### Title

Tau-Atrophy Variability Reveals Phenotypic Heterogeneity in Alzheimer's Disease




### Journal

Annals of Neurology, 90(5)

### ISSN

0364-5134



### Authors

Das, Sandhitsu R

Lyu, Xueying

Duong, Michael Tran

et al.



### Publication Date

2021-11-01

### DOI

10.1002/ana.26233

Peer reviewed

Published in final edited form as:

*Ann Neurol*. 2021 November ; 90(5): 751–762. doi:10.1002/ana.26233.

# Tau-atrophy variability reveals phenotypic heterogeneity in Alzheimer's disease

**Sandhitsu R. Das, Ph.D.**[1], **Xueying Lyu**[4], **Michael Tran Duong**[4], **Long Xie, Ph.D.**[2], **Lauren McCollum, M.D.**[6], **Robin de Flores, Ph.D.**[7], **Michael DiCalogero**[1], **David J. Irwin, M.D.**[1], **Bradford C. Dickerson, M.D.**[5], **Ilya M. Nasrallah, M.D.**[2], **Paul A. Yushkevich, Ph.D.**[2], **David A. Wolk, M.D.**[1], **Alzheimer's Disease Neuroimaging Initiative**[*]

[1]Department of Neurology, University of Pennsylvania, Philadelphia, PA, USA

[2]Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA

[4]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA

[5]Massachusetts General Hospital, Boston, MA, USA

[6]Department of Medicine, University of Tennessee, TN, USA

[7]Université de Caen Normandie, INSERM UMRS U1237, Caen, France

## Abstract

**Objective:** Tau neurofibrillary tangles (T) are the primary driver of downstream neurodegeneration (N) and subsequent cognitive impairment in Alzheimer's disease (AD). However, there is substantial variability in the T-N relationship – manifested in higher or lower atrophy than expected for level of tau in a given brain region. The goal of this study was to determine if region-based quantitation of this variability allows for identification of underlying modulatory factors, including polypathology.

**Methods:** Cortical thickness (N) and $^{18}$F-Flortaucipir SUVR (T) were computed in 104 gray matter regions from a cohort of cognitively-impaired, amyloid-positive (A+) individuals. Region-specific residuals from a robust linear fit between SUVR and cortical thickness were computed as a surrogate for T-N mismatch. A summary T-N mismatch metric defined using residuals were correlated with demographic and imaging-based modulatory factors, and to partition the cohort into data-driven subgroups.

Corresponding author: Sandhitsu Das, Department of Neurology, University of Pennsylvania, 3700 Market Street, Fl 6, Philadelphia, PA 19104, sudas@seas.upenn.edu, Phone: 2674083001.

**Author Contributions**: SD, DW, PY, BD, and IMN contributed to the conception and design of the study; SD, DW, MTD, LX, LM, RF, XL, and MD contributed to acquisition and data analysis; SD, DW, PY, BD, DJI, LX, XL, MTD and IMN contributed to drafting the text or preparing the figures.

[*]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

**Potential Conflicts of Interest**: DW was a site-PI for an Eli Lilly study who produces Florbetapir and Flortaucipir tracers used in this study, and is a consultant for GE Healthcare who produces another amyloid PET ligand. BD reports receiving consulting fees from Eli Lily. IMN ran educational sessions for Biogen. SD, PY, DJI, XL, MTD, LX, LM, RF and MD do not report any related conflicts of interest.

**Results:** The summary T-N mismatch metric correlated with underlying factors such as age and burden of white matter hyperintensity lesions. Data-driven subgroups based on clustering of residuals appear to represent different biologically relevant phenotypes, with groups showing distinct spatial patterns of higher or lower atrophy than expected.

**Interpretation:** These data support the notion that a measure of deviation from a normative relationship between tau burden and neurodegeneration across brain regions in individuals on the AD continuum captures variability due to multiple underlying factors, and can reveal phenotypes, which if validated, may help identify possible contributors to neurodegeneration in addition to tau, which may ultimately be useful for cohort selection in clinical trials.

### Keywords

Tau; PET; MRI

## Introduction

In a shift towards a biological, rather than clinical, definition of Alzheimer's Disease (AD), the National Institute on Aging – Alzheimer's Association (NIA-AA) recently formalized a biomarker driven classification referred to as the A/T/(N) framework[1]. Following this system, individuals receive dichotomous designations for the presence (+) or absence (−) of Beta-amyloid plaques (A), tau-based neurofibrillary tangles (T), and neurodegeneration (N). To be classified as AD, one must be both A+ and T+ paralleling the pathological criteria for AD neuropathologic change (ADNC). Neurodegenerative markers are meant to support biological staging of disease.

Neurofibrillary tangles (NFT) are more tightly linked to downstream neurodegenerative change than amyloid plaques[2–6]. As such, the A/T/(N) framework posits that neurodegeneration without NFTs (T-) is due to the presence of suspected non-AD pathophysiology (SNAP) regardless of the presence or absence of amyloid[7,8].

While the dichotomous designation of these biomarkers provides a simplicity in interpretation, it may blunt the information provided about an individual's underlying disease status. In particular, T and N imaging biomarkers provide continuous and spatially varying information that are likely important for understanding phenotype. As NFTs are the putative primary driver of neurodegeneration in AD, T is expected to be tightly linked to N. Thus, discordance between T and N suggests additional non-AD modulators. While the extreme of T− and N+ supports SNAP, even in a T+ individual, a relatively larger magnitude of N than the magnitude of T also may support concomitant SNAP. Further, the spatial pattern of this discordance may support different underlying non-AD pathologies, given the differential loci of greatest neurodegeneration with these conditions[9]. For example, greater anterior temporal atrophy than expected for the amount of local NFT pathology may suggest concomitant limbic-predominant age-associated TDP-43 encephalopathy (LATE)[10,11] which often co-occurs with AD and primarily affects this region. Alternatively, less neurodegeneration than expected for a given degree of T, may suggest resilience to AD pathology, or brain reserve[12].

Given that co-pathology with AD is common[13] and has implications for clinical interventions and prognosis, and that definitive biomarkers for non-AD molecular pathology are lacking[14], approaches that can operationalize the presence of co-pathology are critical, as are measures that capture resilience. The current study attempts to provide greater precision to the A/T/(N) framework by treating T and N as continuous variables and then determining their relationship in a way that is "*spatially aware*". We first establish the normative relationships between T and N at regions of interest (ROIs) across the cortical mantle in people on the AD continuum (A+). Deviation from this relationship at each ROI constitutes *mismatch*. We predicted that a quantitative measure of this T-N mismatch would be associated with factors that may reflect non-AD neurodegeneration, such as the presence of cerebrovascular disease or age, as well as with cognitive performance. Further, we examined whether using data-driven clustering based on spatial pattern of these T-N mismatch metrics would produce distinct groups that would offer insight into the presence of different concomitant pathologies or brain reserve/resilience.

## Materials and Methods:

### Participants:

Data used in the preparation of this article were obtained from the ADNI database. The current study included 137 A+ cognitively impaired participants with diagnosis of MCI or dementia. The number of datasets were based on data available for download in January 2020. They included 80 MCI (age 76.1 ± 7.5 years, 38 Female, 15.8 ± 2.7 years of education, MMSE score of 27.1 ± 2.4), and 57 Dementia cases (age 77.7 ± 9.5 years, 27 Female, 15.9 ± 2.5 years of education, MMSE score of 22.5 ± 4.1).

### Image acquisition:

ADNI MR imaging included a T1-weighted structural scan of resolution 1.0×1.0× 1.2 $mm^3$, and a FLAIR MRI acquired in the same session with variable spatial resolution as prescribed in the ADNI protocol. Tau PET imaging consisted of a continuous 30-minute brain scan (six 5-minute frames) started 75 minutes following injection of approximately 10 mCi of $^{18}$F-Flortaucipir injection. Most patients (N=106) had an amyloid PET scan, with a 20 min brain scan (four 5-minute frames) performed 50 min after an approximately 10 mCi injection of the radiotracer $^{18}$F-Florbetapir. The rest of the participants received an $^{18}$F-Florbetaben Amyloid PET scan. After an injection of approximately 8.1 mCi of $^{18}$F-Florbetaben and a 90 min uptake phase, a 20 min brain scan (four 5-minute frames) was performed. The respective PET scans acquired closest in time to the structural MRI were analyzed. PET images were downloaded from the ADNI data archive in the most fully pre-processed format with the image description of "Coreg, Avg, Std Img and Vox Siz, Uniform Resolution".

### Image processing:

The pre-processed PET images were aligned with the anatomical MRI via rigid registration using the ANTs normalization software[15] with a mutual information metric. The anatomical MRI was parcellated into ROIs including cortical, subcortical and cerebellar ROIs using a multi-atlas segmentation method[16]. The parcellation scheme is described elsewhere[17]. Mean tracer uptake in cerebellar gray matter and gray+white matter was used as reference

to generate an SUVR map for the entire brain for $^{18}$F-Flortaucipir and $^{18}$F-Florbetapir/$^{18}$F-Florbetaben, respectively. A composite ROI consisting of middle frontal, anterior cingulate, posterior cingulate, inferior parietal, precuneus, supra marginal, middle temporal and superior temporal cortex was used to compute a global SUVR for Amyloid PET scans. Thresholds of SUVR ≥ 1.11 and 1.08 for $^{18}$F-Florbetapir and $^{18}$F-Florbetaben respectively, were then used to determine Amyloid-β status, as previously described[18]. *Composite SUVR provided by ADNI are now available for all but one participant (variable SUMMARYSUVR_WHOLECEREBNORM) and had a correlation of r=0.97 with those used here. Amyloid-β status was identical in all participants*.

The T1-weighted structural MRI was processed using the ANTs cortical thickness pipeline[19] that implements a diffeomorphic registration-based thickness estimation method[20]. ROI-based measures of tau (T) and neurodegeneration (N) were then computed as average $^{18}$F-Flortaucipir SUVR and cortical thickness respectively within 104 bilateral ROIs (52 from each hemisphere).

White matter hyperintensity masks were computed from FLAIR images using a previously validated deep learning-based method that was the top performer in the WMH segmentation challenge[21]. The masks were visually examined by a trained rater and edited as necessary before computing volumes.

## Modeling of Tau (T) – Neurodegeneration (N) relationship: a T-N mismatch metric

Tau-Neurodegeneration relationship was modeled using average gray matter thickness (N) and $^{18}$F-Flortaucipir SUVR (T) for 104 bilateral ROIs in each subject. A robust linear regression[22] was performed at each ROI. To mitigate effects of a skewed distribution of SUVR, the common logarithmic transform was applied to T. Additionally, a bi-square weighting function which can further minimize the influence of outliers was used for robust regression. An example is shown in Figure 1 (a).

For each individual, a summary measure, *T-N mismatch metric*, was calculated by taking the difference between the number of ROIs that had residuals below 1.5 × standard deviation from the regression line (indicating greater than expected N for the given level of T load) and number of ROIs that had residuals above 1.5 standard deviation from the regression line (indicating less than expected N for the given level of T load). Thus, a positive *T-N mismatch metric* suggests generally more atrophy than expected for tau (i.e more vulnerable) while a negative *T-N mismatch metric* suggests greater cortical thickness than expected for tau burden (i.e. resilience or brain reserve). The relationship of this metric with demographic variables and cognitive performance scores were analyzed.

## Clustering for phenotype discovery

The regression residual for each ROI was discretized into a two-element binary vector based on whether it was farther than 1.5 standard deviation away from the regression line and if the residual was negative or positive. These binary vectors were then entered into Ward's hierarchical agglomerative clustering method[23] to generate data-driven groupings of subjects. *The number of clusters (6) was specified empirically by the elbow method*

*proposed by Thorndike*[24] *that uses the percentage of variance explained as a function of number of clusters and optimizes within-cluster similarity.*

## Results

Pearson correlation between the summary T-N mismatch metric, defined as the difference in number of ROIs with negative versus positive residuals of greater than 1.5 × standard deviation, and age and WMH volume were calculated (Table 1, Figure 1 (b),(d)). As expected, age ($r=0.38$, $P<10^{-5}$) and WMH volume ($r=0.35$, $p<0.0001$) were strongly correlated with the mismatch metric, such that older age and greater WMH volume were associated with a higher T-N mismatch metric. In other words, these factors were associated with more atrophy than expected for a given level of tau pathology within this study population. Moreover, when the relationship between WMH volume and T-N mismatch was covaried for age, the correlation remained significant ($r = 0.31$, $p<0.001$).

To determine whether the T-N mismatch metric was associated with the degree of tau pathology, we examined correlations with $^{18}$F-Flortaucipir uptake in areas typically affected by AD: entorhinal cortex, inferior temporal cortex, precuneus, and angular gyrus. We found no significant relationship of the T-N mismatch metric to tau deposition in these regions, supporting the notion that T-N mismatch is not a reflection of AD severity, as determined by tau PET.

We also examined the relationship of this summary T-N mismatch metric with performance in various cognitive tasks, with age, gender, years of education and IT tau SUVR as covariates. The latter, an early region of tau pathology, was used as a surrogate measure of tau load to determine the degree to which mismatch influenced performance beyond AD NFT pathology. Cognitive tasks included a global measure (MMSE), measures of memory (Logical Memory and AVLT delayed recall), language (category fluency), visuomotor speed (Trailmaking Test A) and executive function (Trailmaking Test B). These results are summarized in Table 2. Significant correlations were found with MMSE, AVLT, category fluency scores and Trailmaking Test A while Trailmaking Test B was borderline.

Correlational analyses between T-N mismatch metric and demographic, imaging, and cognitive performance measures reported above were repeated with an additional covariate of composite amyloid PET SUVR. Amyloid SUVR was not a significant predictor for any of the analyses.

### Data driven clustering of participants based on T-N mismatch metric

Participants were grouped into six clusters using hierarchical clustering. They appear to represent different imaging phenotypes with distinct spatial topography, which may be a proxy for underlying disease phenotypes, as summarized in Table 3. Overall, the groups differed in age, burden of white matter hyperintensity (WMH), clinical status, and MMSE. However, they did not differ in tau burden in key AD ROIs (e.g. inferotemporal cortex, precuneus, angular gyrus) based on $^{18}$F-Flortaucipir uptake.

Group 1 was the largest group (n=61) and had atrophy and tau load measures closest to the regression line. We labeled this group “canonical” since it defines the expected relationship between neurodegeneration and tau burden. Group 2 (n=24) displayed greater atrophy than expected for $^{18}$F-Flortaucipir uptake particularly in anterior temporal/temporal limbic regions (Figure 2). The group was slightly older than the canonical group with significantly greater WMH burden (p=0.01) and slightly lower MMSE. A number of cases in Group 2 had particularly severe anterior hippocampal/temporal polar atrophy suggestive of temporal lobe pathology such as concomitant TDP-43[10,11] or hippocampal sclerosis[25], as described in autopsy studies. Alternatively, Group 3 (n=13) appeared to be a resilient imaging phenotype with less atrophy relative to tau, particularly in lateral cortical regions (Figure 2). While Group 3 was significantly younger than the canonical group (p=0.01), this group did include older individuals who appeared to have minimal atrophy despite significant $^{18}$F-Flortaucipir uptake as illustrated by the case in Figure 2. This group also had higher mean MMSE scores than the canonical group. Group 4 (n=3) was quite small and had severe, diffuse atrophy greater than expected for tau in both limbic and cortical regions. This group was marked by being the oldest (age range of 82–94). Group 5 (n=25) displayed resilience, particularly in temporal and frontal regions. This group was similar in age to the canonical group but had less WMH. However, Group 5 did not have significant difference in MMSE from the canonical group. Finally, Group 6 (n=11) had widespread higher atrophy relative to tau. This group was marginally older than the canonical group, but displayed significantly higher WMH volume (p=0.01) and lower MMSE. Some individuals within this group had quite severe WMH, as exemplified by the case in Figure 2. *The entirety of the T-N regression residual data used for clustering is shown in* Figure 3, *along with a summary of measure of percentage of participants that had a suprathreshold residual for each ROI.*

We also performed an analysis of longitudinal changes in MMSE across the groups using linear mixed effects model. Followup period of cognitive data used was between 1–4 years and a maximum of 4 time points. Time*cluster interaction was entered as a predictor and participants as random effect; covariates included age, gender and years of education. Time*cluster interaction was significant (p = 0.04) indicating the groups differed in their rate of decline.

## Sensitivity analyses

We conducted the following sensitivity analyses to assess the robustness of the proposed residual-based approach: 1) When T and N measures were averaged across left and right hemispheres, the T-N mismatch metric had a correlation of r=0.99 with that using unilateral measures, and clustering produced groups with similar characteristics, with the vast majority of subjects continuing to be clustered together. This is pictorially shown in a tanglegram display in Figure 5(top), 2) When non-discretized residuals were directly fed into hierarchical clustering, again qualitatively similar groups were obtained, as shown in the tanglegram display in Figure 5(bottom), 3) When we weighted the contribution of each ROI (# of suprathreshold residuals) with ROI size, the resulting T-N mismatch metric was still highly correlated (r=0.97) with those without such weighting.

## Discussion:

Explicitly modeling the variability in the relationship between tau burden and neurodegeneration can be a useful tool for understanding phenotype and may provide important clues to additional factors that modulate neurodegeneration and cognitive decline. Determining such factors, e.g., the presence of co-pathology, is essential in both clinical research and practice to achieve a precision-based approach to diagnosis, prognosis, and management. While the dichotomous designation of biomarkers in the A/T/(N) framework has proven to be a powerful concept for AD classification, it fails to utilize the continuous and spatially varying information that in-vivo multimodal imaging can provide. Here we demonstrate that a measure of deviation from a normative relationship between tau burden and neurodegeneration across brain regions in individuals on the AD continuum captures variability due to multiple underlying factors.

The basic conceptual principle motivating the approach described here is that NFT pathology is tightly linked to AD-related neurodegeneration, as has been demonstrated in numerous in vivo and ex vivo studies[26–30]. In contrast, measures of beta-amyloid plaque pathology do not strongly correlate with neurodegeneration[31]. Thus, the load of NFT pathology in the setting of AD (i.e. presence of amyloid and tau pathology as defined by the A/T/(N) framework) should provide an estimation of the degree to which brain structure would be expected to be affected specifically by AD. Deviation from this expectation would suggest that additional factors such as co-pathology or brain reserve are modulating this relationship. While this assumption may only approximate reality and more complex relationships may be present, the basic notion is likely to be correct and is consistent with the findings of this study.

### T-N mismatch metric has meaningful associations

As presented in Tables 1 and 2, a simple scalar measure summarizing the degree of mismatch between the amount of tau pathology and neurodegeneration across the whole cerebrum is a useful metric to study underlying factors that may contribute to the heterogeneity of AD. This scalar measure used spatial information to account for the extent of this mismatch across the brain (i.e. the number of ROIs), but is insensitive to the spatial pattern. Nonetheless, we found that the T-N mismatch metric was significantly correlated with two factors that are likely to be drivers of atrophy beyond that of AD-related pathology; age[32] and cerebrovascular disease[33,34], as measured by WMH burden[35]. Moreover, the extent of this mismatch was associated with the degree of cognitive impairment across several domains, even after controlling for age and $^{18}$F-Flortaucipir uptake in the inferior temporal cortex, a surrogate for NFT burden. As one would predict, more atrophy than expected for a given level of tau was associated with poorer cognition while less atrophy was associated with better performance in the face of the same tau burden. The fact that the T-N mismatch metric did not correlate with $^{18}$F-Flortaucipir uptake in key AD-related regions supports the notion that our measure of mismatch is not confounded by disease severity in a non-linear relationship with structural neurodegeneration.

The current result supports a prior study by Ossenkoppele et al.[36] which used a global measure of T and N, whole cortical $^{18}$F-Flortaucipir uptake and mean cortical thickness,

respectively, and a global cognitive measure, MMSE, to calculate a brain and cognitive reserve metric. Consistent with our findings, Ossenkoppele et al. reported that brain reserve was associated with age, such that younger individuals had more brain reserve, and that cognitive reserve was related to cortical thickness. Additionally, they found that female sex was associated with higher brain reserve. We also found a trend towards sex related to our continuous measure of mismatch. Thus, the findings here provide convergence in an independent dataset, but also demonstrate that these relationships occur independent of overall tau burden. *Further, the sensitivity analyses using non-discretized residuals and bilaterally averaged ROI measures points to the robustness of the general approach of using T-N residuals*.

### T-N mismatch metric can potentially reveal phenotypes

The scalar T-N mismatch metric captures the extent of, but not the spatial pattern of mismatch. This pattern may provide additional insights, particularly related to the presence of co-pathology in cases with more neurodegeneration than expected for a level of tau pathology. To explore this, we used a data-driven approach to cluster individuals based on the spatial pattern of T-N mismatch. The clustering yielded six groups with the largest, not surprisingly, demonstrating little T-N mismatch that we referred to as a canonical group. Three groups (2,4, and 6) displayed greater neurodegeneration than expected for the level of tau, but differed in spatial patterns. One group in particular (Group 2) had excessive atrophy in temporal-limbic structures, particularly anterior hippocampus and temporal pole. While requiring pathologic confirmation in future work, this group had a pattern of atrophy suggestive of concomitant TDP-43 pathology consistent with descriptions of limbic-predominant age-associated TDP-43 encephalopathy[10,11], or TDP-43 Type C pathology which is associated with semantic dementia[37]. The two other groups with higher levels of neurodegeneration had differing degrees of severity, but were generally diffuse with one group associated with high levels of WMHs. *Post-hoc analysis of longitudinal changes in MMSE also found a significantly steeper rate of decline in these groups compared to the canonical group, further supporting the notion that they represent vulnerable phenotypes*.

While we expected evidence of resilience as well, we found two different groups based on pattern of regions that had greater cortical thickness than expected. These groups may be explained by having higher brain reserve, or relative resistance to the presence of NFT pathology, perhaps through less neuropil or neuronal dropout[38]. The separation of these two groups may reflect different mechanisms of resilience that if studied on a larger scale could provide important insights into environmental or genetic factors that could contribute to reducing risk or slowing disease progression.

This data-driven approach needs to be validated using post-mortem pathological analysis, to determine if the imaging phenotypes found indeed map on to disease phenotypes. Nonetheless, it has implications for clinical interventions and prognosis, particularly as a surrogate for neurodegenerative pathologies for which no in-vivo biomarkers exist[39]. On the other hand, this approach can also help researchers identify and study resilience or brain reserve as a phenotype. If validated with longitudinal cognitive data, this also may be a valuable prognostic tool for clinicians.

Heterogeneity of underlying phenotypes is a significant factor that reduces power in clinical trials[40]. As an increasing number of new therapies that specifically target pathological tau species are tested[41], this approach may help exclude trial participants that might have significant non-tau co-morbidity (such as vascular disease or other proteinopathy) that might be driving their neurodegeneration. This could potentially reduce cost and provide a clearer understanding of the efficacy of a potential therapy.

### Limitations and future work

The approach presented here for utilizing the richer spatial, continuously varying, information about T and N biomarkers to study variability is necessarily limited by the composition of the study cohort. The current composition of the ADNI research cohort has relatively strict inclusion criteria that limits the number of underlying factors that can be examined, including a relatively restricted range of disease severity. Future work will apply this framework to more heterogeneous cohorts that encompass a more diverse phenotypic landscape, including earlier onset disease[42], non-amnestic presentations, and greater inclusion of groups with comorbidities, such as vascular risk factors such as diabetes.

Alternative approaches to defining a T-N mismatch metric could also be considered. The use of ROI-based regression residuals, as we have done, is easy to operationalize and implement. One could use a similar approach using voxel-level regressions instead. Another way to estimate a mismatch would be to first build a model of direct image-to-image (tau PET to MRI) translation using deep neural networks[43,44] instead of using biomarker measurements from both modalities as we have done here, and then use a metric of deviation from the predicted model output as a mismatch measure.

Any data-driven clustering approach is exploratory in nature, and the results will likely vary depending on cohort composition and the underlying disease phenotypes. Nonetheless, it is encouraging that the clusters that emerged based on T-N residuals appear to differ from each other based on underlying *factors other than the pair of imaging measurements used to derive them*. These factors included those likely related to vulnerability and resilience, e.g. a marker of vascular disease (white matter hyperintensity) and age respectively.

Finally, the relatively small cohort used here limits strong conclusions about the stability of these groups, but does provide a proof-of-principle that accounting for the relationship of T and N provides phenotypic groupings sensitive to biological modifiers of these relationships. In addition to more heterogeneous populations, future studies will need to apply this approach to larger cohorts and study longitudinal outcomes to determine the clinical significance of mismatch. Ultimately, the study of cross-modality relationships of different biological constructs, including A, T, and N, may allow for quantitative assessments linked to disease heterogeneity[45].

## Acknowledgements:

Data collection and sharing for this project was funded by the Alzheimer's Disease Neur oimaging 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 Biomed ical Imaging and Bioengineering, and through generous contributions from the following : AbbVie,Alzheimer's Association; Alzheimer's Drug DiscoveryFoundation; Araclon Biote ch; 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; Me rck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack T echnologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; S ervier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadia n 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 Un iversity of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

## References:

1. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimer's Dement. 2018;14(4):535–62. [PubMed: 29653606]
2. La Joie R, Visani A, Bourakova V, et al. AV1451-PET cortical uptake and regional distribution predict longitudinal atrophy in Alzheimer's disease. Alzheimer's Dement 2017;13.
3. Schonhaut D, Ossenkoppele R, Schöll M, et al. Associations of [18F]AV1451 Tau PET with age, ApoE genotype, and cognition in Alzheimer's disease. Alzheimer's Dement 2015;
4. Knopman DS, Lundt ES, Therneau TM, et al. Entorhinal cortex tau, amyloid-β, cortical thickness and memory performance in non-demented subjects. Brain 2019;142(4):1148–60. [PubMed: 30759182]
5. Schöll M, Lockhart SN, Schonhaut DR, et al. PET Imaging of Tau Deposition in the Aging Human Brain. Neuron 2016;89(5):971–82. [PubMed: 26938442]
6. Das SR, Xie L, Wisse LEM, et al. Longitudinal and cross-sectional structural magnetic resonance imaging correlates of AV-1451 uptake. Neurobiol Aging 2018;66:49–58. [PubMed: 29518752]
7. Mormino EC, Papp KV, Rentz DM, et al. Heterogeneity in Suspected Non-Alzheimer Disease Pathophysiology Among Clinically Normal Older Individuals. JAMA Neurol 2016;73(10):1185–91. [PubMed: 27548655]
8. Wisse LEM, Butala N, Das SR, et al. Suspected non-AD pathology in mild cognitive impairment. Neurobiol Aging 2015;36(12).
9. Ossenkoppele R, Cohn-Sheehy BI, La Joie R, et al. Atrophy patterns in early clinical stages across distinct phenotypes of Alzheimer's disease. Hum Brain Mapp 2015;36(11):4421–37. [PubMed: 26260856]
10. Larner AJ, Griffiths TD. Limbic-predominant age-related TDP-43 encephalopathy (LATE) [Internet]. Brain. 2019;142(8):E42. [PubMed: 31243434]
11. de Flores R, Wisse LEM, Das SR, et al. Contribution of mixed pathology to medial temporal lobe atrophy in Alzheimer's disease. Alzheimer's Dement 2020;16(6):843–52. [PubMed: 32323446]
12. Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, et al. Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimer's Dement 2018;1–7.
13. Robinson JL, Richardson H, Xie SX, et al. The development and convergence of co-pathologies in Alzheimer's disease. Brain 2020;
14. Young PNE, Estarellas M, Coomans E, et al. Imaging biomarkers in neurodegeneration: Current and future practices [Internet]. Alzheimer's Res. Ther 2020;12(1):49. [PubMed: 32340618]
15. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 2011;54(3):2033–44. [PubMed: 20851191]
16. Wang H, Suh JW, Das S, Pluta J, Altinay M, Yushkevich P. Regression-Based Label Fusion for Multi-Atlas Segmentation. Conf Comput Vis Pattern Recognit Work IEEE Comput Soc Conf Comput Vis Pattern Recognit Work 2011;20(1):1113–20.
17. Landman BA, Warfield S. MICCAI 2012: Grand challenge and workshop on multi-atlas labeling. In: Proc. International Conference on Medical Image Computing and Computer Assisted Intervention, MICCAI. 2012.

18. Landau SM, Breault C, Joshi AD, et al. Amyloid-β imaging with Pittsburgh compound B and florbetapir: comparing radiotracers and quantification methods. J Nucl Med 2013;54(1):70–7. [PubMed: 23166389]
19. Tustison NJ, Cook PA, Klein A, et al. Large-scale evaluation of ANTs and FreeSurfer cortical thickness measurements. Neuroimage 2014;99.
20. Das SR, Avants BB, Grossman M, Gee JC. Registration based cortical thickness measurement. Neuroimage 2009;45(3).
21. Kuijf HJ, Casamitjana A, Collins DL, et al. Standardized Assessment of Automatic Segmentation of White Matter Hyperintensities and Results of the WMH Segmentation Challenge. IEEE Trans Med Imaging 2019;38(11):2556–68. [PubMed: 30908194]
22. Huber PJ, Wiley J, New York Chichester Brisbane Toronto Singapore S. Robust Statistics [Internet]. 1981.
23. Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc 1963;58(301):236–44.
24. Thorndike RL. Who belongs in the family? Psychometrika 1953;18(4):267–76.
25. Botha H, Mantyh WG, Murray ME, et al. FDG-PET in tau-negative amnestic dementia resembles that of autopsy-proven hippocampal sclerosis. Brain 2018;141(4):1201–17. [PubMed: 29538658]
26. 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(524):5732.
27. Honer WG, Dickson DW, Gleeson J, Davies P. Regional synaptic pathology in Alzheimer's disease. Neurobiol Aging 1992;13(3):375–82. [PubMed: 1625766]
28. Bobinski M, Wegiel J, Wisniewski HM, et al. Neurofibrillary pathology — correlation with hippocampal formation atrophy in Alzheimer disease. Neurobiol Aging 1996;17(6):909–19. [PubMed: 9363803]
29. Gómez-Isla T, Hollister R, West H, et al. Neuronal loss correlates with but exceeds neurofibrillary tangles in Alzheimer's disease. Ann Neurol 1997;41(1):17–24. [PubMed: 9005861]
30. Nagy Z, Jobst KA, Esiri MM, et al. Hippocampal pathology reflects memory deficit and brain imaging measurements in alzheimers disease: Clinicopathologic correlations using three sets of pathologic diagnostic criteria. Dement Geriatr Cogn Disord 1996;7(2):76–81.
31. Iaccarino L, Tammewar G, Ayakta N, et al. Local and distant relationships between amyloid, tau and neurodegeneration in Alzheimer's Disease. NeuroImage Clin 2018;
32. Hou Y, Dan X, Babbar M, et al. Ageing as a risk factor for neurodegenerative disease [Internet]. Nat. Rev. Neurol 2019;15(10):565–81. [PubMed: 31501588]
33. Toledo JB, Arnold SE, Raible K, et al. Contribution of cerebrovascular disease in autopsy confirmed neurodegenerative disease cases in the National Alzheimer's Coordinating Centre. Brain 2013;136(9):2697–706. [PubMed: 23842566]
34. Lao PJ, Brickman AM. Multimodal neuroimaging study of cerebrovascular disease, amyloid deposition, and neurodegeneration in Alzheimer's disease progression. Alzheimer's Dement Diagnosis, Assess Dis Monit 2018;10:638–46.
35. Ferreira D, Shams S, Cavallin L, et al. The contribution of small vessel disease to subtypes of Alzheimer's disease: a study on cerebrospinal fluid and imaging biomarkers. Neurobiol Aging 2018;70:18–29. [PubMed: 29935417]
36. Ossenkoppele R, Lyoo CH, Jester-Broms J, et al. Assessment of Demographic, Genetic, and Imaging Variables Associated with Brain Resilience and Cognitive Resilience to Pathological Tau in Patients with Alzheimer Disease. JAMA Neurol 2020;77(5):632–42. [PubMed: 32091549]
37. Grossman M Primary progressive aphasia: Clinicopathological correlations [Internet]. Nat. Rev. Neurol 2010;6(2):88–97. [PubMed: 20139998]
38. Arenaza-Urquijo EM, Vemuri P. Resistance vs resilience to Alzheimer disease [Internet]. Neurology. 2018;90(15):695–703. [PubMed: 29592885]
39. Sancesario GM, Bernardini S. Diagnosis of neurodegenerative dementia: where do we stand, now? Ann Transl Med 2018;6(17):340–340. [PubMed: 30306079]

40. Devi G, Scheltens P. Heterogeneity of Alzheimer's disease: Consequence for drug trials? [Internet]. Alzheimer's Res. Ther 2018;10(1).

41. VandeVrede L, Boxer AL, Polydoro M. Targeting tau: Clinical trials and novel therapeutic approaches. Neurosci Lett 2020;731(July 2019):134919. [PubMed: 32380145]

42. Mendez MF. Early-onset Alzheimer disease and its variants. Contin Lifelong Learn Neurol 2019;25(1):34–51.

43. Sikka A, Peri SV, Bathula DR. MRI to FDG-PET: Cross-modal synthesis using 3d u-net for multi-modal alzheimer's classification. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Verlag; 2018. p. 80–9.

44. Choi H, Lee DS. Generation of structural MR images from amyloid PET: Application to MR-less quantification. J Nucl Med 2018;59(7):1111–7. [PubMed: 29217736]

45. Young AL, Marinescu RV., Oxtoby NP, et al. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. Nat Commun 2018;9(1):1–16. [PubMed: 29317637]

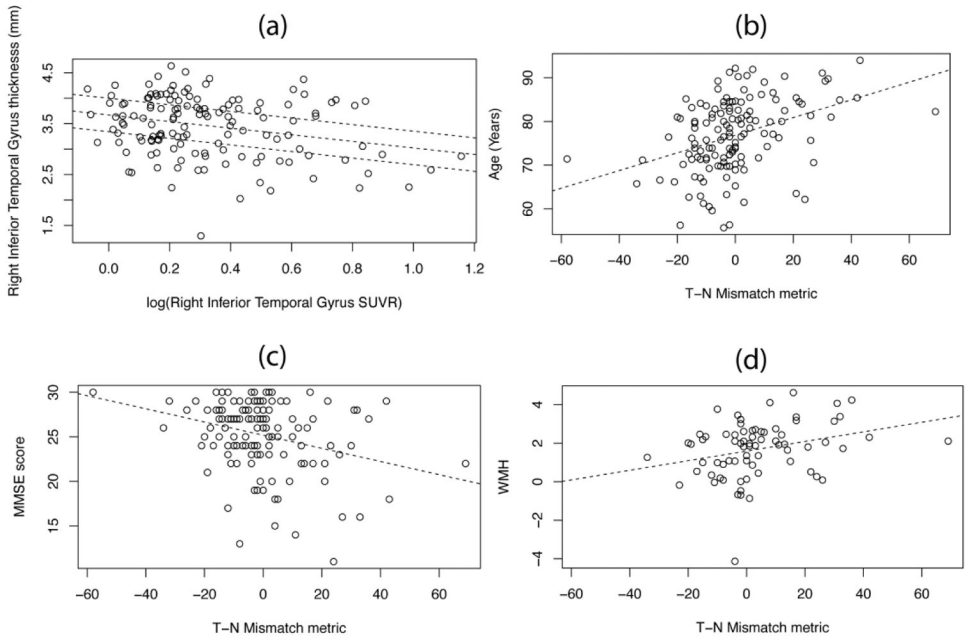


**Figure 1.**

Scatter plots showing (a) Tau tracer SUVR vs. cortical thickness in the right inferior temporal gyrus ROI along with robust regression fit and ± 1.5 standard deviation lines, (b) T-N mismatch metric vs. age, (c) T-N mismatch metric vs. MMSE scores, and (d) T-N mismatch metric vs. WMH volume.

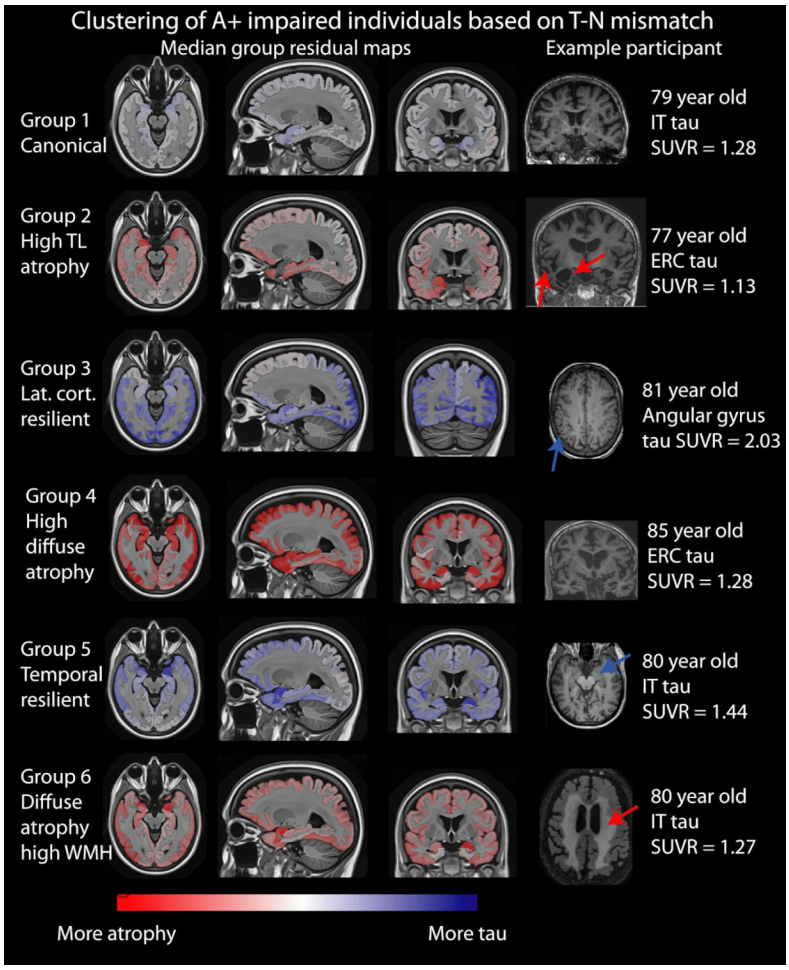


**Figure 2:**

Median ROI-wise group residual maps (columns 1–3). More saturated color represents higher residual, indicating higher (red) or lower (blue) atrophy relative to tau burden. Column 4 shows representative examples with arrows indicating features characteristic of the group.

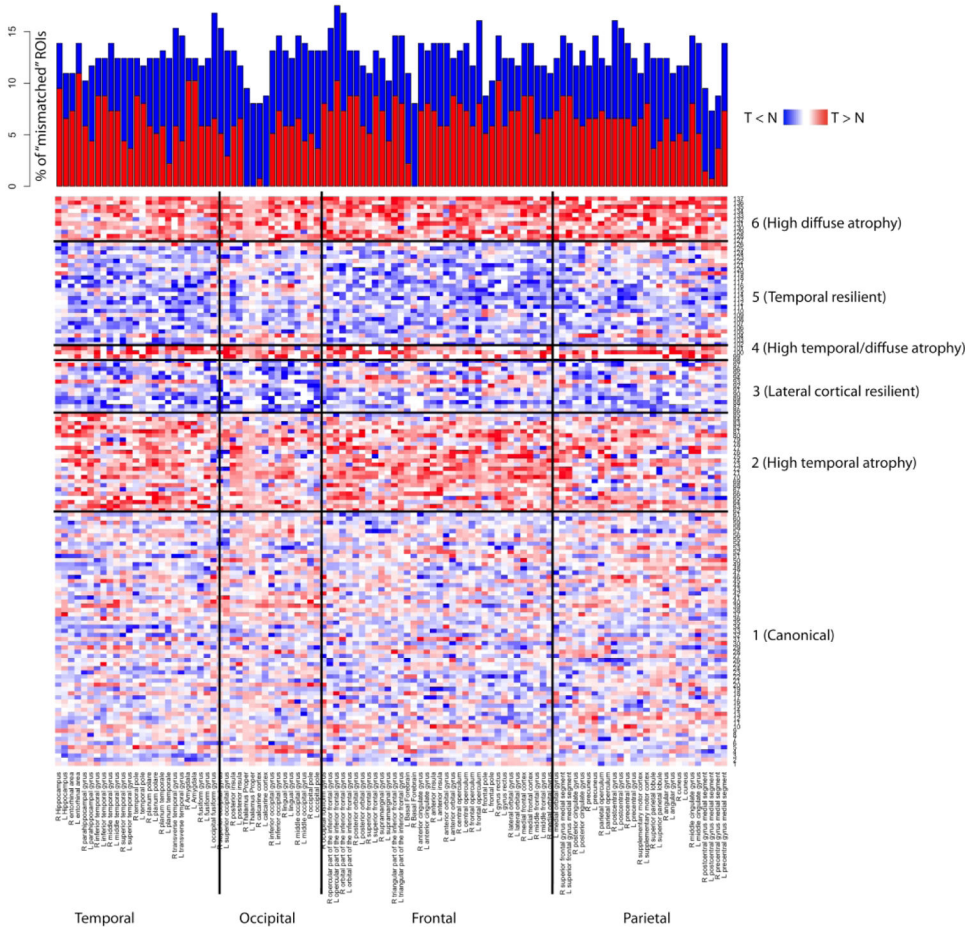


**Figure 3.**

T-N regression residuals for each ROI and each participant visualized on a heatmap. ROIs on the x-axis are sorted by lobes, participants on the y-axis are sorted by clusters. Shades of red show positive residuals (more atrophy) and shades of blue show negative residuals (more tau). Stacked bar graph in the top panel shows percentage of ROIs showing “mismatch”, as defined by the residual being greater than 1.5 standard deviation away from the regression line.

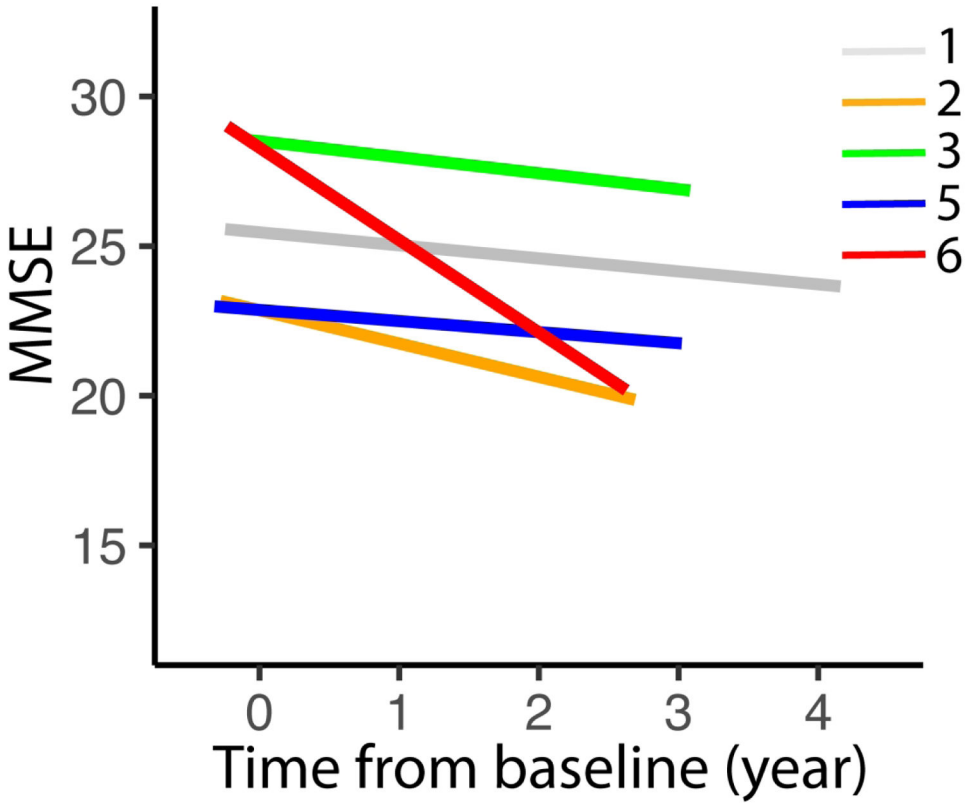


**Figure 4.**

Trajectory of cognitive decline for different groups of participants. Slopes shown for each group is estimated by linear mixed-effects model. Groups 2 and 6 are the vulnerable groups showing steeper decline than group 1 (canonical). Groups 3 and 5 are the resilient groups. Group 4 is not shown because of the small group size. Note that the variable maximum time periods for different groups reflect differences in duration of followup data available.

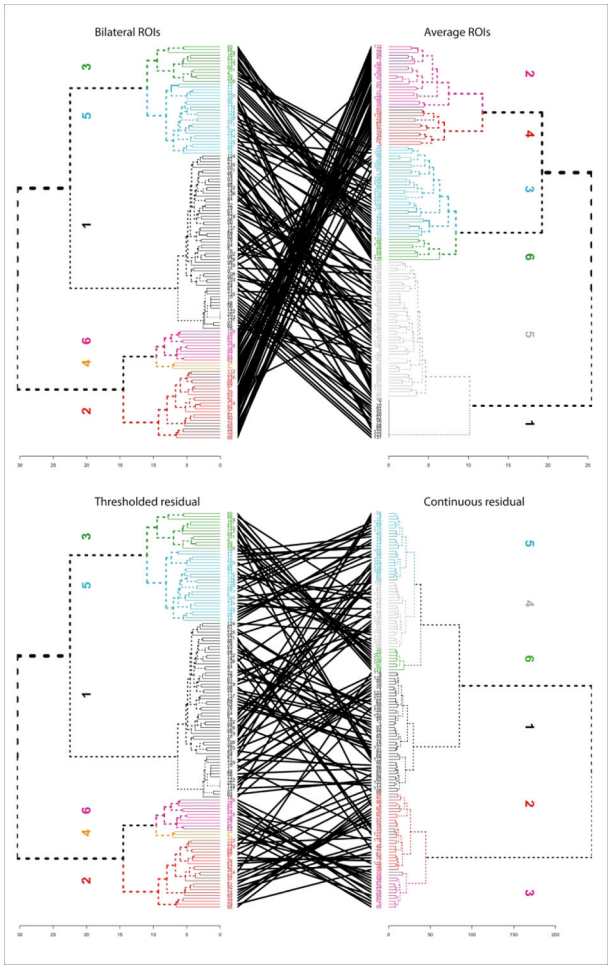


**Figure 5.**

Tanglegram displays comparing cluster memberships. Top panel: comparison memberships between using left and right hemispheric ROI measures separately (left) vs. using averaged measures (right). Bottom panel: comparison using discretized regression residuals (left) vs. the raw continuous regression residuals (right). In both panels, color for each cluster on the right was chosen to be similar to the cluster on the left that had the most overlap in membership for ease of visualization. Cluster labels (1–6) are arbitrary, as output by the clustering algorithm.

**Table 1:**

Summary of demographic and imaging variables within tertiles of T-N mismatch metric and its correlation

| Variable | Lower tertile (resilient) | Middle tertile | Upper tertile (vulnerable) | Pearson r/chi-square | p-value |
|---|---|---|---|---|---|
| Age (Years) | 73.3 | 76.4 | 80.9 | 0.38 | $< 10^{-5}$ |
| Gender (F/M) | 26/20 | 24/22 | 15/30 | 5.53 | 0.06 |
| ApoE4 status (0/1/2 allele) | 12/14/12 | 10/18/8 | 16/18/5 | 4.88 | 0.29 |
| MMSE | 26.2 | 25.9 | 23.5 | −0.30 | 0.0005 |
| WMH volume (log $mm^3$) | 1.30 | 1.21 | 2.25 | 0.35 | < 0.0001 |
| ERC tau ($^{18}$F-Flortaucipir SUVR) | 1.35 | 1.34 | 1.31 | 0.03 | 0.69 |
| IT tau ($^{18}$F-Flortaucipir SUVR) | 1.43 | 1.45 | 1.46 | 0.003 | 0.96 |
| Precuneus tau ($^{18}$F-Flortaucipir SUVR) | 1.25 | 1.26 | 1.32 | 0.04 | 0.63 |
| Angular gyrus tau ($^{18}$F-Flortaucipir SUVR) | 1.38 | 1.34 | 1.39 | 0.01 | 0.88 |

Mean (age, MMSE, WMH volume and regional tau burden) or categorical counts (gender and ApoE4 status) for participants trichotomized into tertiles of T-N mismatch metric. Upper tertile has the highest relative atrophy than expected for tau (vulnerable). Rightmost columns show Pearson correlation of age, MMSE and the imaging measures with the T-N mismatch metric is also shown. For the two categorical variables gender and ApoE4 status, the two result of chi-square test between the three tertiles are shown instead.

**Table 2:**

Correlation of cognitive scores with T-N mismatch metric with age, gender, years of education and IT tau uptake as covariates

| Cognitive test | Pearson r | p-value |
|---|---|---|
| MMSE | −0.30 | 0.0005 |
| Logical Memory Test (Delayed Recall) | −0.15 | 0.09 |
| AVLT 5-min Delayed Recall | −0.24 | 0.006 |
| Category fluency | −0.23 | 0.01 |
| Trail making Test A | 0.19 | 0.03 |
| Trail making Test B | 0.17 | 0.06 |

## Table 3

Characteristics of groups defined by data-driven clustering of residuals from T-N. Bottom row indicates variables that differ significantly between groups.

| Group (N) | Description | Diagnosis MCI/ Dementia | Age | Gender (F/M) | Log White Matter Hyper-intensity | MMSE | IT Tau SUVR | T-N Mismatch Metric (median) |
|---|---|---|---|---|---|---|---|---|
| Group 1 (61) | Canonical (low residuals) | 38/23 | 76.8±8.0 | 32/29 | 8.38 | 25.8±3.2 | 1.43±0.4 | −2 |
| Group 2 (24) | High temporal/ limbic atrophy (pattern suggestive of TDP-43) | 9/15 | 79.9±7.5 | 10/14 | 9.47* | 24.2±3.7 | 1.46±0.4 | 10 |
| Group 3 (13) | Resilient (less atrophy relative to tau in lateral cortical) | 10/3 | 69.7±8.8* | 9/4 | 8.02 | 27.0±3.4 | 1.46±0.4 | −14 |
| Group 4 (3) | High temporal/ limbic and diffuse atrophy (oldest group; ?TDP + vascular + age-related change) | 1/2 | 87.2±6.1 | 0/3 | 7.88* | 23.0±5.6 | 1.46±0.2 | 43 |
| Group 5 (25) | Resilient Temporal | 18/7 | 75.1±6.8 | 11/14 | 7.70 | 25.1±4.4 | 1.46±0.3 | −14 |
| Group 6 (11) | High diffuse atrophy (lowest cognition, high WMH) | 4/7 | 79.3±10.2 | 3/8 | 9.73* | 22.6±6.0 | 1.45±0.5 | 27 |
| Group Difference | | $p < 0.05$ | $p< 0.01$ | $p=0.15$ | $p = 0.06$ | $p<0.05$ | $p> 0.05$ | - |

* indicates significant group difference compared with Group 1 (two-sided t-test).