

# Advancing Tau PET Quantification in Alzheimer Disease with Machine Learning: Introducing THETA, a Novel Tau Summary Measure

Robel K. Gebre<sup>1</sup>, Alexis Moscoso Rial<sup>2,3</sup>, Sheelakumari Raghavan<sup>1</sup>, Heather J. Wiste<sup>4</sup>, Fiona Heeman<sup>2,3</sup>, Alejandro Costoya-Sánchez<sup>5-7</sup>, Christopher G. Schwarz<sup>1</sup>, Anthony J. Spychalla<sup>1</sup>, Val J. Lowe<sup>1</sup>, Jonathan Graff-Radford<sup>8</sup>, David S. Knopman<sup>8</sup>, Ronald C. Petersen<sup>4,8</sup>, Michael Schöll<sup>2,3,9</sup>, Melissa E. Murray<sup>10</sup>, Clifford R. Jack Jr.<sup>1</sup>, and Prashanthi Vemuri<sup>1</sup>, for the Alzheimer's Disease Neuroimaging Initiative

<sup>1</sup>Department of Radiology, Mayo Clinic, Rochester, Minnesota; <sup>2</sup>Department of Psychiatry and Neurochemistry, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; <sup>3</sup>Wallenberg Centre for Molecular and Translational Medicine, University of Gothenburg, Gothenburg, Sweden; <sup>4</sup>Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota; <sup>5</sup>Universidad de Santiago de Compostela, Santiago de Compostela, Spain; <sup>6</sup>Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas, Instituto de Salud Carlos III, Madrid, Spain; <sup>7</sup>Nuclear Medicine Department and Molecular Imaging Group, Instituto de Investigación Sanitaria de Compostela, Travesía da Choupana s/n, Santiago de Compostela, Spain; <sup>8</sup>Department of Neurology, Mayo Clinic, Rochester, Minnesota; <sup>9</sup>Dementia Research Centre, Institute of Neurology, University College London, London, United Kingdom; and <sup>10</sup>Department of Neuroscience, Mayo Clinic, Jacksonville, Florida

Alzheimer disease (AD) exhibits spatially heterogeneous 3- or 4-repeat tau deposition across participants. Our overall goal was to develop an automated method to quantify the heterogeneous burden of tau deposition into a single number that would be clinically useful. **Methods:** We used tau PET scans from 3 independent cohorts: the Mayo Clinic Study of Aging and Alzheimer's Disease Research Center (Mayo,  $n = 1,290$ ), the Alzheimer's Disease Neuroimaging Initiative (ADNI,  $n = 831$ ), and the Open Access Series of Imaging Studies (OASIS-3,  $n = 430$ ). A machine learning binary classification model was trained on Mayo data and validated on ADNI and OASIS-3 with the goal of predicting visual tau positivity (as determined by 3 raters following Food and Drug Administration criteria for <sup>18</sup>F-flortaucipir). The machine learning model used region-specific SUV ratios scaled to cerebellar crus uptake. We estimated feature contributions based on an artificial intelligence-explainable method (Shapley additive explanations) and formulated a global tau summary measure, Tau Heterogeneity Evaluation in Alzheimer's Disease (THETA) score, using SUV ratios and Shapley additive explanations for each participant. We compared the performance of THETA with that of commonly used meta-regions of interest (ROIs) using the Mini-Mental State Examination, the Clinical Dementia Rating-Sum of Boxes, clinical diagnosis, and histopathologic staging. **Results:** The model achieved a balanced accuracy of 95% on the Mayo test set and at least 87% on the validation sets. It classified tau-positive and -negative participants with an AUC of 1.00, 0.96, and 0.94 on the Mayo, ADNI, and OASIS-3 cohorts, respectively. Across all cohorts, THETA showed a better correlation with the Mini-Mental State Examination and the Clinical Dementia Rating-Sum of Boxes ( $\rho \geq 0.45$ ,  $P < 0.05$ ) than did meta-ROIs ( $\rho < 0.44$ ,  $P < 0.05$ ) and discriminated between participants who were cognitively unimpaired and those who had mild cognitive impairment with an effect size of 10.09, compared with an effect size of 3.08 for meta-ROIs. **Conclusion:** Our proposed approach identifies positive tau PET scans and provides a

quantitative summary measure, THETA, that effectively captures heterogeneous tau deposition observed in AD. The application of THETA for quantifying tau PET in AD exhibits great potential.

**Key Words:** neurology; Alzheimer disease; PET; machine learning; summary measure; tau

**J Nucl Med 2024; 00:1-8**

DOI: 10.2967/jnumed.123.267273

Alzheimer disease is characterized by the accumulation of  $\beta$ -amyloid plaques and neurofibrillary tangles in the brain. Neurofibrillary tangles are composed of hyperphosphorylated tau proteins that progress along predictable patterns, originating in the transentorhinal cortex and spreading to the limbic system and eventually to the neocortex. The spread of tau leads to cognitive impairment and dementia (1). The current understanding of Alzheimer disease (AD) pathophysiology and neurodegeneration suggests that neurofibrillary tangle accumulation correlates with disease progression and precedes clinical symptoms, making tau a promising biomarker for disease diagnosis and clinical trial designs (2). However, pathologic and imaging evidence reveals heterogeneous tau deposition patterns in the AD brain across individuals (3,4).

PET imaging is used to visualize and assess tau deposition in neurodegenerative disorders using radioligands that bind to paired helical filaments of neurofibrillary tangles to detect and track tau pathology in vivo (5). Preclinical AD PET studies indicate that tau spreads throughout cortical regions along multiple trajectories (6). Common tau PET quantification methods include meta-regions of interest (ROIs) in the medial temporal lobe (MTL), temporal lobe, and neocortex (7). However, these methods may overlook tau outside the meta-ROIs and underestimate focal tau depositions.

In this study, we followed visual rating criteria that relied on the density and distribution of tau identified using the Food and Drug Administration (FDA)-approved radiotracer <sup>18</sup>F-flortaucipir (Tauvid;

Received Dec. 25, 2023; revision accepted Jun. 25, 2024.

For correspondence or reprints, contact Prashanthi Vemuri (vemuri.prasanthi@mayo.edu).

Published online Jul. 25, 2024.

COPYRIGHT © 2024 by the Society of Nuclear Medicine and Molecular Imaging.

Lilly) for AD tau pathology at the B3 level (Braak stages V and VI) (8). These criteria consider focal tau deposition and could overcome the limitations of the meta-ROI-based methods. We hypothesized that a machine learning (ML) model can be developed to identify tau PET-positive scans by incorporating the visual ratings that account for the spatial heterogeneity of tau to achieve improved quantification. We aimed to, first, develop the model on a single-site dataset using regional SUV ratios (SUVRs) and visual assessments as targets and validate it on 2 independent datasets; second, compare this model with temporal, MTL, and neocortical meta-ROIs based on clinical scores and histopathologic staging; and third, develop a summary measure by leveraging spatial heterogeneity captured by the model.

## MATERIALS AND METHODS

### Study Participants

We included participants who had undergone  $^{18}\text{F}$ -flortaucipir tracer tau PET in the Mayo Clinic Study of Aging and Mayo Alzheimer's Disease Research Center dataset (referred to as Mayo;  $n = 1,290$ ) (9), the Alzheimer's Disease Neuroimaging Initiative phase 2 or 3 (ADNI;  $n = 831$ ), or the Open Access Series of Imaging Studies phase 3 (OASIS-3;  $n = 430$ ) (10). Individuals with frontotemporal dementia were excluded. The Mayo Clinic Study of Aging is a population-based study in Olmsted County, Minnesota, and the Alzheimer's Disease Research Center is a longitudinal clinical practice study. Both studies have been approved by the Mayo Clinic and Olmsted Medical Center Institutional Review Boards. The ADNI initiative was launched in 2003 with the primary goal of testing whether serial MRI, PET, other biologic markers, and clinical and neuropsychological assessment can be combined to measure progression of mild cognitive impairment (MCI) and AD ([www.adni-info.org](http://www.adni-info.org)). OASIS-3 is a longitudinal study of cognitively normal individuals at various stages of cognitive decline who have undergone MRI and PET (<https://www.oasis-brains.org/>).

### Image Acquisition and Preprocessing

In Mayo, ADNI, and OASIS-3, 370 MBq of  $^{18}\text{F}$ -flortaucipir were administered, with acquisition times varying: Mayo had a 20-min acquisition at 80 min after injection (11), whereas ADNI and OASIS-3 had a 30-min acquisition between 75 and 105 min after injection.

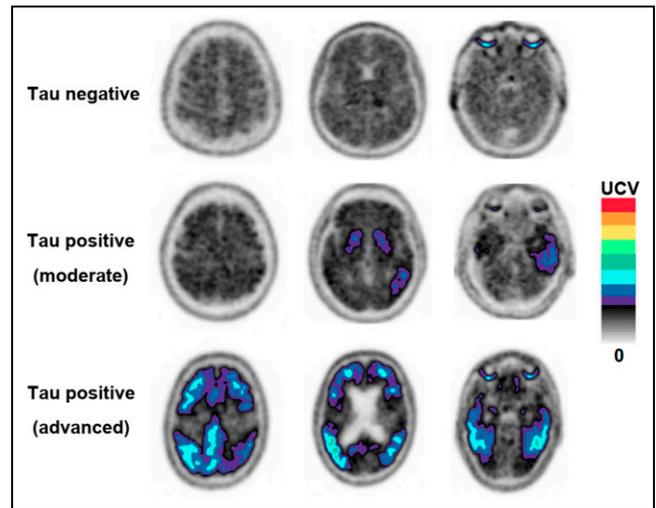
The tau PET scans were rigidly coregistered to corresponding T1-weighted MRI scans, and median values were taken for each region. The MRI scans were initially tissue-class-segmented and divided into atlas regions using the MICALT-ADIR122 atlas (12). Cortical and subcortical regions were referenced to the cerebellar crus median uptake to form SUVR units. These regional SUVR units were used both to form the meta-ROIs and as inputs to our ML models.

### Visual Assessment of Tau PET Scans

Three trained raters independently visually assessed tau PET images scaled to average counts in a 2-dimensional cerebellum ROI. FDA-approved official criteria for visual assessment were followed (8). A positive scan was one with increased neocortical tracer uptake isolated to the posterolateral temporal or occipital or parietal/precuneus regions with or without frontal activity. A negative scan was one with no increased neocortical activity or increased neocortical activity isolated to the mesial temporal, anterolateral temporal, or frontal regions (Fig. 1).

### Tau PET Status Using Meta-ROIs

The temporal meta-ROI (cut point, 1.23 SUVR) was the voxel-weighted average of the median uptake in the entorhinal, amygdala, parahippocampal, fusiform, inferior temporal, and middle temporal regions normalized to the median uptake of the cerebral crus gray matter region used as reference (7). MTL (cut point, 1.30 SUVR) was the



**FIGURE 1.** Visual assessment of tau PET scans used in study. Elevated signals are shown in color. Upper contrast value (UCV) for each participant was calculated from scaled mean cerebellar counts adjusted to match desired levels of color transitions.

unweighted average of MTL tau PET uptake in the left and right entorhinal cortex and amygdala, and the neocortical meta-ROI (cut point, 1.37 SUVR) was the voxel-weighted average in the left and right middle temporal and inferior temporal gyri (13).

### Training and Interpreting the ML Model

The model inputs were regional SUVs, whereas the targets were binary visual ratings. The Mayo dataset was used for training, with 80% ( $n = 1,038$ ) for training and 20% ( $n = 252$ ) for testing, partitioned using stratified splitting (Supplemental Fig. 1; supplemental materials are available at <http://jnm.snmjournals.org>). External validation was done on ADNI and OASIS-3.

A multilayer stack ensemble ML technique using repeated k-fold bagging was applied to train the models (AutoGluon, version 0.8.2) (14). Shapley additive explanation (SHAP), which is a model-agnostic explainable AI technique that calculates input contributions to the model predictions, was used to explain our models (15). SHAP was chosen because of its better performance as reported previously (16).

### Developing the Tau Summary Measure THETA (Tau Heterogeneity Evaluation in Alzheimer's Disease)

We developed a novel tau summary measure, THETA (<https://github.com/ADIR-Lab/THETA>), calculated as a linear combination of 2 parts: the sum of the SHAP ( $\varphi_i$ ) values of the model input SUVs ( $x_i$ ), and the weighted sum of the inputs with their SHAPs that fall within the first and 99th percentiles ( $\hat{\varphi}_i$ ) (Eq. 1). SHAPs represent the individual contribution of each brain region ( $i$ ) to the model's prediction. The first part captures overall feature importance by summing the SHAP values ( $\sum_{i=1}^m \varphi_i$ ) across  $m$  number of regions. It sums up to the predicted binary classification as a consequence of the property of SHAP (15). The second part ( $\sum_{i=1}^m \hat{\varphi}_i x_i$ ) focuses on the weighted contribution of the brain regions whose SHAPs fall within the percentile range. This part scales the SUVs on the basis of their importance for classification and can be important for encoding disease heterogeneity:

$$\Theta = \sum_{i=1}^m \varphi_i + \sum_{i=1}^m \hat{\varphi}_i x_i. \quad \text{Eq. 1}$$

The intuition behind Equation 1 was such that THETA can express the degree of tau deposition quantitatively while accounting for tau spatial heterogeneity. For example, in visually tau-positive concordant

**TABLE 1**  
Population Characteristics

Variable	Mayo	ADNI	OASIS-3
Total number	1,290	831	430
Mean age (y)	67 (SD, 14)	72 (SD, 8)	70 (SD, 8)
Male	706 (55)	399 (48)	186 (43)
Female	584 (45)	432 (52)	244 (57)
CU	957 (74)	455 (55)	371 (86)
MCI	173 (13)	283 (34)	11 (3)
AD	121 (9)	93 (11)	48 (11)
Dementia with Lewy bodies	37 (3)	—	—
APOE4-positive	425 (34)	287 (40)	168 (39)
β-amyloid-positive	512 (40)	335 (43)	133 (32)
Tau-positive			
Visual	245 (19)	230 (28)	61 (14)
MTL	243 (19)	255 (31)	80 (19)
Neocortical	202 (16)	183 (22)	57 (13)
Temporal	476 (37)	418 (50)	159 (37)
MTL and neocortical	183 (14)	170 (20)	49 (11)
Temporal and MTL	235 (18)	242 (29)	74 (17)
Temporal and neocortical	202 (16)	183 (22)	57 (13)

Data are number followed by percentage in parentheses, except for age.

Mayo cases, regional THETA values were higher for mid-temporal and inferior temporal regions in most participants because they displayed a typical AD tau deposition pattern and their rankings clearly showed spatial spread (Supplemental Fig. 2). In these participants, the first part summed up to one and the second component positively weighted the mid-temporal and inferior temporal regions. This becomes useful when, for example, the inferior temporal region is important in 2 concordant cases differing only by their degree of tau PET SUVR. In such a scenario, THETA could appropriately reflect the degree of tau deposition.

Lastly, the repeatability of THETA was assessed, and a high degree of agreement was found when compared with scores generated by various ML models (Supplemental Fig. 3).

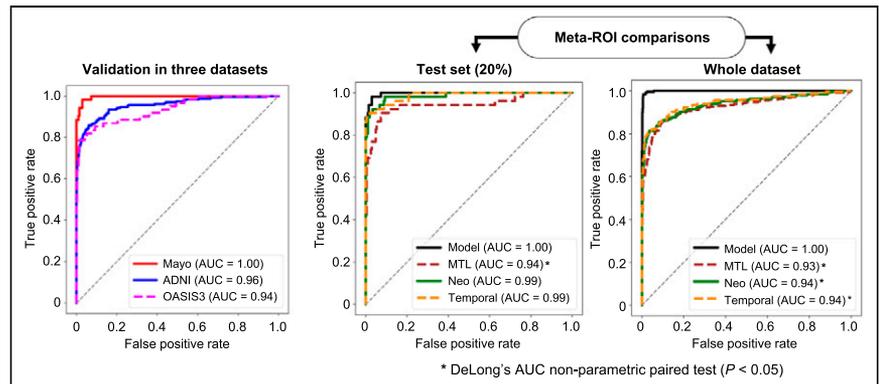
### Statistical Analysis

Model performance was evaluated using metrics such as the Matthew correlation coefficient, balanced accuracy, precision, recall, and the F1 score. The classification was assessed using area under the receiver operating characteristic curve (AUC), and curves were compared using the DeLong AUC nonparametric paired test. True-positive and true-negative rates were used to assess misclassification rates. The Spearman  $\rho$  accompanied by the Choi nonparametric test was used to analyze correlation. The separation between tau positives and tau negatives was examined using Cohen  $d$  and 2-tailed independent-samples  $t$  testing with Bonferroni adjustment.

## RESULTS

### Characteristics of Study Population

In the Mayo cohort (average age, 67 y [SD, 14 y]), 55% were male and 74% were cognitively unimpaired (CU); in ADNI (average age, 72 y [SD, 8 y]), 48% were male and 55% were CU; and in OASIS-3 (average age, 70 y [SD, 8 y]), 43% were male and 86% were CU (Table 1). There was overall sample variability in tau status, with 19% of visibly tau-positive cases in Mayo, 28% in ADNI, and 14% in OASIS-3. Additionally, MTL and neocortical tau positives were 14% for Mayo, 20% for ADNI, and 11% for OASIS-3 (Table 1).



**FIGURE 2.** Classification performance of model in all 3 datasets and comparison to meta-ROIs in test set and whole dataset.

**TABLE 2**  
Models Trained on Mayo and Validated on ADNI and OASIS-3

Dataset	MCC	Balanced accuracy	F1 score	Precision	Recall
Mayo training (80%)	90.68 (3.71)	94.87 (2.51)	92.20 (3.14)	96.36 (2.53)	88.52 (4.87)
Mayo testing (20%)	87.68 (3.43)	92.49 (1.99)	90.00 (2.44)	93.88 (3.60)	86.45 (4.20)
ADNI testing	78.68 (1.69)	86.25 (1.36)	83.05 (1.60)	94.56 (2.63)	74.18 (3.28)
OASIS-3 testing	73.57 (2.87)	82.48 (1.95)	75.79 (2.64)	88.81 (5.19)	66.40 (4.25)

MCC = Matthew correlation coefficient.  
Data are mean and SD.

#### Model Trained on Visual Ratings for Predicting Tau Positivity

After multiple cross-validation runs, the best model based on its highest F1 score was selected. This model showed excellent balanced accuracy, achieving scores of 98.58% during training and 95.43% during testing on the Mayo dataset. It also performed well on the ADNI and OASIS-3 datasets, with scores of 87.74% and 87.03%, respectively.

Tau status was classified with AUCs of 1.00, 0.96, and 0.94 on the Mayo, ADNI, and OASIS-3 datasets (Fig. 2; Table 2).

#### Model Performance in Comparison to Meta-ROIs

The meta-ROIs classified tau positivity with an AUC of 0.99 in the Mayo test set (20%) and 0.94 overall (Fig. 2). The model outperformed the meta-ROIs in the Mayo dataset, with the lowest misclassification rates for both tau-positive and tau-negative cases (3.67% and 0.48%, respectively). In the ADNI dataset, the temporal meta-ROI showed the least tau-positivity misclassification (6.96%). In the OASIS-3 dataset, the model misclassified tau negatives at the lowest rate (1.36%) but was second best in

tau-positive misclassification (24.59%), surpassed by the temporal meta-ROI (18.03%) (Table 3).

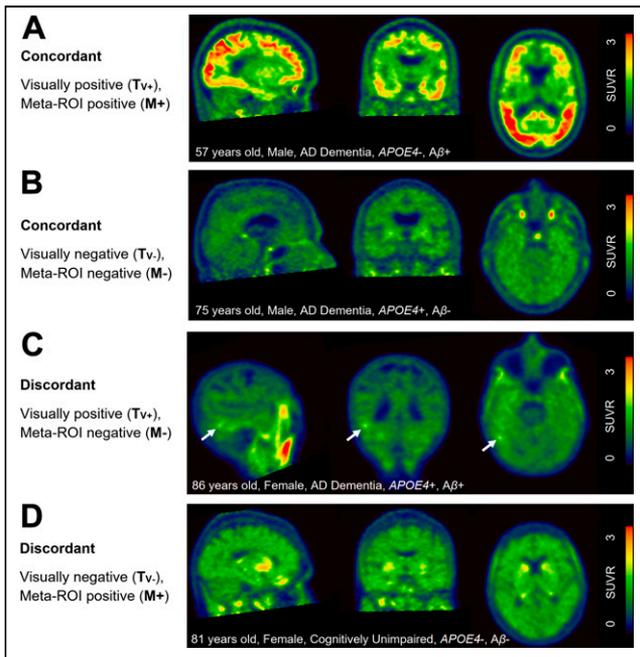
#### THETA

*Quantification of Spatial Heterogeneity in Tau PET Scans.* We demonstrate tau spatial heterogeneity captured by THETA in 2 subpopulations: discordant and concordant. The discordant group included cases in which the visual rating disagreed with one or more of the meta-ROI classifications, whereas the concordant group included cases in which both methods agreed (Fig. 3).

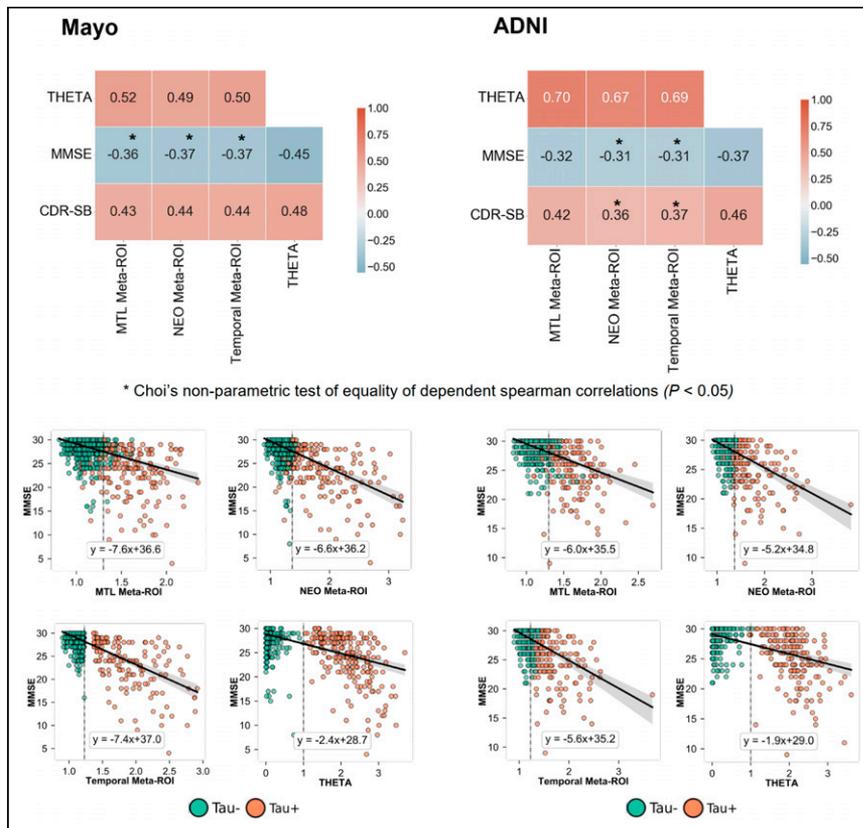
THETA's formulation accurately reflected the heterogeneous contributions of all regions in the discordant cases having mild signals and similar regional contributions. In this group, the input contributions were distributed among different regions rather than being focused specifically on the meta-ROIs (Supplemental Figs. 2 and 4). In the tau-positive concordant cases, our ML model's hot-spot regions were the top predictors that constituted the meta-ROIs. In these cases, THETA maintained the importance of top regions, preserving spatial heterogeneity.

**TABLE 3**  
Comparison of Meta-ROIs and ML Model

Tau-positivity comparison	True-positive rate (%)	True-negative rate (%)	True-positive (n)	True-negative (n)	1 – true-positive rate (%)	1 – true-negative rate (%)
<b>Mayo</b>						
Visual vs. temporal	0.93	0.76	227	796	7.34	23.83
Visual vs. MTL	0.78	0.95	191	993	22.04	4.97
Visual vs. neocortical	0.76	0.98	185	1,028	24.49	1.63
Visual vs. model	0.96	1.00	236	1,040	3.67	0.48
<b>ADNI</b>						
Visual vs. temporal	0.93	0.66	214	397	6.96	33.94
Visual vs. MTL	0.78	0.88	180	526	21.74	12.48
Visual vs. neocortical	0.70	0.96	161	579	30.00	3.66
Visual vs. model	0.78	0.98	179	587	22.17	2.33
<b>OASIS-3</b>						
Visual vs. temporal	0.82	0.70	50	260	18.03	29.54
Visual vs. MTL	0.66	0.89	40	329	34.43	10.84
Visual vs. neocortical	0.66	0.95	40	352	34.43	4.61
Visual vs. model	0.75	0.98	46	364	24.59	1.36



**FIGURE 3.** Examples of concordant groups (with agreement between visual rating and meta-ROIs) (A and B) and discordant groups (C and D). Although meta-ROIs can miss visually positive scans in which SUVR is lower than cut point (C), visual assessment does not consider isolated increased activity in MTL (D). Arrows indicate increased tracer uptake.  $A\beta = \beta$ -amyloid.



**FIGURE 4.** Comparison of meta-ROIs and THETA to clinical measures (Mini-Mental State Examination [MMSE] and Clinical Dementia Rating–Sum of Boxes [CDR-SB]). Tau negativity (T-) and tau positivity (T+) are visual assessments, and meta-ROI cut points are shown by dotted vertical lines. OASIS-3 results are shown in Supplemental Figure 6.

*Association of THETA with Clinical Disease Severity Markers.* THETA correlated with the Mini-Mental State Examination significantly better than all meta-ROIs for the Mayo dataset based on the Choi nonparametric test of Spearman correlations (THETA  $\rho = -0.45$ , meta-ROIs  $\rho \geq -0.36$ , Choi  $P < 0.05$ ) (Fig. 4). For the ADNI dataset, the correlation of THETA with the Mini-Mental State Examination was significantly better than the neocortical and temporal meta-ROIs (THETA  $\rho = -0.37$ , neocortical and temporal  $\rho = -0.31$ ,  $P < 0.05$ ). THETA did not significantly correlate with the Clinical Dementia Rating–Sum of Boxes compared with the meta-ROIs in the Mayo and OASIS-3 datasets but correlated significantly with the neocortical and temporal meta-ROIs in ADNI (THETA  $\rho = 0.46$ , meta-ROIs  $\rho \geq 0.36$ , Choi  $P < 0.05$ ). A similar pattern of association with clinical indices was also observed in Mayo-concordant groups. However, in Mayo-discordant groups, THETA demonstrated a significantly negative correlation with the Mini-Mental State Examination (THETA  $\rho = -0.43$ , meta-ROIs  $\rho \geq -0.07$ , Choi  $P < 0.05$ ) and a significantly positive correlation with the Clinical Dementia Rating–Sum of Boxes (THETA  $\rho = 0.37$ , meta-ROIs  $\rho = 0.05$ , Choi  $P < 0.05$ ) compared with the meta-ROIs (Fig. 5).

*Separation of Tau-Positive from Tau-Negative Scans.* THETA provided greater separation between visually tau-positive and -negative cases, a characteristic attributed to its development from a model trained on the binary visual ratings; however, the meta-ROIs did not provide greater separation either according to visual classifications or according to their respective cut points (Fig. 4).

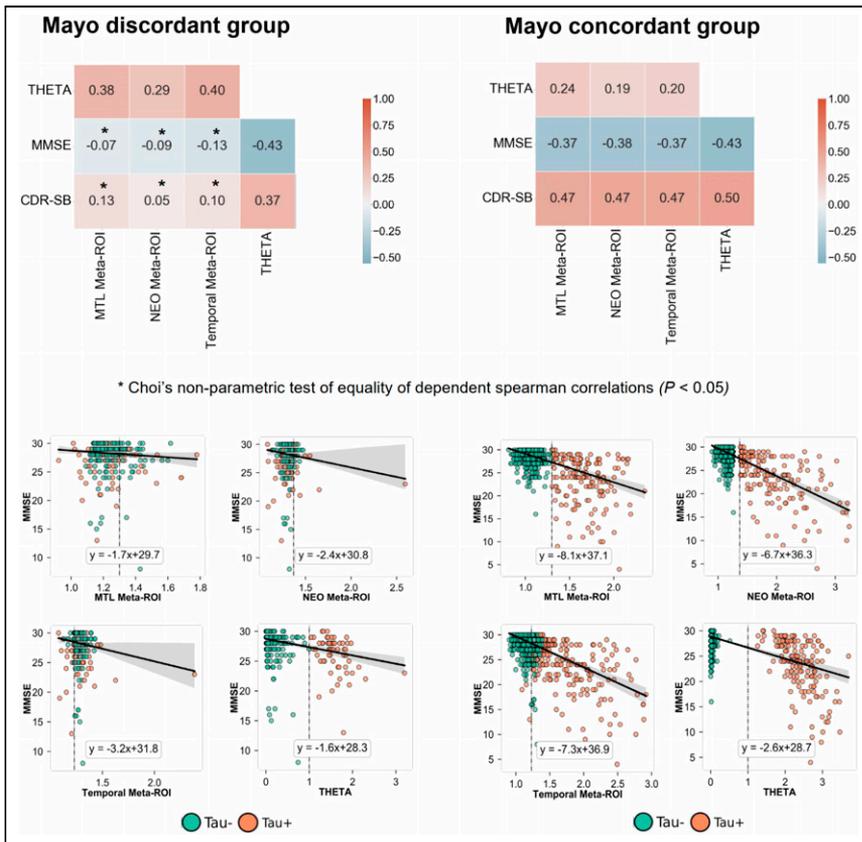
When THETA was compared with the temporal meta-ROI across various clinical diagnostic outcomes, the separation between tau-

positive and tau-negative cases was similar for AD dementia cases. However, for CU and MCI, there was an overlap in tau status for the temporal meta-ROI, whereas THETA showed better separation between cases (Fig. 6). For instance, in ADNI, the difference between tau-positive and tau-negative temporal meta-ROI values for CU and MCI participants had an effect size of 3.08 ( $t = 16.50$ ,  $P < 0.001$ ) and 2.23 ( $t = 16.76$ ,  $P < 0.001$ ), respectively. THETA showed a larger effect size of 10.09 ( $t = 54.09$ ,  $P < 0.001$ ) and 6.83 ( $t = 51.36$ ,  $P < 0.001$ ), respectively (Fig. 6).

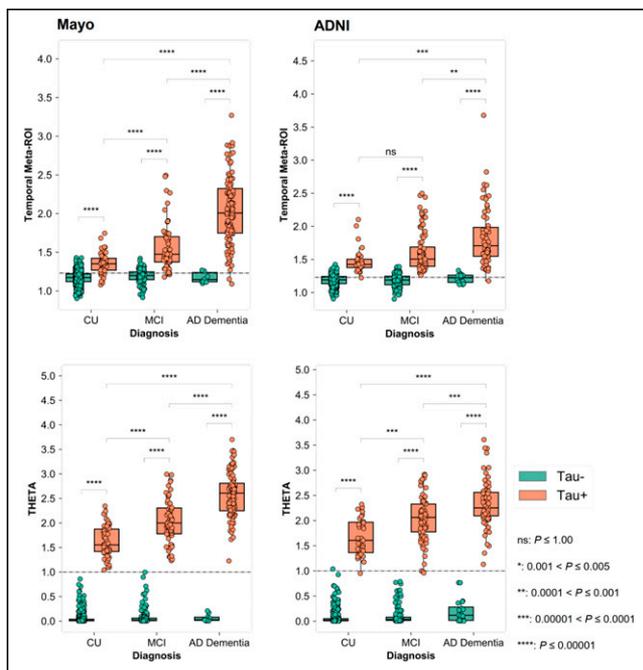
*Association of THETA with Braak Staging.* Mayo participants with Braak staging (I–VI) ( $n = 90$ ) were included for histopathologic validation of THETA (Supplemental Table 3). The association of THETA with Braak staging was similar to the meta-ROIs ( $\rho = 0.87$  for THETA,  $P < 0.05$  vs.  $P \leq 0.83$ ;  $P < 0.05$  for meta-ROIs). However, improved separation of clinical diagnostic groups was observed in THETA compared with the meta-ROIs (Fig. 7).

## DISCUSSION

Tau pathology, as captured by tau PET scans, is a key indicator of disease severity in AD (17). Current tau PET quantitative techniques have limitations in addressing the heterogeneity of tau deposition because



**FIGURE 5.** Comparison of meta-ROIs and THETA to clinical measures (Mini-Mental State Examination [MMSE] and Clinical Dementia Rating–Sum of Boxes [CDR-SB]) for discordant and concordant Mayo groups. Discordant groups showed disagreement between visual rating and meta-ROIs, whereas concordant groups showed agreement between visual and meta-ROI methods. Tau negativity (T-) and tau positivity (T+) are visual assessments, and meta-ROI cut points are shown by dotted vertical lines.

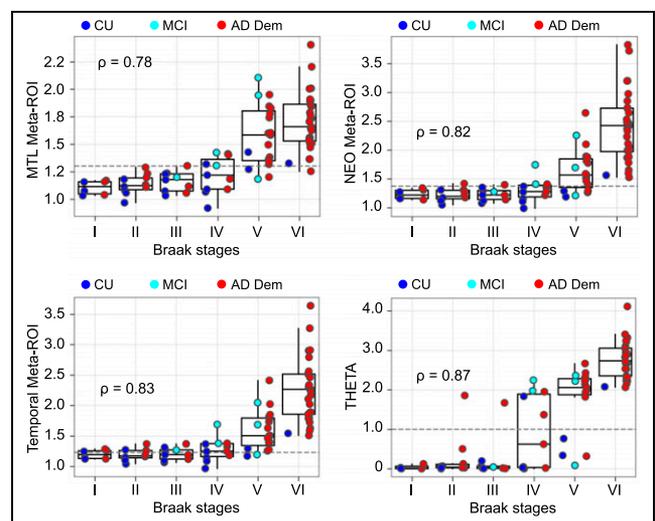


**FIGURE 6.** Comparison of temporal meta-ROI and THETA in diagnostic groups (CU, MCI, and AD dementia). Tau negativity (T-) and tau positivity (T+) are based on visual assessment. ns = not statistically significant.

they focus on regions with typically high tau uptake while ignoring the spatial variance of tau burden. Using visual assessment by 3 raters as the gold standard in a large single-site dataset (Mayo), we developed an ML model to accurately classify tau PET scans. We then validated our model in 2 independent datasets (ADNI and OASIS-3) and developed a summary metric, THETA, to consider tau SUVRs across the entire brain and map well to disease progression. The clinical validations showed that THETA was better associated with indicators of cognitive impairment, provided a greater magnitude of separation between clinical diagnosis groups, and was better associated with AD pathology-based staging than are meta-ROIs.

The FDA approved  $^{18}\text{F}$ -florbetapir for detecting neurofibrillary tangles at Braak stages V and VI (8). Effective quantification of tau PET has paramount importance as it leads to more sensitive assessments of disease severity. Given that visual evaluation is a clinically accepted practice, a model incorporating multirater assessment as a gold standard can be sensitive in identifying positive scans. Moreover, using such a model to develop a summary metric can be beneficial. Our work takes advantage of this model to accurately quantify tau PET scans using regional SUVRs to address heterogeneous tau deposition across the entire brain.

Our results show that THETA mapped onto clinical diagnostic and cognitive indices comparably to or better than meta-ROIs while retaining pattern-based information. This was further examined by inspecting the feature importance of participants with neocortical versus MTL tau positivity. For participants with neocortically dominant tau positivity,



**FIGURE 7.** Association of meta-ROIs and THETA to Braak stages. Cut points are shown by dotted lines. NEO = neocortex.

the inferior temporal cortex was the top predictor, whereas in participants with MTL tau positivity—a region notably affected by tau deposition—the entorhinal cortex was a key predictor (Supplemental Fig. 4). Furthermore, THETA demonstrated a clearer separation of clinical diagnoses at higher Braak stages (V and VI) better than the meta-ROIs. Given that a higher Braak stage is not always associated with cognitive decline, particularly in resilient groups (16), THETA effectively filtered out higher Braak (V) tau-negative CU participants, indicating its utility to quantify tau burden and better associate with clinical diagnosis (Fig. 7). Nonetheless, the histopathology sample used here was small, and further studies are needed to investigate the utility of THETA in identifying resilient and resistant groups.

Meta-ROIs and visual ratings have different approaches to measuring tau pathology. Meta-ROIs focus on the entorhinal cortex, which can overestimate tau positivity, whereas visual assessments can miss isolated tau in the MTL and anterolateral temporal lobe (18). Visual assessments offer the possibility of assessing tau burden throughout the brain and can be superior because the meta-ROIs rely on specific regions. Nonetheless, we found in our Mayo cohort that all 3 meta-ROIs underperformed in the visual identification of tau-positive cases. The neocortical meta-ROI's true negative rate was consistent across all 3 datasets, whereas the MTL did better in the identification of true negatives in Mayo and decreased in performance in ADNI and OASIS-3.

This study had some strengths and limitations. A strength is that both the ML model and THETA were validated on 2 independent datasets. A limitation is that visual assessment was FDA-approved specifically for moderate to higher levels of tau. Although this method ensures exclusion of off-target binding, the trained model could be less sensitive to early tau changes. Further studies are needed to assess the effectiveness of THETA to identify early tau changes. Another limitation is that the model's performance was lower for the ADNI and OASIS-3 validation sets because of cohort differences. However, combining them and training a new model achieved a balanced accuracy of more than 94% and an AUC of more than 0.99 (Supplemental Table 4; Supplemental Fig. 5). Finally, THETA exhibits sensitivity for visual tau positives, which can be a strength or a limitation. A longitudinal sensitivity analysis found no significant difference in performance from meta-ROIs (Supplemental Fig. 6). Lastly, changing the cut points for meta-ROIs could change the results presented for the meta-ROIs.

## CONCLUSION

We were able to develop an ML model to predict tau status on single-site data and validate it on external datasets. We also developed a tau summary measure, THETA, which captured the spatial heterogeneity of tau and showed higher sensitivity to clinical disease severity. Overall, this study provided promising results for using ML to improve the detection and quantification of tau pathology in AD.

## DISCLOSURE

Support was received from a Mayo Clinic Radiology Research fellowship; NIH grants R01 NS097495, R01 AG056366, U01 AG006786, P50 AG016574, RF1AG069052, UF1NS125417, R37

AG011378, R01 AG041851, R01 AG073282, NIH P30 AG066444, P50 AG00561, P30 NS09857781, P01 AG026276, P01 AG003991, R01 AG043434, UL1 TR000448, and R01 EB009352; a GHR Foundation grant; the Alexander Family Alzheimer's Disease Research Professorship of Mayo Foundation; the Elsie and Marvin Deikelbom Family Foundation U.S.A.; Opus building NIH grant C06 RR018898; and Rochester Epidemiology Project (REP) medical records-linkage system resources, supported by the National Institute on Aging (NIA; AG 058738), by the Mayo Clinic Research Committee, and by fees paid annually by REP users. Data collection and sharing for this project were funded by the ADNI (NIH grant U01 AG024904) and DOD ADNI (Department of Defense award W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging, by the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from AbbVie, the Alzheimer's Association, the 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](http://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 funders had no role in the design or conduct of the study. No other potential conflict of interest relevant to this article was reported.

## ACKNOWLEDGMENTS

Data used in preparation of this article were obtained from the ADNI database ([adni.loni.usc.edu](http://adni.loni.usc.edu)). As such, the investigators within the ADNI contributed to the design and implementation of the ADNI or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found online ([https://adni.loni.usc.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](https://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)). We acknowledge Valentina Garibotto, Ismini Maita, Valle Camacho, Martijn van Essen, Andrés Perissinotti, Omar Rodríguez, and Kohl L. Johnson Sparman for visually rating some participants. We thank AVID Radiopharmaceuticals, Inc., for its support in supplying the AV-1451 precursor, for chemistry production advice and oversight, and for the needed FDA regulatory cross-filing permission and documentation. Data were provided in part by OASIS-3\_AV1451; the longitudinal multimodal neuroimaging principal investigators were Tammie L.S. Benzinger, Daniel Marcus, and John C. Morris; NIH P30 AG066444, AW000093; AV-1451 doses were provided by Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly.

## KEY POINTS

**QUESTIONS:** Can an ML model trained using visual ratings identify positive tau PET scans and aid in developing a summary metric that considers spatial heterogeneity?

**PERTINENT FINDINGS:** Our model, trained on a single-site cohort, was accurate in classifying tau-positive status and was able to generalize well to external cohorts. The summary metric THETA was able to associate with clinical cognitive markers, clinical diagnoses, and histopathology comparably to or better than meta-ROIs.

**IMPLICATIONS FOR PATIENT CARE:** Our model and THETA can enhance information extracted from tau PET for improved clinical diagnoses, disease progression tracking, AD subtyping, and identification of therapeutic targets, for better personalized and effective patient care.

## REFERENCES

1. Braak H, Alafuzoff I, Arzberger T, Kretschmar H, Del Tredici K. Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol (Berl)*. 2006;112:389–404.
2. La Joie R, Visani AV, Baker SL, et al. Prospective longitudinal atrophy in Alzheimer's disease correlates with the intensity and topography of baseline tau-PET. *Sci Transl Med*. 2020;12:eaa5732.
3. Vogel JW, Young AL, Oxtoby NP, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nat Med*. 2021;27:871–881.
4. Murray ME, Graff-Radford NR, Ross OA, Petersen RC, Duara R, Dickson DW. Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *Lancet Neurol*. 2011;10:785–796.
5. Schöll M, Lockhart SN, Schonhaut DR, et al. PET imaging of tau deposition in the aging human brain. *Neuron*. 2016;89:971–982.
6. Young CB, Winer JR, Younes K, et al. Divergent cortical tau positron emission tomography patterns among patients with preclinical Alzheimer disease. *JAMA Neurol*. 2022;79:592–603.
7. Jack CR, Wiste HJ, Weigand SD, et al. Defining imaging biomarker cut points for brain aging and Alzheimer's disease. *Alzheimers Dement*. 2017;13:205–216.
8. Jie CV, Treyer V, Schibli R, Mu L. Tauvid™: the first FDA-approved PET tracer for imaging tau pathology in Alzheimer's disease. *Pharmaceuticals (Basel)*. 2021;14:110.
9. Roberts RO, Geda YE, Knopman DS, et al. The Mayo Clinic study of aging: design and sampling, participation, baseline measures and sample characteristics. *Neuroepidemiology*. 2008;30:58–69.
10. LaMontagne PJ, Benzinger TLS, Morris JC, et al. P3-083: OASIS-3—longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease. medRxiv website. <https://www.medrxiv.org/content/10.1101/2019.12.13.19014902v1>. Published December 15, 2019. Accessed July 15, 2024.
11. Schwarz CG, Thorneau TM, Weigand SD, et al. Selecting software pipelines for change in flortaucipir SUVR: balancing repeatability and group separation. *Neuroimage*. 2021;238:118259.
12. Schwarz CG, Gunter JL, Wiste HJ, et al. A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. *Neuroimage Clin*. 2016;11:802–812.
13. Ossenkoppele R, Pichet Binette A, Groot C, et al. Amyloid and tau PET-positive cognitively unimpaired individuals are at high risk for future cognitive decline. *Nat Med*. 2022;28:2381–2387.
14. Erickson N, Mueller JW, Shirkov A, et al. AutoGluon-Tabular: robust and accurate AutoML for structured data. ArXiv website. <https://arxiv.org/abs/2003.06505>. Published March 13, 2020. Accessed July 9, 2024.
15. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: *NIPS'17: Proceedings of the 31<sup>st</sup> International Conference on Neural Information Processing Systems*. Curran Associates Inc.; 2017:4768–4777.
16. Band SS, Yarahmadi A, Hsu C-C, et al. Application of explainable artificial intelligence in medical health: a systematic review of interpretability methods. *Inform Med Unlocked*. 2023;40:101286.
17. Kolb HC, Andrés JI. Tau positron emission tomography imaging. *Cold Spring Harb Perspect Biol*. 2017;9:a023721.
18. Soleimani-Meigooni DN, Rabinovici GD. Tau PET visual reads: research and clinical applications and future directions. *J Nucl Med*. 2023;64:822–824.