broad cortical involvement in V/VI. (C) Boxplots of average MTL SUVR for each Braak Stage group. Median MTL SUVR increases from low values in I/II, to intermediate levels in III/IV, to highest values in V/VI, matching the uptake differences visible in Panel B.

## Synthetic tau PET generated using non-SCAN compliant MRIs from the NACC dataset

Finally, we assessed if CoMA-UNet could generate high fidelity synthetic tau PET in a heterogeneous sample with non-standardized MRI imaging protocols and no available plasma ptau-217 data. Using the NACC non-SCAN dataset, comprising non-SCAN-compliant MRIs (n = 475) with post-mortem assessment, we generated a synthetic tau PET volume for each subject using the CoMA-UNet. Across adjacent Braak stage contrasts, the synthetic images recapitulated

the expected neuropathological progression. Relative to Braak I/II, subjects in Braak III/IV exhibited significantly greater simulated uptake within the medial temporal lobe, within the Braak V/VI group there were further increases extending from the medial temporal lobe into temporal neocortex and association cortices (Figure 6A, B). Effect sizes were largest in temporal regions, with the III/IV versus V/VI comparison revealing a marked neocortical spread consistent with late-stage tau deposition (Figure 6A, B**)**. Complementing the voxel-wise contrasts, an ordinal regression of $\widehat{MetaTempTau}$ signal against Braak stage demonstrated a monotonic increase in the predicted probability of higher Braak stage groups with increasing temporal tau ($\beta = 2.96, 95\% CI [2.10, 3.81], z = 6.78, p < 0.001$), indicating that a simple summary of the synthetic images also tracks ordinal pathology (Figure 6C**;** Supp Figure 9, 10).

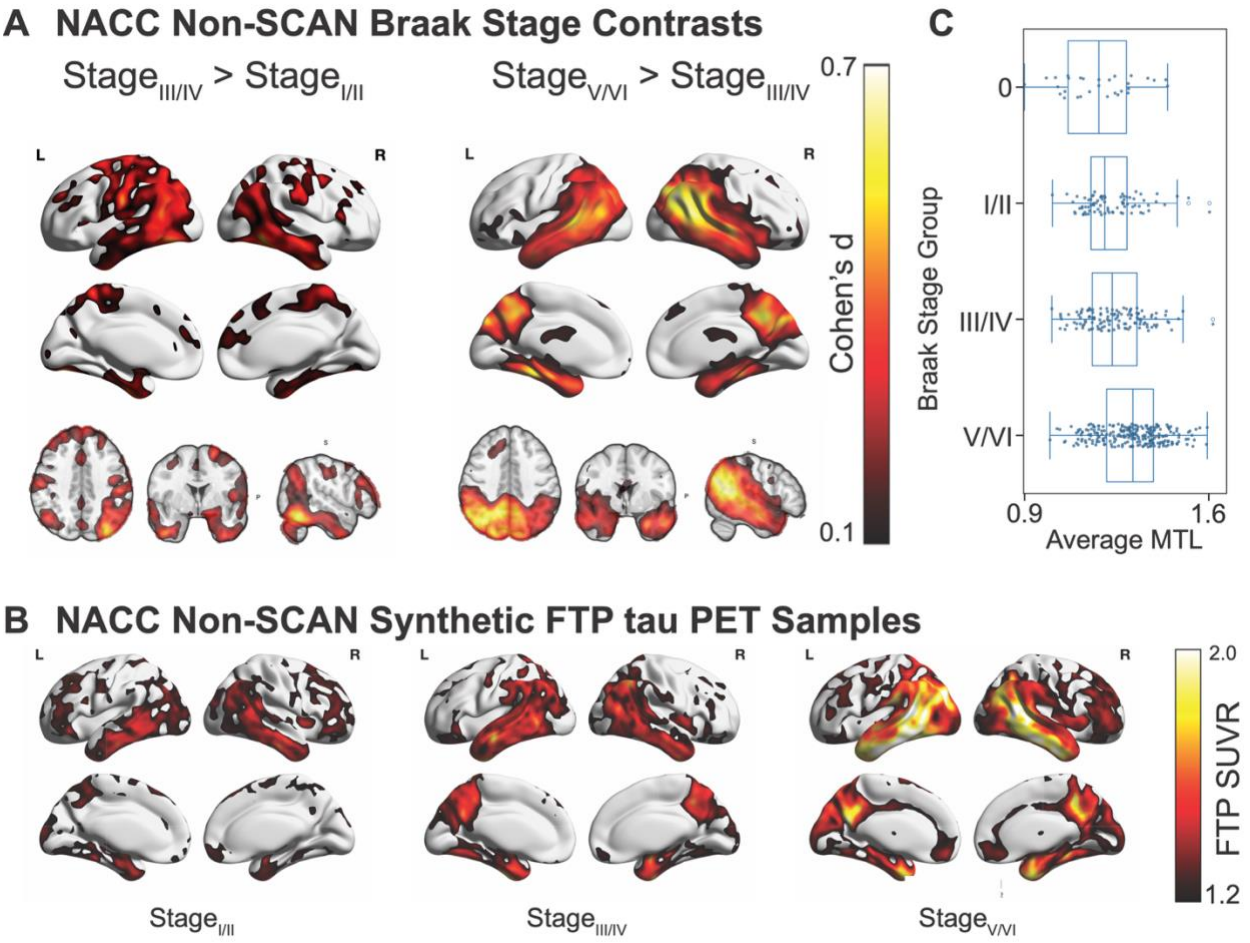


*Figure 6* Braak-stage effects in NACC non-SCAN participants using synthetic (FTP) tau PET. (A) Voxel-wise Cohen's $d$ maps showing regional differences in synthetic tau deposition between Braak Stage I/II vs. III/IV (left) and III/IV vs. V/VI (right). The I/II vs III/IV contrast shows strongest effects in medial temporal regions extending into inferior temporal cortex, whereas the III/IV vs V/VI contrast shows broad neocortical association involvement. (B) Representative synthetic tau PET images from individuals in Braak Stage groups I/II, III/IV, and V/VI (left to right), illustrating the progressive spatial spread of tau. (C) Boxplots of average meta-temporal (MTL) SUVR across Braak Stage groups (I/II, III/IV, V/VI), demonstrating monotonic increases in regional tau signal with advancing Braak stage.

# Discussion

Here we introduce CoMA-UNet, a novel approach to incorporate non-PET covariate information in an MRI to tau-PET image synthesis. We show that the synthetic images produced using the CoMA-UNet were of high fidelity and markedly outperformed existing MRI-to-PET models. Furthermore, we show that the CoMA-UNet produced synthetic images that generalized across cohorts and tracers while preserving cognition-linked signals sufficient to support downstream

clinical tasks. Critically, we provide an in-depth validation of the synthetic tau-PET images derived out-of-sample to the gold standard neuropathological Braak staging.

CoMA-UNet synthesizes full volumetric tau PET images from structural MRI by incorporating subject-level covariates from non-PET variables (e.g. demographics, genetics, plasma markers) into the decoder of its UNet component. Through covariate-parameterized prompts anatomical priors are combined with ROI-level tau estimates through auxiliary models, yielding anatomically coherent SUVR patterns and accommodating covariate missingness. Previous deep learning approaches for volumetric MRI-to-PET using classical U-Net variants[20–22], transformer backbones[23–26], generative-adversarial paradigms, or drift-diffusion omit explicit covariate conditioning, relying on pure image-to-image mappings[18,27–35]. Prediction of tau-PET from non-imaging inputs such as plasma biomarkers, or from coarse imaging features such as MRI-derived regional summaries has been increasingly explored[15–17], yet these approaches typically produce region-level or categorical estimates rather than full-resolution voxel-wise maps. In contrast, CoMA-UNet directly generates spatially resolved tau-PET representations, enabling finer-grained correspondence with ground-truth signal. This yielded a clear benefit over alternate convolutional and transformer-based architectures achieving substantially higher regional associations with ground truth tau PET (average ROI correlation [$Corr_{AVG}$] CoMA-UNet: r= 0.52, others: r<0.15). Moreover, CoMA-UNet showed strong generalizability in the external NACC cohort, capturing variance in an unseen tracer while outperforming other comparable modelling approaches on the same task[18]. Furthermore, we observed that longitudinal synthetic tau PET images derived using the CoMA-UNet architecture in a held-out sample faithfully tracked within subject change in tau accumulation in key AD regions overcoming challenges in the longitudinal modelling of tau PET [36,37]. Together, these findings demonstrate the potential of CoMA-UNet to generate high fidelity synthetic tau PET.

We show through a series of downstream tasks that the synthetic tau PET generated by CoMA-UNet preserves cognition-linked signals sufficient for diagnostic classification and capturing participant-level variance in cognition. Synthetic tau PET images showed strong voxel-wise association with MMSE in association cortices, reflecting established findings that neocortical tau best explains cognitive decline[3,38–40]. Regionally, this association was significantly stronger than that of regional MRI features with MMSE in held-out Aβ positive subjects, aligning with results suggesting that tau PET is more predictive of cognition than MRI[41–43]. Comparing correlations between model derived and ground truth temporal meta-ROI SUVR with MMSE showed no statistical difference and when extended to categorical diagnosis, classifiers based on synthetic temporal meta-ROI SUVR achieved accuracy and AUC values that matched or exceeded those using ground truth temporal meta-ROI SUVR. This benchmarks synthetic tau PET generated by CoMA-UNet against prior work showing the utility of tau PET in discriminating different levels of clinical impairment[10,44].

We present a potential use of synthetic tau PET imaging to retrospectively re-stratify clinical trials based on heterogeneity of baseline tau when tau PET was not available. When re-stratifying Aβ -positive cognitively normal older participants from the A4 solanezumab trial[19,45] into tertiles of synthetic temporal meta-ROI SUVR we revealed significant divergence in PACC trajectories, with participants in the highest tertile showing the steepest cognitive decline. This is of clinical relevance as the successful TRAILBLAZER trial for Donanemab showed when participants were stratified by baseline tau burden treatment benefits varied[46]. Together, this offers an interesting avenue to apply synthetic tau PET to other completed clinical trials for AD, re-stratifying based on heterogeneity in baseline tau-PET and reassessing treatment effects. In doing so, retrospective subgroup discovery, counterfactual enrichment that prioritizes intermediate tau burden, and

hypothesis testing on how tau modulates effect sizes, can all be achieved using MRI-only datasets for future tau-informed trial design.

We provide the most compelling account of the biological validity of tau-PET surrogates to date showing a tight coupling to the underlying neuropathological staging of AD. In two highly heterogenous held-out cohorts with neuropathology, voxel-wise Braak-group contrasts (I/II vs III/IV; III/IV vs V/VI) reproduced the canonical early medial-temporal tau focus, with progressive spread to posterior cingulate and retrosplenial cortex, lateral temporal and parietal association areas, and ultimately broader neocortex[40,47]. This finding was robust to the covariate feature set available (i.e. with or without plasma ptau-217) and the standardization of MRI imaging. Namely, in the out-of-sample NACC cohort consisting of over 475 scanning sessions with non-SCAN compliant MRIs we accurately recovered the underlying Braak staging assessed at autopsy. This provides strong evidence that the synthetic tau PET generated using the CoMA-UNet aligns with unobserved ground truth tau PET that shows antemortem-postmortem concordance and Braak stage-like regional progression[3,38,48,49]. Taken together, these results indicate that even under heterogeneous, lower-quality MRI and covariate inputs, CoMA-UNet captures anatomically coherent Braak-ordered tau distributions, supporting its use in large retrospective MRI repositories lacking tau PET or complete covariates.

The synthetic tau PET generated through CoMA-UNet has several limitations. First, we did not have in our training sample atypical AD patients, as such characteristic atypical patterns of tau deposition exhibited in non-amnestic variants, such as Posterior cortical atrophy and Primary Progressive Aphasia dementias[50], are unlikely to be well represented. However, we were able to capture the full AD continuum based on Braak staging training on a low proportion of clinically impaired participants (n=228; 25%), this suggests that fine-tuning the model on even small datasets that encompass atypical pathologies may extend the fidelity of the synthetic tau PET

images to rarer, atypical topographies. Second, relative to modern vision-based deep learning models such as Stable Diffusion[33], CoMA-UNet contains significantly fewer trainable parameters. This lightweight base configuration of CoMA-UNet presented here makes it widely accessible and deployable, although limiting its maximal performance[51–53]. The framework, however, naturally supports parameter scaling, as additional covariate-parameterized prompts can be introduced to enable the learning of a broader set of covariate-specific representations[54,55]. Such extensions are most appropriate for covariates that define subject strata whose expected ground-truth distributions differ fundamentally from those of other groups[50,56–58]. Thus, while the present implementation is constrained by parameter count, the CoMA-UNet architecture provides a conceptually simple and extensible pathway for scaling. Finally, our current implementation of CoMA-UNet was trained exclusively to synthesize FTP tau PET from MRI. Given that tau PET tracers exhibit distinct profiles of off-target binding and region–specific sensitivity to tau pathology[59–61], extending the model to jointly learn across multiple tracers represents an important direction for future work.

We present and validate the CoMA-UNet pipeline for high fidelity MRI to tau-PET synthesis with varied non-PET covariate data available. This novel approach out-performs other deep learning approaches, producing biologically valid, and clinically useful synthetic tau PETs. Through the application of this pipeline, retrospective analyses to investigate regional tau PET pathology in MRI-only or partially missing datasets is possible. Through these synthetic tau PET images, new avenues for predictive modelling are available allowing the aggregation of legacy data resources with publicly available tau PET samples or the precision re-stratification of completed clinical trials.

# Methods

## Participants and Data

### Datasets

Cross-sectional and cross-validation cohorts were drawn from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the A4 study, comprising ADNI participants across the cognitive spectrum (cognitively normal, mild cognitive impairment, and Alzheimer's dementia) and cognitively normal A4 participants, each with available T1-weighted MRI and [$^{18}$F]flortaucipir (FTP) tau PET imaging. The ADNI-A4 training dataset (n = 1250 image pairs) was randomly partitioned into five folds, balanced for meta-temporal tau burden and amyloid-β (Aβ) positivity. All analyses performed within this dataset used a standard five-fold procedure: for each fold, CoMA-UNet was trained on four folds and used to synthesize tau PET only for that fold's held-out subjects. For external dataset analyses, CoMA-UNet was trained once on the entirety of the ADNI-A4 training set. The cross-sectional validation dataset (n = 200 image pairs) comprised participants from ADNI (n = 100) and A4 (n = 100). For longitudinal evaluation, we used Longitudinal validation dataset which comprised held out 298 serial observations from 126 ADNI participants with multiple MRI-PET pairs to assess within-subject change. Separately, for post-mortem validation, we held out 99 distinct ADNI participants with autopsy-confirmed Braak stage data. An additional external A4 validation cohort included unseen A4 participants with available T1 MRI but no corresponding tau PET, used to test generalization to missing-modality settings. Plasma biomarkers were quantified using Fujirebio Lumipulse assays (Aβ42 /Aβ40 and p-tau217 in participants from the ADNI and A4 cohorts) and Quanterix Simoa assays (Aβ42 /Aβ40, p-tau217, neurofilament light [NfL], and glial fibrillary acidic protein [GFAP] in participants from the ADNI cohort). For further external validation, we included participants from the national

Alzheimer's Coordinated Center (NACC), encompassing both SCAN-compliant and non-SCAN-compliant MRI acquisitions. The NACC SCAN (n = 1409) cohort had paired tau PET using either [$^{18}$F]flortaucipir (n = 958) or [$^{18}$F]MK-6240 (n = 451), providing heterogeneity in tracer kinetics and quantification scales. The NACC non-SCAN cohort (n = 475) consisted of participants with autopsy-confirmed Braak staging but no PET imaging, enabling neuropathological validation. Plasma biomarkers were unavailable for NACC participants.

## Covariates & Feature Sets

The synthesis of the tau PET images from MRI proceeded in three stages, each using a distinct set of covariates. In the first stage, CatBoost models were used to estimate regional tau burden, MetaTempTau, Aβ status, and MMSE. These predictors use demographic (age, sex, education, marital status, race, height, weight), vital sign (systolic and diastolic blood pressure, pulse, respiratory rate, temperature), genetic (APOE4 allele count), and plasma biomarkers (p-tau217, GFAP, NfL, Aβ40, and Aβ42), when available. In the second stage, the deep learning model integrated these MetaTempTau, Aβ status, and MMSE estimates through a conditional decoder. In the third stage, the regional tau estimates are used in a final modulation step. Note that plasma biomarkers were not provided directly to CoMA-UNet. For model performance analyses and all downstream evaluation tasks, MetaTempTau and regional tau values were derived exclusively from synthetic tau PET images.

## Regions of Interest

In our analyses we consider $R = 32$ regions of interest (ROI) taken from the Desikan-Killiany (DK) Atlas (Supp Table 1). In particular, these ROIs span the parietal, temporal, and include the Hippocampus and Amygdala. For each of the ROIs, we consider the left and right hemispheres separately. In addition to these regions, we also utilize a Meta ROI, denoted as $MetaTempTau$,

consisting of the Amygdala, Inferior and Middle Temporal, Entorhinal, Parahippocampal, and Fusiform regions. The SUVR in the $MetaTempTau$ is calculated as the weighted average of SUVR in each of its constituent regions.

# Models

## Deep learning model for synthetic tau PET generation

A novel deep learning model, Covariate-modulated Attention UNet, was used to synthesize tau PET scans from MRIs. The architecture of the model consists of an Attention UNet with conditional up-sampling steps integrating subject-specific covariate data and estimated pathological markers. In particular, we replace all up-sampling Conv(·) operations in the Decoder with Conditional Convolution operators[62], integrating age, sex, education, as well as estimated MMSE and $MetaTempTau$.

To further increase the inductive bias and enhance the predictive capability of CoMA-UNet, estimated regional average tau values obtained from CatBoost[63] models are incorporated. The core idea is to modulate the initial predictions using a dynamic prompt that is both parameterized and tailored based on individual covariates and estimated regional tau load. Through the integration of these individual average tau values directly into the prediction process, the model becomes more sensitive to variations in disease pathology among different subjects. This modulation framework thus enables the model to capture subtle variations in the data associated with covariates and pathological states, leading to more accurate and individualized assessments and enhancing its predictive performance.

Formally, given an input MRI $M_i$, the CoMA-UNET produces an initial prediction $Z_i \in \mathbb{R}^{p_1 \times p_2 \times p_3}$. This output is refined by computing the final prediction $\hat{P}_i$ as $\hat{P}_i$ =ReLU(Conv($Z_i \oplus$ConvBlocks($\tilde{F}^{c_i} \oplus Z_i$)), where ReLU(·) is the Rectified Linear Unit activation function, and $\oplus$ represents concatenation along the channel dimension. The function ConvBlocks(·) consists of a sequence of three convolutional operations, each followed by PReLU activation, dropout, and instance normalization, processing the concatenated features to extract higher-level representations. The modulated dynamic prompt is computed as $\tilde{F}^{c_i} = F +$ ConvBlocks$\left(F^{c_i} \oplus \hat{S}_i \oplus \hat{I}_i\right)$, which adjusts a base dynamic prompt $F$ using three subject-specific components. The tensor $\hat{I}_i \in \mathbb{R}^{p_1 \times p_2 \times p_3}$ contains CatBoost-estimated regional tau SUVR values $\widehat{tau}_{r,i}$ placed in voxels corresponding to brain region $r \in R$, and zeros elsewhere, thereby masking out non-informative background. The tensor $\hat{S}_i$ contains corresponding regional uncertainty estimates $\hat{\sigma}_{r,i}$ for each $\widehat{tau}_{r,i}$, assigned to the same voxels. The parameterized dynamic prompt $F^{c_i}$, selected based on the subject's covariate $c_i$. In this work, we set $c_i$ to be the Aβ positivity status estimated by a CatBoost model. The convolution Conv(·) applies further processing to the combined features of $\hat{I}_i$ and $\tilde{F}^{c_i}$, through a convolutional layer followed by PReLU activation, dropout, and instance normalization. This formulation allows the model to incorporate individual pathological markers and covariate information directly into the feature space, enabling it to adjust its predictions based on personalized disease profiles and improving its ability to capture individual differences for enhanced prediction accuracy.

The CoMA-UNet was trained by minimizing a ROI-weighted MSE, $L_{Gen}$, in which the squared residuals in voxels corresponding to ROI $r$ are scaled by pre-specified non-zero weights, $w_r$. This loss function measures the difference between the ground-truth tau PET and the generated synthetic tau PET; when $L_{Gen} = 0$, the synthetic tau PET and the actual tau PET are the same. Concretely, over a set of $n$ MRI and actual tau PET pairs, $\{(M_i, P_i)\}_{i=1}^{n}$ , we define

$L_{Gen} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{|V|}\sum_{v\in V} w(v)\left(P_{i,v} - \hat{P}_{i,v}\right)^2$, where $\hat{P}_{i,v}$ is the value of voxel $v$ in the synthetic tau PET generated from $M_i$, $V$ is the set of all voxel coordinates, and $w(v)$ is a function which either assigns a user-specified value of $w_r \in \mathbb{R}$ if the voxel $v$ is in ROI $r$, or a value of 1 otherwise.

## Prediction of covariates and tau values

Independent CatBoost predictors were trained using plasma biomarker, family history, demographic, vital, and MRI-derived regional summary features to estimate temporal meta-ROI SUVR ($MetaTempTau$), as well as regional tau SUVR estimates ($\widehat{tau}_{r,i}$) and their associated standard deviation ($\hat{\sigma}_{r,i}$), assuming normally distributed errors. Using the same feature set, additional CatBoost models were separately trained to predict Aβ status, and a KNN model was used to predict MMSE scores. Missing covariate data were imputed using a KNN algorithm, with the number of neighbors equal to the square root of the total number of features, and each neighbor weighted uniformly. CatBoost and KNN model training and evaluation employed the same training and testing partitions as the deep learning model.

## Image Processing

All T1-weighted MRI volumes were preprocessed using FreeSurfer v7.1[64], which performed skull stripping, intensity normalization, cortical surface reconstruction, and subcortical segmentation of the native-space anatomical images. This process generated subject-specific anatomical volumes and corresponding Desikan-Killiany parcellations, providing voxel-wise anatomical label maps for subsequent region-based sampling. Each intensity-normalized MRI was then resampled to the model's fixed grid of 128x128x128 voxels ($p_1 = p_2 = p_3 = 128$) using nearest-neighbor interpolation, symmetrically padded along the axial axis, and centrally cropped before a single channel dimension was added to match the network's expected input. Tau PET images using either [$^{18}$F]Flortaucipir (FTP) or [$^{18}$F]MK-6240 (MK) radiotracers corresponding to the T1 MRIs

were processed following the Berkeley PET Imaging Pipeline (B-PIP)[65], which includes framewise motion correction, frame averaging, co-registration to the subject's MRI, and normalization to the inferior cerebellar gray reference region for standardized uptake value ratio (SUVR) computation. For select downstream evaluations, specified in the Evaluation section, the model's output volumes were spatially normalized to MNI space (157x189x156 voxels) using nonlinear registration in SPM12 and resampled with third-order B-spline interpolation.

## Metrics

Similarity between ground-truth tau PET and synthetic tau PET was assessed across cohorts using regional SUVR calculated from $R$ ROIs in addition to full-image wide metrics. Pearson's correlation, Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE) were calculated in each ROI, with MAPE and MAE also being calculated across the entire image. To further quantify the quality of the synthetic images we calculate the average of the regional Pearson correlations ($Corr_{AVG}$) and the structural similarity index (SSIM). SSIM is a metric used to measure the similarity between two images. SSIM is designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have been shown to be inconsistent with human visual perception. Instead, SSIM considers changes in structural information, luminance, and contrast. The SSIM index is a decimal value between -1 and 1, where 1 indicates perfect similarity. It is calculated based on the comparison of three main components of the images: luminance (l), contrast (c), and structure (s). We give the precise mathematical definitions of each of the described performance metrics in the Supplementary.

# Training & Evaluation

## Training

The model was trained on an Nvidia Tesla V100 (16G VRAM) GPU using PyTorch[66]. Optimization of the generative loss function $L_{Gen}$ was performed using AdamW with an initial learning rate of 0.001, weight decay of 0.001, and a reduce-on-plateau learning rate scheduler, with batch sizes of two. Early stopping was performed after the model didn't improve on a held-out validation subset for five consecutive epochs. Model parameters were first tuned using a five-fold cross-validation framework, in which each fold was trained exclusively on its training partition and evaluated on the held-out partition via strict out-of-fold predictions. Following hyperparameter selection, a final model was retrained on the full cross-validation dataset (n = 1250) using the same optimization procedure and was subsequently used for all downstream validation analyses.

All auxiliary predictors for ROI-level tau, Aβ status, and cognitive score were exclusively fit on the same training split as the deep learning model. That is, in cross-validation, each auxiliary model was re-fit within every outer fold in lockstep with the DL model; for cross-sectional analyses, models were fit on the full cross-sectional training set. Thus, all evaluations relied strictly on out-of-sample predictions from every model component.

## Evaluation

During five-fold cross-validation over the 1,250 samples in the ADNI-A4 training dataset, the model was trained on four folds (80%) and tested exclusively on the held-out fold (20%). Model performance for each fold was then quantified using the similarity metrics described in the Supplementary Metrics Section, including SSIM, MAPE, MAE, ROI-focused variants of MAPE

and MAE, as well as regional correlation and $Corr_{AVG}$. We present the cross-validation performance of the CoMA-UNet as benchmarked on the ADNI-A4 training dataset against several other model with alternative architectures. The CoMA-UNet model was then trained on the entirety of ADNI-A4 training dataset and was retained for evaluation on the downstream tasks to assess the clinical utility of synthetic tau PET scans. We elaborate each downstream task in the subsequent sections.

### Longitudinal Predictions

We evaluated how well synthetic tau PET-derived $\widehat{MetaTempTau}$ tracked longitudinal changes in the corresponding actual PET-derived $MetaTempTau$ in held-out ADNI subjects with repeated visits (n = 126 subjects, with total 298 observations). We fit a linear mixed-effects model that relates the true metric to the synthetic one while allowing subject-specific deviations $MetaTempTau_{it} = \beta_0 + \beta_1 \widehat{MetaTempTau}_{it} + b_{0i} + b_{1i}\widehat{MetaTempTau}_{it} + \epsilon_{it}$, with random intercepts and random slopes $(b_{0i}, b_{1i})$ for subject $i$. Here, $\beta_0$ and $\beta_1$ represent the population-level fixed intercept and slope relating synthetic MetaTempTau to MetaTempTau, and $\epsilon_{it}$ denotes the within-subject residual error. Models were estimated by restricted maximum likelihood for coefficient reporting. We report fixed-effect estimates (slope/intercept) with standard errors and two-sided Wald $p$-values, variance components for random effects, and the correlation between random intercepts and slopes.

### AD Diagnosis Classification

We conducted pairwise classifications across cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease Dementia (Dem) groups in the ADNI samples within the ADNI-A4 training and cross-sectional validation datasets. Samples from A4 were excluded in this analysis to prevent the model from being biased against population differences. For these pairwise classification tasks, we trained logistic regression models using only $\widehat{MetaTempTau}$ and

$MetaTempTau$. In the ADNI-A4 training dataset, we used a five-fold cross-validation scheme, computing $\widehat{MetaTempTau}$ from the synthetic tau PET images generated from the MRIs in each held-out fold. Logistic-regression classifiers were trained on the remaining four folds and evaluated on the held-out fold. In the external cross-sectional validation dataset, $\widehat{MetaTempTau}$ was classified using logistic-regression classifiers trained on the $\widehat{MetaTempTau}$ derived from synthetic tau PET images produced for held-out samples in the cross-validation. Model performance was assessed using receiver operating characteristic area under the curve (ROC-AUC) scores. To quantify voxel-level discriminative effect size across the brain, we computed voxel-wise Cohen's d values in spatially normalized synthetic tau PET for each diagnostic cohort comparison.

### Cognitive Baseline Prediction

To assess how well synthetic tau PETs relate to baseline cognition in comparison to actual tau PETs, we compared correlations of $\widehat{MetaTempTau}$-MMSE and $MetaTempTau$-MMSE with the cross-sectional cohort, quantifying this similarity using Steiger's Z. We additionally evaluated the voxel-wise correlation between spatially normalized synthetic tau PET scans and MMSE. Finally, to assess incremental value over structural MRI, we extracted ROI-level SUVR and corresponding MRI regional volumes from the Desikan-Killiany atlas. For each ROI and Aβ stratum, we compared SUVR-MMSE versus MRI-MMSE correlations using Steiger's Z for dependent correlations. Multiple comparisons across ROIs were controlled with the Benjamini-Hochberg False Discovery Rate procedure.

### Braak Stage Contrasts

To test whether synthetic tau PET recapitulates the canonical neuropathological progression of tau, we performed adjacent-stage voxel-wise contrasts in the ADNI-NP cohort (n=99), and separately, in the NACC non-SCAN cohort (n = 475). Synthetic PET volumes were spatially

normalized to MNI space and grouped according to Braak stages I/II, III/IV, and V/VI. We conducted two independent contrasts (I/II vs. III/IV; III/IV vs. V/VI) using two-sample independent t-tests at each voxel, and additionally computed Cohen's d effect-size maps. As a complementary ROI analysis, we summarized medial-temporal (MTL) SUVR from each synthetic volume and fit an ordinal (proportional-odds) regression of MTL SUVR on ordered Braak group.

### A4 population restratification

To examine whether model-derived measures could refine treatment group stratification, we first modeled PACC trajectories as a function of treatment assignment (placebo versus solanezumab) in the A4 validation cohort. We then re-stratified participants within each treatment arm according to $\widehat{MetaTempTau}$ tertiles derived from synthetic tau PET scans, generated both with and without incorporation of plasma biomarker information to explore the impact of plasma-informed versus MRI-only synthetic PET stratifications. This approach enabled assessment of cognitive trajectories within more biologically informed subgroups, providing a framework to evaluate whether synthetic imaging markers could enhance sensitivity to treatment effects beyond conventional randomization strata. Trajectory estimation followed the same statistical framework used in the original A4 study. Specifically, a linear mixed-effects model with natural cubic splines was employed to capture nonlinear longitudinal change, estimated by restricted maximum likelihood[45].

### Evaluation under unseen radiotracer in held-out subject population

We performed a rigorous evaluation of the proposed model's generalization by assessing performance in the NACC-SCAN cohort (n = 1409), which contained both tau PET acquired with FTP tracer (n = 958), and an unseen tracer, MK-6240 (n = 451). We did not provide the model with plasma biomarkers in this experiment. For each tracer and ROI, we computed Pearson's r between predicted SUVR from synthetic PET and observed SUVR from the corresponding

ground-truth PET, along with $R^2$. Since FTP and MK differ in binding properties and quantitative scales, we restricted evaluation metrics to these scale-invariant, correlation-based measures.

## Data and Code Availability

Python scripts, model checkpoints, and help files are available on GitHub. NACC data can be requested and downloaded at https://naccdata.org. Data from A4 are available to download on the LONI website at https://ida.loni.usc.edu. Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (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 ADNI has been to test whether serial magnetic resonance imaging MRI), positron emission tomography (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 (AD). For up-to-date information, see www.adni-info.org.

# Acknowledgements

J.G. is supported by the Alzheimer's Association (23AARF-1026883). 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 the following: 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 its affiliated company 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. The A4 Study was a secondary prevention trial in preclinical Alzheimer's disease, aiming to slow cognitive decline associated with brain amyloid accumulation in clinically normal older individuals. The A4 Study was funded by a public-private-philanthropic partnership, including funding from the National Institutes of Health-National Institute on Aging, Eli Lilly and Company, Alzheimer's Association, Accelerating Medicines Partnership, GHR Foundation, an anonymous foundation, and additional private

donors, with in-kind support from Avid Radiopharmaceuticals, Cogstate, Albert Einstein College of Medicine and the Foundation for Neurologic Diseases. The companion observational Longitudinal Evaluation of Amyloid Risk and Neurodegeneration (LEARN) Study was funded by the Alzheimer's Association and GHR Foundation. The A4 and LEARN Studies were led by Dr. Reisa Sperling at Brigham and Women's Hospital, Harvard Medical School, and Dr. Paul Aisen at the Alzheimer's Therapeutic Research Institute (ATRI) at the University of Southern California. The A4 and LEARN Studies were coordinated by ATRI at the University of Southern California, and the data are made available under the auspices of Alzheimer's Clinical Trial Consortium through the Global Research & Imaging Platform (GRIP). The complete A4 Study Team list is available on: https://www.actcinfo.org/a4-study-team-lists/. We would like to acknowledge the dedication of the study participants and their study partners who made the A4 and LEARN Studies possible. The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADRCs: P30 AG062429 (PI James Brewer, MD, PhD), P30 AG066468 (PI Oscar Lopez, MD), P30 AG062421 (PI Bradley Hyman, MD, PhD), P30 AG066509 (PI Thomas Grabowski, MD), P30 AG066514 (PI Mary Sano, PhD), P30 AG066530 (PI Helena Chui, MD), P30 AG066507 (PI Marilyn Albert, PhD), P30 AG066444 (PI John Morris, MD), P30 AG066518 (PI Jeffrey Kaye, MD), P30 AG066512 (PI Thomas Wisniewski, MD), P30 AG066462 (PI Scott Small, MD), P30 AG072979 (PI David Wolk, MD), P30 AG072972 (PI Charles DeCarli, MD), P30 AG072976 (PI Andrew Saykin, PsyD), P30 AG072975 (PI David Bennett, MD), P30 AG072978 (PI Neil Kowall, MD), P30 AG072977 (PI Robert Vassar, PhD), P30 AG066519 (PI Frank LaFerla, PhD), P30 AG062677 (PI Ronald Petersen, MD, PhD), P30 AG079280 (PI Eric Reiman, MD), P30 AG062422 (PI Gil Rabinovici, MD), P30 AG066511 (PI Allan Levey, MD, PhD), P30 AG072946 (PI Linda Van Eldik, PhD), P30 AG062715 (PI Sanjay Asthana, MD, FRCP), P30 AG072973 (PI Russell Swerdlow, MD), P30 AG066506 (PI Todd Golde, MD, PhD), P30 AG066508 (PI Stephen Strittmatter, MD, PhD), P30 AG066515 (PI Victor Henderson, MD, MS), P30 AG072947 (PI Suzanne Craft, PhD), P30 AG072931 (PI Henry Paulson, MD, PhD), P30

AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Justin Miller, PhD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD). The NACC database is funded by NIA/NIH Grant U24 AG072122. SCAN is a multi-institutional project that was funded as a U24 grant (AG067418) by the National Institute on Aging in May 2020. Data collected by SCAN and shared by NACC are contributed by the NIA-funded ADRCs as follows:

Arizona Alzheimer's Center - P30 AG072980 (PI: Eric Reiman, MD); R01 AG069453 (PI: Eric Reiman (contact), MD); P30 AG019610 (PI: Eric Reiman, MD); and the State of Arizona which provided additional funding supporting our center; Boston University - P30 AG013846 (PI Neil Kowall MD); Cleveland ADRC - P30 AG062428 (James Leverenz, MD); Cleveland Clinic, Las Vegas – P20AG068053; Columbia - P50 AG008702 (PI Scott Small MD); Duke/UNC ADRC – P30 AG072958; Emory University - P30AG066511 (PI Levey Allan, MD, PhD); Indiana University - R01 AG19771 (PI Andrew Saykin, PsyD); P30 AG10133 (PI Andrew Saykin, PsyD); P30 AG072976 (PI Andrew Saykin, PsyD); R01 AG061788 (PI Shannon Risacher, PhD); R01 AG053993 (PI Yu-Chien Wu, MD, PhD); U01 AG057195 (PI Liana Apostolova, MD); U19 AG063911 (PI Bradley Boeve, MD); and the Indiana University Department of Radiology and Imaging Sciences; Johns Hopkins - P30 AG066507 (PI Marilyn Albert, Phd.); Mayo Clinic - P50 AG016574 (PI Ronald Petersen MD PhD); Mount Sinai - P30 AG066514 (PI Mary Sano, PhD); R01 AG054110 (PI Trey Hedden, PhD); R01 AG053509 (PI Trey Hedden, PhD); New York University - P30AG066512-01S2 (PI Thomas Wisniewski, MD); R01AG056031 (PI Ricardo Osorio, MD); R01AG056531 (PIs Ricardo Osorio, MD; Girardin Jean-Louis, PhD); Northwestern University - P30 AG013854 (PI Robert Vassar PhD); R01 AG045571 (PI Emily Rogalski, PhD); R56 AG045571, (PI Emily Rogalski, PhD); R01 AG067781, (PI Emily Rogalski, PhD); U19 AG073153, (PI Emily Rogalski, PhD); R01 DC008552, (M.-Marsel Mesulam, MD); R01

AG077444, (PIs M.-Marsel Mesulam, MD, Emily Rogalski, PhD); R01 NS075075 (PI Emily Rogalski, PhD); R01 AG056258 (PI Emily Rogalski, PhD); Oregon Health & Science University - P30 AG066518 (PI Lisa Silbert, MD, MCR); Rush University - P30 AG010161 (PI David Bennett MD); Stanford – P30AG066515; P50 AG047366 (PI Victor Henderson MD MS); University of Alabama, Birmingham – P20; University of California, Davis - P30 AG10129 (PI Charles DeCarli, MD); P30 AG072972 (PI Charles DeCarli, MD); University of California, Irvine - P50 AG016573 (PI Frank LaFerla PhD); University of California, San Diego - P30AG062429 (PI James Brewer, MD, PhD); University of California, San Francisco - P30 AG062422 (Rabinovici, Gil D., MD); University of Kansas - P30 AG035982 (Russell Swerdlow, MD); University of Kentucky - P30 AG028283-15S1 (PIs Linda Van Eldik, PhD and Brian Gold, PhD); University of Michigan ADRC - P30AG053760 (PI Henry Paulson, MD, PhD) P30AG072931 (PI Henry Paulson, MD, PhD) Cure Alzheimer's Fund 200775 - (PI Henry Paulson, MD, PhD) U19 NS120384 (PI Charles DeCarli, MD, University of Michigan Site PI Henry Paulson, MD, PhD) R01 AG068338 (MPI Bruno Giordani, PhD, Carol Persad, PhD, Yi Murphey, PhD) S10OD026738-01 (PI Douglas Noll, PhD) R01 AG058724 (PI Benjamin Hampstead, PhD) R35 AG072262 (PI Benjamin Hampstead, PhD) W81XWH2110743 (PI Benjamin Hampstead, PhD) R01 AG073235 (PI Nancy Chiaravalloti, University of Michigan Site PI Benjamin Hampstead, PhD) 1I01RX001534 (PI Benjamin Hampstead, PhD) IRX001381 (PI Benjamin Hampstead, PhD); University of New Mexico - P20 AG068077 (Gary Rosenberg, MD); University of Pennsylvania - State of PA project 2019NF4100087335 (PI David Wolk, MD); Rooney Family Research Fund (PI David Wolk, MD); R01 AG055005 (PI David Wolk, MD); University of Pittsburgh - P50 AG005133 (PI Oscar Lopez MD); University of Southern California - P50 AG005142 (PI Helena Chui MD); University of Washington - P50 AG005136 (PI Thomas Grabowski MD); University of Wisconsin - P50 AG033514 (PI Sanjay Asthana MD FRCP); Vanderbilt University – P20 AG068082; Wake Forest - P30AG072947 (PI Suzanne Craft, PhD); Washington University, St. Louis - P01 AG03991 (PI John Morris MD); P01 AG026276 (PI John Morris MD); P20 MH071616 (PI Dan Marcus); P30

AG066444 (PI John Morris MD); P30 NS098577 (PI Dan Marcus); R01 AG021910 (PI Randy Buckner); R01 AG043434 (PI Catherine Roe); R01 EB009352 (PI Dan Marcus); UL1 TR000448 (PI Brad Evanoff); U24 RR021382 (PI Bruce Rosen); Avid Radiopharmaceuticals / Eli Lilly; Yale - P50 AG047270 (PI Stephen Strittmatter MD PhD); R01AG052560 (MPI: Christopher van Dyck, MD; Richard Carson, PhD); R01AG062276 (PI: Christopher van Dyck, MD); 1Florida - P30AG066506-03 (PI Glenn Smith, PhD); P50 AG047266 (PI Todd Golde MD PhD)

# References


1. 2025 Alzheimer's disease facts and figures. *Alzheimer's & Dementia* **21**, (2025).
2. Safiri, S. *et al.* Alzheimer's disease: a comprehensive review of epidemiology, risk factors, symptoms diagnosis, management, caregiving, advanced treatments and associated challenges. *Front Med (Lausanne)* **11**, (2024).
3. Jack, C. R. *et al.* Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's & Dementia* **20**, 5143–5169 (2024).
4. Thal, D. R. *et al.* Estimation of amyloid distribution by [18F]flutemetamol PET predicts the neuropathological phase of amyloid β-protein deposition. *Acta Neuropathol* **136**, 557–567 (2018).
5. Macedo, A. C. *et al.* The Use of Tau PET to Stage Alzheimer Disease According to the Braak Staging Framework. *Journal of Nuclear Medicine* **64**, 1171–1178 (2023).
6. Therriault, J. *et al.* Biomarker-based staging of Alzheimer disease: rationale and clinical applications. *Nat Rev Neurol* **20**, 232–244 (2024).
7. Jagust, W. Imaging the evolution and pathophysiology of Alzheimer disease. *Nature Reviews Neuroscience* vol. 19 687–700 Preprint at https://doi.org/10.1038/s41583-018-0067-3 (2018).
8. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol* **82**, 239–259 (1991).
9. Therriault, J. *et al.* Biomarker modeling of Alzheimer's disease using PET-based Braak staging. *Nat Aging* **2**, 526–535 (2022).
10. Schöll, M. *et al.* PET Imaging of Tau Deposition in the Aging Human Brain. *Neuron* **89**, 971–982 (2016).
11. Maass, A. *et al.* Comparison of multiple tau-PET measures as biomarkers in aging and Alzheimer's disease. *Neuroimage* **157**, 448–463 (2017).
12. Betthauser, T. J. *et al.* In vivo characterization and quantification of neurofibrillary tau PET radioligand 18 F-MK-6240 in humans from Alzheimer disease dementia to young controls. *Journal of Nuclear Medicine* **60**, 93–99 (2019).
13. Ashton, N. J. *et al.* Diagnostic Accuracy of a Plasma Phosphorylated Tau 217 Immunoassay for Alzheimer Disease Pathology. *JAMA Neurol* **81**, 255–263 (2024).
14. Bucci, M., Chiotis, K. & Nordberg, A. Alzheimer's disease profiled by fluid and imaging markers: tau PET best predicts cognitive decline. *Mol Psychiatry* **26**, 5888–5898 (2021).
15. Karlsson, L. *et al.* Machine learning prediction of tau-PET in Alzheimer's disease using plasma, MRI, and clinical data. *Alzheimer's and Dementia* **21**, (2025).
16. Lew, C. O., Zhou, L., Mazurowski, M. A., Murali Doraiswamy, P. & Petrella, J. R. MRI-based Deep Learning Assessment of Amyloid, Tau, and Neurodegeneration Biomarker Status across the Alzheimer Disease Spectrum. *Radiology* **309**, (2023).
17. Giorgio, J. *et al.* A robust and interpretable machine learning approach using multimodal biological data to predict future pathological tau accumulation. *Nat Commun* **13**, (2022).
18. Lee, J. *et al.* Synthesizing images of tau pathology from cross-modal neuroimaging using deep learning. *Brain* **147**, 980–995 (2024).

19. Sperling, R. A. *et al.* Trial of Solanezumab in Preclinical Alzheimer's Disease. *New England Journal of Medicine* **389**, 1096–1107 (2023).
20. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) 234–241 (Springer, Cham, 2015). doi:10.1007/978-3-319-24574-4_28.
21. Williams, C. *et al.* A Unified Framework for U-Net Design and Analysis. http://arxiv.org/abs/2305.19638 (2023).
22. Oktay, O. *et al.* Attention U-Net: Learning Where to Look for the Pancreas. http://arxiv.org/abs/1804.03999 (2018).
23. Hatamizadeh, A. *et al.* UNETR: Transformers for 3D Medical Image Segmentation. http://arxiv.org/abs/2103.10504 (2021).
24. Cai, Y. *et al.* Swin Unet3D: a three-dimensional medical image segmentation network combining vision transformer and convolution. *BMC Med Inform Decis Mak* **23**, (2023).
25. Yang, Y.-Q. *et al.* Swin3D: A Pretrained Transformer Backbone for 3D Indoor Scene Understanding. http://arxiv.org/abs/2304.06906 (2023).
26. Lai, X. *et al. Stratified Transformer for 3D Point Cloud Segmentation*. *arXiv preprint arXiv:2203.14508* https://arxiv.org/abs/2203.14508 (2022).
27. Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 5967–5976 (2017). doi:10.1109/CVPR.2017.632.
28. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* 2223–2232 (2017). doi:10.1109/ICCV.2017.244.
29. Zhang, L., Chen, X., Zhang, J., Dong, R. & Ma, K. Contrastive Deep Supervision. http://arxiv.org/abs/2207.05306 (2022).
30. Theodorou, B. *et al.* MRI2PET: Realistic PET Image Synthesis from MRI for Automated Inference of Brain Atrophy and Alzheimer's. *medRxiv* https://doi.org/10.1101/2025.04.23.25326302 (2025) doi:10.1101/2025.04.23.25326302.
31. Zheng, X. *et al.* Generating synthetic brain PET images of synaptic density based on MR T1 images using deep learning. *EJNMMI Phys* **12**, 30 (2025).
32. Dosovitskiy, A. *et al.* An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. http://arxiv.org/abs/2010.11929 (2020).
33. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 10684–10695 (2022). doi:10.1109/CVPR52688.2022.01074.
34. Ho, J., Jain, A. & Abbeel, P. *Denoising Diffusion Probabilistic Models*. *arXiv preprint arXiv:2006.11239* https://arxiv.org/abs/2006.11239 (2020) doi:10.48550/arXiv.2006.11239.
35. Saharia, C. *et al. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding*. *arXiv preprint arXiv:2205.11487* https://arxiv.org/abs/2205.11487 (2022).

36. Bollack, A. *et al.* Longitudinal amyloid and tau PET imaging in Alzheimer's disease: A systematic review of methodologies and factors affecting quantification. *Alzheimer's & Dementia* **19**, 5232–5252 (2023).
37. Visser, D. *et al.* Longitudinal Tau PET Using $^{18}$ F-Flortaucipir: The Effect of Relative Cerebral Blood Flow on Quantitative and Semiquantitative Parameters. *Journal of Nuclear Medicine* **64**, 281–286 (2023).
38. Soleimani-Meigooni, D. N. *et al.* 18F-flortaucipir PET to autopsy comparisons in Alzheimer's disease and other neurodegenerative diseases. *Brain* **143**, 3477–3494 (2020).
39. Leuzy, A. *et al.* Comparison of Group-Level and Individualized Brain Regions for Measuring Change in Longitudinal Tau Positron Emission Tomography in Alzheimer Disease. *JAMA Neurol* **80**, 614 (2023).
40. Ioannou, K. *et al.* Tau PET positivity predicts clinically relevant cognitive decline driven by Alzheimer's disease compared to comorbid cases; proof of concept in the ADNI study. *Mol Psychiatry* **30**, 587–599 (2025).
41. Ossenkoppele, R. *et al.* Associations between tau, Aβ, and cortical thickness with cognition in Alzheimer disease. *Neurology* **92**, (2019).
42. Groot, C. *et al.* Tau Positron Emission Tomography for Predicting Dementia in Individuals With Mild Cognitive Impairment. *JAMA Neurol* **81**, 845 (2024).
43. Singleton, E. H. *et al.* Longitudinal tau aggregation, atrophy, and cognitive decline in Alzheimer's disease. *Alzheimers Dement* **21**, e70435 (2025).
44. St-Onge, F., Chapleau, M., Breitner, J. C. S., Villeneuve, S. & Pichet Binette, A. Tau accumulation and its spatial progression across the Alzheimer's disease spectrum. *Brain Commun* **6**, (2023).
45. Sperling, R. A. *et al.* Amyloid and Tau Prediction of Cognitive and Functional Decline in Unimpaired Older Individuals: Longitudinal Data from the A4 and LEARN Studies. *J Prev Alzheimers Dis* **11**, 802–813 (2024).
46. Sims, J. R. *et al.* Donanemab in Early Symptomatic Alzheimer Disease. *JAMA* **330**, 512 (2023).
47. Berron, D. *et al.* Early stages of tau pathology and its associations with functional connectivity, atrophy and memory. *Brain* **144**, 2771–2783 (2021).
48. DeTure, M. A. & Dickson, D. W. The neuropathological diagnosis of Alzheimer's disease. *Mol Neurodegener* **14**, 32 (2019).
49. Smith, R. *et al.* The accumulation rate of tau aggregates is higher in females and younger amyloid-positive subjects. *Brain* **143**, 3805–3815 (2020).
50. Graff-Radford, J. *et al.* New insights into atypical Alzheimer's disease in the era of biomarkers. *Lancet Neurol* **20**, 222–234 (2021).
51. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. (2015).
52. Tan, M. & Le, Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. (2020).
53. Dehghani, M. *et al.* Scaling Vision Transformers to 22 Billion Parameters. (2023).
54. Li, X. L. & Liang, P. Prefix-Tuning: Optimizing Continuous Prompts for Generation. (2021).

55. Jia, M. *et al.* Visual Prompt Tuning. (2022).
56. Hampel, H. *et al.* The Amyloid-β Pathway in Alzheimer's Disease. *Mol Psychiatry* **26**, 5481–5503 (2021).
57. Phillips, J. S. *et al.* Rates of longitudinal change in $^{18}$F-flortaucipir PET vary by brain region, cognitive impairment, and age in atypical Alzheimer's disease. *Alzheimer's & Dementia* **18**, 1235–1247 (2022).
58. Jack, C. R. *et al.* Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. *Alzheimer's & Dementia* **20**, 5143–5169 (2024).
59. Groot, C., Villeneuve, S., Smith, R., Hansson, O. & Ossenkoppele, R. Tau PET Imaging in Neurodegenerative Disorders. *J Nucl Med* **63**, 20S-26S (2022).
60. Gogola, A. *et al.* Direct Comparison of the Tau PET Tracers 18F-Flortaucipir and 18F-MK-6240 in Human Subjects. *J Nucl Med* **63**, 108–116 (2022).
61. Okamura, N. *et al.* The development and validation of tau PET tracers: current status and future directions. *Clin Transl Imaging* **6**, 305–316 (2018).
62. Yang, B., Bender, G., Le, Q. V. & Ngiam, J. CondConv: Conditionally Parameterized Convolutions for Efficient Inference. http://arxiv.org/abs/1904.04971 (2019).
63. Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. CatBoost: Unbiased boosting with categorical features. in *Advances in Neural Information Processing Systems (NeurIPS)* 6638–6648 (2018).
64. Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–81 (2012).
65. Landau, S. M. *et al.* Positron emission tomography harmonization in the Alzheimer's Disease Neuroimaging Initiative: A scalable and rigorous approach to multisite amyloid and tau quantification. *Alzheimers Dement* **21**, e14378 (2025).
66. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. (2019).