Prediction of clinical and biomarker conformed Alzheimer's disease and mild cognitive impairment from multi-feature brain structural MRI

Binyin Li (libinyin@126.com)
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital https://orcid.org/0000-0003-1953-382X

Miao Zhang
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Joost Riphagen
Athinoula A Martinos Center for Biomedical Imaging

Kathryn Morrison Yochim
Athinoula A Martinos Center for Biomedical Imaging

Biao Li
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

Jun Liu
Shanghai Jiao Tong University Medical School Affiliated Ruijin Hospital

David H. Salat
Massachusetts General Hospital

Research

Keywords: neuroimage, classification, multi-feature MRI, age detrending

Posted Date: April 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-23368/v1

Prediction of clinical and biomarker conformed Alzheimer's disease and mild cognitive impairment from multi-feature brain structural MRI

Binyin Li[1,2,4], Miao Zhang[3], Joost Riphagen[1], Kathryn Morrison Yochim[1], Biao Li[3], Jun Liu[2], David H. Salat[1,4,5], for the Alzheimer's Disease Neuroimaging Initiative*

Affiliations:
1. MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts
2. Department of Neurology, Ruijin Hospital affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China
3. Department of Nuclear Medicine, Ruijin Hospital affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China
4. Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts
5. Neuroimaging Research for Veterans Center, VA Boston Healthcare System, Boston, Massachusetts
6. Wellcome Trust Centre for Integrative Neuroimaging, University of Oxford, Oxford, UK

Corresponding author:
Binyin Li;
1. Department of Neurology, Ruijin Hospital affiliated with Shanghai Jiao Tong University School of Medicine, Shanghai, 200025, China
2. MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, Massachusetts
3. Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts

Email: libinyin@126.com

*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/wpcontent/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Background: Structural neuroimaging has been applied towards identification of individuals with Alzheimer's disease (AD) and mild cognitive impairment (MCI). However, these methods are greatly impacted by age limiting their utility for detection of preclinical pathology. Therefore, careful consideration of age effects in the modeling of AD degenerative patterns could provide more sensitive detection of the earliest stages of brain disease.

Methods: We built linear models for age based on multiple combined structural features (cortical thickness, subcortical structural volumes, ratio of gray to white matter signal intensity, white matter signal abnormalities, total intracranial volume) in 272 healthy adults across a wide age range (D1: age 36-108). These models were then used to create a new support vector machine (SVM) training model with 10-fold cross validation in 136 AD and 268 control participants (D2) based on deviations from the expected age-effects found in the initial sample. Subsequent validation assessed the accuracy of the SVM model to correctly classify AD patients in a new dataset (D3). Finally, we applied the classifier to individuals with MCI to evaluate prediction for early impairment and longitudinal cognitive change.

Results: Optimal cross-validation accuracy was 93.07% in the D2, compared to 91.83% without age detrending in D1. In the validation dataset (D3), the classifier obtained an accuracy of 84.85% (56/66), sensitivity of 85.36% (35/41) and specificity of 84% (21/25). In the MCI dataset, we observed significantly greater longitudinal cognitive decline in MCI who were classified as more 'AD-like' (MCI-AD), and this effect was pronounced in individuals who were late MCI. The top five contributive features were volumes of left hippocampus, right hippocampus, left amygdala, the thickness of left and right medial temporal & parahippocampus gyrus.

Conclusions: Linear detrending for age in SVM for combined structural features resulted in good performance for classification of AD and generalization of MCI prediction. Such procedures should be employed in future work.

Keywords: neuroimage, classification, multi-feature MRI, age detrending

Background

Alzheimer's disease (AD) is the most common cause of dementia in older adults and is characterized by abnormal pathologic proteins (e.g. amyloid and tau) that are presumed to promote neural injury. Neural injury is regarded as one of the hallmarks of AD and this can be assessed in patients *in vivo* as regional atrophy on structural magnetic resonance imaging (MRI). Compared to the more direct biomarkers of amyloid and tau that can be visualized by positron emission tomography (PET) or quantified in cerebrospinal fluid, structural MRI is less costly, less invasive and available in most clinical settings. Thus, utilization of structural MRI in the accurate and sensitive detection of AD is an important goal.

Various structural features obtained from simple T1 weighted MRI are impacted by typical aging as well as the degenerative processes of AD (1-15) and such features have been used in statistical classification of individuals with a clinical diagnosis of AD (1, 16-18). For example, patterns of atrophy measured by computational models of cerebral cortical thickness can provide spatially distinct features to be utilized in classification models (1, 19-21). Similarly, subcortical volumetric atrophy can provide metrics that are sensitive to neurodegenerative processes in AD as well as the impact of typical aging, with the hippocampus, amygdala, and thalamus being particularly affected in AD (22-24). In addition to morphometric properties, additional 'microstructural' features can be obtained from standard structural MRI through quantification of signal properties such as the gray to white matter signal intensity ratio (GWR) at each point along the cortical surface. The GWR is significantly regionally increased with age particularly throughout frontal regions demonstrating a reduction in the gray-white matter contrast (12) and is also impacted in regions known to be affected by AD pathology (6). The GWR can be used in the detection of mild cognitive impairment (MCI) progression (15), and this measure has been demonstrated to strengthen results for cortical thickness when included in statistical models (25). Finally, white matter signal abnormalities (WMSAs; aka white matter hyperintensities) are presumed to be due to small vessel disease (26), more highly prevalent in individuals with AD (2, 3, 5), and known to impact clinical trajectories as well as potentially modulate clinical status for a given level of primary AD pathology (27). In summary, these results demonstrate the range of structural features that can be obtained from standard T1 imaging and used in the classification of AD. However, given the overlapping nature of aging and AD related effects on these features, careful modeling is required to assure that classifications are not biased due to this strong contributing factor.

Machine learning algorithms provide powerful methodology for detection of individuals with AD from structural MRI given favorable properties for handling high-dimensional data and efficient feature selection (28, 29). Support vector machine (SVM) is currently the most widely used procedure in neuroimaging studies to classify AD (1, 30-32). It is possible that such procedures using simple structural imaging can provide an initial screening to determine whether follow up basement with more direct, yet costly and invasive biomarker procedures are necessary. Although relatively successful, these procedures have mostly been applied in clean research

samples and have some limitations in generalization. First, given the impact of age on structural imaging measures, it is possible that these effects bias prior results. This could be particularly consequential in samples expanding past the typical research sample age range (e.g. for identifying early onset AD or AD in the oldest old). Although prior studies have noted the need to consider age as a confounding factor contributing to misclassification of individuals with AD (1, 33, 34), such information was not incorporated directly into the mode. For example, we previously found that SVM using structural imaging features in AD vulnerable regions resulted in older control individuals being more likely to be classified as AD (1). Second, biomarker analysis in our previous work suggested that patients may have a clinical diagnosis that is inconsistent with their pathology status (e.g. clinical diagnosis of control but having high levels of AD biomarkers (1)). Thus, novel methods for classification performance must be assessed in addition to the clinical diagnosis including biomarker confirmation and assessment of longitudinal cognitive trajectories. According to the ATN criterion for the AD research frame (35), prior work has rarely used an AD specific biomarker in diagnosis during the validation phase of classification.

In this work, we implement a novel the linear detrending of structural MRI features for age based on age trends in cognitively unimpaired individuals with a wide range of age prior to statistical evaluation using SVM classification. Linear models were chosen based on studies for age correction (33, 34) in healthy participants, and we optimized coefficient and offset for each age-dependent structural feature. We then applied the linear models to a novel dataset to classify individuals with AD. Deviations of each structural feature from age-regression line were trained and tested by SVM, as well as ranked by their contribution to the SVM classifier. We validated the classifier on a new dataset with AD and controls, who were diagnosed by both β-amyloid and cognitive assessments. Finally, we applied the classifier to MCI patients to evaluate its prediction for MCI longitudinal cognitive decline.

Methods

Dataset

Dataset one for linear detrending models (D1): We used structural MRI data of 272 healthy adults from the Human Connectome Project Lifespan/ Aging cohort (HCP-A) (146 females, age: 36 - 108). T1 weighted multi-echo MPRAGE with prospective navigator motion correction MRI parameters included TE=1.8/3.6/5.4/7.2 ms (multi-echo), TR = 2500 ms, field of view = 256 ×256 mm^2, number of slices = 208, voxel size = 0.8 × 0.8 × 0.8 mm^3, and flip angle = 8° (36, 37).

Dataset two for training and testing classifier (D2): 3.0T T1-weighted images from AD and control participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (adni.loni.usc.edu) were used for training classifier. A total of 404 participants (268 controls and 136 AD) were considered in this study similar to description in our prior work (1). 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 MCI and early AD.

Dataset three for validation (D3): We evaluated 66 participants using a whole-body PET/MR scanner (Biograph mMR; Siemens Healthcare, Erlangen, Germany) with an 8-channel head/neck coil in Shanghai Jiao Tong University affiliated Ruijin Hospital, China. PET scans used [18F] florbetapir (AV45) tracer for imaging Aβ. Individuals were considered Aβ positive if global AV45 standardized uptake value ratio > 1.228 (38). The simultaneous MR models included sagittal three-dimensional (3D) T1-weighted magnetization-prepared rapid acquisition gradient-echo sequence (T1 MPRAGE): TR = 1900 ms, TE = 2.44 ms, angle = 9°, 192 slices (gap, 0.5 mm) covering the whole brain, FOV 256 × 256 mm, matrix 256 × 256, voxel size = 1.0 × 1.0 × 1.0 mm^3. All patients performed Mini-Mental State Examination (MMSE, Chinese Version) (39), Beijing version of Montreal cognitive assessment (MoCA), the Chinese version of Addenbrooke's cognitive examination-revised (ACER) (40), and global clinical dementia rating. Forty-one participants were diagnosed as AD according to the National Institute on Aging-Alzheimer's Association (NIA-AA) workgroups (41). This part was approved by the Ethics Committee, Shanghai Jiao Tong University affiliated Ruijin Hospital, China.

MCI dataset: We included 180 early MCI (EMCI) and 96 late MCI (LMCI) patients defined by the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset (adni.loni.usc.edu) with complete baseline and follow-up neuropsychological assessments. The timelines had 6 visits: baseline, month 6, month 12, month 24, month 36 and month 48. Patients were evaluated by MMSE, MoCA, Alzheimer's Disease Assessment Scale Cognitive Subscale (ADAS-13) (42), Rey's auditory verbal learning test (RAVLT) learning and forgetting score, and trail making test-B (TMT-B) in each visit. We applied validated classifier from the above datasets to MCI and evaluated its predictive performance at baseline. Demographics for the entire study are presented in Table 1 and Table 2. We additionally examined whether participants classified as AD-like (MCI-AD) exhibited greater longitudinal cognitive decline compared to MCI classified as more like control brains.

Structural features

The preprocessing pipeline of structural images was performed using the FreeSurfer image analysis suite (https://surfer.nmr.mgh.harvard.edu/)(43-45). As described in prior our publications (1, 12), the processing included removal of non-brain tissue, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures, intensity normalization, tessellation of the gray matter white matter boundary, automated topology correction, and surface deformation following intensity gradients to optimally place the gray/white and gray/CSF borders at the location where the greatest shift in intensity defines the transition to the other tissue class.

Cortical thickness: Once the cortical models were complete, a number of deformable procedures can be performed for further data processing and analysis including surface inflation (44), registration to a spherical atlas which utilizes individual cortical folding patterns to match cortical

geometry across participants (46), parcellation of the cerebral cortex into units based on gyral and sulcal structure (47, 48) , and creation of a variety of surface-based data including maps of curvature and sulcal depth. Procedures for the measurement of cortical thickness have been validated against histological analysis (49) and manual measurements (14, 50).

Subcortical segmentation volumes: Automatic subcortical segmentation of a brain volume was based upon the existence of an atlas containing probabilistic information on the location of structures (51). Estimated total intracranial volume (eTIV) and volumes of 16 regions of interest were extracted, including cerebral white matter, cerebral cortex, thalamus, caudate, putamen, pallidum, hippocampus and amygdala in each hemisphere. The WMSAs were labeled using a probabilistic procedure subsequently extended to label white matter lesions. Total WMSA (hypointensity) volume was then calculated for each hemisphere; these were averaged together to create a single WMSA volume for each participant. This procedure has demonstrated sensitivity in measuring white matter damage in individuals with AD (52) and has been utilized in several prior studies of brain aging and AD(53-56).

Gray-white matter ratio (GWR): The gray matter values as a ratio to bordering white matter values provided a unit that was normalized for the local imaging environment. As described previously (6, 12), tissue intensities were measured 30% through the thickness of the cortical ribbon, normal to the gray/white border for gray matter, and 1 mm subjacent to the gray/white border along the surface normal for white matter. The 30% sampling procedure was utilized to be conservatively close to the gray/white border and the white matter sampling voxel (which should minimize potential spurious effects that could arise in sampling from more remote locations) and was additionally able to adjust in regions of low cortical thickness (as opposed to using a constant value across the entire border which could be problematic for thinner cortical areas).

We extracted averaged cortical thickness and GWRs in 74 labels per hemisphere from the Destrieux atlas (57) respectively, and thus each participant had 296 features from the brain surface. The subcortical segmentation volumes (16 ROIs, eTIV, WMSA) were log-transformed, and finally a total of 314 features for each participant were prepared for analysis.

Feature selection by linear detrending for age

Linear regression models were created between age and each structural feature from D1. We evaluated all linear regression models (n = 314), and preserved the coefficients (β) and offsets (x_0) only when the feature significantly correlated with age at $p < 0.05$ without correction.

$$\text{feature} = \beta \times \text{age} + x_0$$

Then ages from D2 and D3 were put into the corresponding linear models, and then we calculated the deviation as the difference between the actual feature and predicted one by coefficients (β) and offsets (x_0) in each model. The following m is the number of significant models evaluated above.

$$\varepsilon_m = \text{feature}_m - (\beta_m \times \text{age} + x_{m0})$$

After detrending, each participant had a vector of m deviations which represented the features without linear age effect. We had a 136 × m matrix for the AD group and 268 × m matrix for controls in D2, as well as 66 × m matrix for all participants in D3.

Support vector machine classifier

Support vector machine (SVM) is a commonly utilized supervised, multivariate classification method. The SVM classifier finds a hyperplane maximizing the margin between groups. The problem of AD detection in D2 using SVM was formulated as a binary classification problem. In this study, we used the SVM implementation publicly available in LibSVM (csie.ntu.edu.tw/cjlin/libsvm). The cost parameter C and kernel parameter γ of the nonlinear Gaussian function in the SVM classifier were optimized using cross-validation via the grid-search approach (58). The grid search was performed over the ranges $C = 2^{-5}, 2^{-4}, \ldots 2^{15}$, $\gamma = 2^{-15}, 2^{-14}, \ldots, 2^{5}$. The optimized set of parameters was then used to train the SVM classifier by the input of features after linear detrending from D2. We used the 10-fold cross-validation to obtain an unbiased estimate of the classifier performance. During each fold the classifier was developed using data from 90% of the participants and tested using data from the remaining 10% of the participants.

In this study, we adopted the F-score method in the libsvm feature selection tool to further evaluate feature contribution for classification. The F-score method has been generally used in pattern recognition systems to select the optimal feature subset (59, 60). A larger F-score value indicates that the feature has more discriminative power. Finally, we use the well-developed classifier to predict each participant in D3 by the input of 66 × m matrix. The overall accuracy, specificity and sensitivity were reported as the generalization of the classifier. We also compared the results by putting different kinds of features into the SVM classifier.

Statistics

Group differences in cross-sectional demographics were assessed using t-tests for continuous and Chi-squared tests for nominal data. We applied linear model to adjust age effect in the MCI baseline comparison. To compare the longitudinally cognitive changes in MCI, we performed a mixed model procedure for repeated measures, and evaluated the interaction between classified group and follow-up timelines (Pinheiro J, R Core Team, R package version 3.1-144).

Results

Linear detrending for age in D1

Among the 148 surface parcellations, cortical thickness in 10 regions was not significantly correlated with age and then excluded, including middle-anterior part of the cingulate gyrus and sulcus, long insular gyrus and central sulcus of the insula, inferior temporal gyrus, anterior transverse collateral sulcus in the left hemisphere, fronto-marginal gyrus and sulcus, cuneus, orbit

gyrus, and pericallosal sulcus in the right hemisphere, and bilateral short insular gyri. In the 16 subcortical segmentations, only left caudate was excluded from the first-level selection. The GWR in 148 parcellations were all found correlated with age and included (Figure 1), as well as WMSA and eTIV. The p values for each regression here was not corrected for multiple comparisons as the primary goal of this analysis was to determine age trends for each structure. In all, 303 age-related features were found, and their coefficients with age and corresponding offsets were used for calculating deviations in AD and controls.

Patterns of deviation in D2

The deviation matrix of AD was 136 × 303, and that of controls was 268 × 303 (138 features in thickness, 148 in GWRs, 15 in subcortical volumes, WMSA and eTIV). In the heatmap, each column represented one participant, and each row represented regional difference between the actual measure and predicted one from age regression best-fit line (Min-max normalization across two groups, Figure 2). Among 138 surface regions, cortical thickness of AD deviated more negatively than that of controls in 126 regions (mean error difference: - 0.478 to - 0.029, $p<0.05$), while the deviations were similar in the rest 21 regions. AD also deviated more negatively in the age-subcortical volume best-fit line in 8 regions, except left pallidum. AD had more positive deviations from the age-GWR fit line in 129 surface regions (mean error difference: 0.003 to 0.005, $p<0.05$) and from age-WMSA line as well (Additional tables show this in more detail [see Additional file 1]).

Performance of SVM classifier

To illustrate the performance of the SVM classifier, deviations in 138 parcellations of mean cortical thickness, 148 GWRs, 15 subcortical volumes, WMSA and eTIV were extracted from the 404 participants in D2 and used as features. A total of 303 features were used for classification. The sensitivity, specificity, and accuracy were calculated to measure the performance of the SVM classifier. Sensitivity is defined as the proportion of true positives that are correctly identified by the test and specificity is defined as the proportion of true negatives that are correctly identified by the test. Accuracy is calculated as the proportion of true results (both true positives and true negatives) by the test. In D2, the optimal cross-validation accuracy of 93.07%, total accuracy of 97.03%, sensitivity of 0.941 (128/136), and specificity of 0.985 (264/268) were obtained using the SVM classifier. Without linear detrending for age before SVM, the optimal cross-validation accuracy was 91.83%, total accuracy 96.29%, sensitivity 0.919 (125/136), and specificity 0.985 (264/268).

The top 5 features whose F-score values were larger than 0.4 were volumes of left hippocampus (F-score, 0.696), right hippocampus (0.605), left amygdala (0.572), thickness of left (0.550) and right (0.414) medial temporal & parahippocampus gyrus. The volume of right amygdala (0.382), left middle temporal gyrus thickness (0.366), left and right temporal pole thickness (0.356), and volume of left superior temporal gyrus (0.311) ranked from 6 to 10 (Figure

3). The greatest impacts from GWR were in left (0.275) and right (0.228) medial temporal & parahippocampus gyrus.

The optimal cross-validation accuracy decreased to 90.84% when only cortical thickness included in the D2 SVM model, to 84.65% when only GWR included, to 89.10% with only subcortical volumes, and to 83.17% with only WMSA. Similarly, the total accuracy also decreased by 0.74%, 3.47%, 6.93%, 13.86% with only one of the above measures in the D2 SVM model.

In D3, we had 41 ADs with positive β amyloid deposition and poor cognitive performance and 25 controls with negative β amyloid and normal cognition. When we applied the classifier to D3 for validation, the classifier reported an accuracy of 84.85% (56/66), sensitivity of 85.36% (35/41) and specificity of 84% (21/25). Similar to D2, overall accuracy also decreased when only one of these structural measures included in the classifier, by 3.03%, 15.1%, 9.10%, 21.21% respectively.

Table 1 Subjects demographics from the three datasets

	Dataset 1	Dataset 2 (ADNI)		Dataset 3	
	HCP	AD	Controls	AD	Controls
Number of subjects	272	136	268	41	25
Sex (Female/Male)	146 / 126	57 / 79	148 / 120	19 / 22	12 / 13
Age (year)	62.7 ± 16.8	74.2 ± 8.2	72.9 ± 6.0	68.7 ± 9.0	68.5 ± 6.1
Education (year)	15.3 ± 4.5	15.7 ± 2.5	16. 6 ± 2.5	12.3 ± 3.3	13.2 ± 3.1
MoCA	26.2 ± 2.6	17.2 ± 4.5	25.8 ± 2.4	16.2 ± 6.8	27.2 ± 1.8
MMSE	-	23.0 ± 2.1	29.1 ± 1.1	20.7 ± 6.2	27.9 ± 2.0
ACE-R	-	-	-	60.5 ± 21.9	84.6 ± 18.6

Prediction of MCI by the classifier

In the 180 EMCI, 29 patients were classified as AD (EMCI-AD), and in the 96 LMCI, 41 were classified as AD (LMCI-AD). The EMCI-AD and LMCI-AD both had worse performance in MoCA and ADAS-13 at baseline compared with EMCI and LMCI classified as control (EMCI-CN, LMCI-CN) respectively ($p<0.001$, after adjusting for age). LMCI-AD additionally performed worse in baseline MMSE, learning of RAVLT and TMT-B ($p = 0.006$, 0.009, and 0.015 after adjusting for age, Figure 4). Longitudinally, EMCI-AD had a greater decline in ADAS-13 and learning of RAVLT by month 36 compared to EMCI-CN, and LMCI-AD had a more significant decrease in ADAS-13 and MoCA by month 24 (Table 2). To perform a more conservative analysis further minimizing age effects in the classification results, we examined only age and education matched EMCI-CN/AD and LMCI-CN/AD cases. Similar effects of classification were found on cognition measured by MoCA and ADAS-13 since month 48 in EMCI and MoCA, MMSE and ADAS-13 since month 24 in LMCI (Table 3).

Table 2 The classification of MCI and longitudinally cognitive changes

	EMCI			LMCI		
	AD	CN	P-value	AD	CN	P-value
N	29	151	-	41	55	-
Age	72.9 ± 6.0	70.2 ± 6.9	0.031	73.3 ± 6.4	69.0 ± 7.3	0.002
Sex (Female/Male)	12/17	66/85	0.978	23/18	26/29	0.516
Education	16.5 ± 3.0	16.5 ± 2.6	0.407	16.9 ± 2.6	16.2 ± 2.6	0.167

	Interaction* (p-value)	EMCI# Effect estimate (Beta)	95% CI of Beta	Interaction* (p-value)	LMCI# Effect estimate (Beta)	95% CI of Beta
ADAS-13	m36 (0.020)	3.60	1.25 ~ 5.95	m24 (<0.001)	5.07	2.10 ~ 8.03
	m48 (<0.001)	8.84	6.38 ~11.3	m36 (<0.001)	9.32	6.10 ~ 12.5
				m48 (<0.001)	11.2	7.64 ~ 14.8
MMSE	m06 (0.003)	-1.32	-2.20 ~ -0.45	m24 (0.002)	-2.00	-3.30 ~ -0.71
	m12 (0.006)	-1.21	-2.08 ~ -0.33	m36 (<0.001)	-4.71	-6.12 ~ -3.31
	m24 (0.004)	-1.31	-2.20 ~ -0.41	m48 (<0.001)	-6.55	-8.11 ~ -4.98
	m36 (<0.001)	-2.10	-3.02 ~ -1.17			
	m48 (<0.001)	-2.33	-3.29 ~-1.36			
MOCA	-	-	-	m24 (0.001)	-2.25	-3.63 ~ -0.88
				m36 (<0.001)	-3.62	-5.10 ~ -2.13
				m48 (<0.001)	-4.49	-6.16 ~ -2.80
RAVLT-learning	m36 (0.047)	-1.14	-2.26 ~ -0.01	-	-	-
RAVLT-forgetting	m12 (0.002)	-1.92	-3.16 ~ -0.67	-	-	-
	m24 (0.013)	-1.61	-2.88 ~ -0.33			
	m36 (0.013)	-1.69	-3.02 ~ -0.35			
TMT-B	-	-	-	m12 (0.024)	27.3	3.62 ~ 51.1
				m24 (<0.001)	57.6	33.0 ~ 82.2
				m36 (<0.001)	48.8	21.5 ~ 76.0
				m48 (0.030)	35.8	3.51 ~ 68.0

Adjusted for age, sex and years of education

* Interaction between follow-up and classification, with baseline as reference, only shown with $p<0.05$;

m, month of follow-up.

Table 3 Cognitive changes in age and education matched group for EMCI-AD and LMCI-AD

	EMCI			LMCI		
	AD	CN	P-value	AD	CN	P-value
N	29	29	-	41	41	-
Age	72.9 ± 6.0	73.6 ± 7.2	0.702	73.3 ± 6.4	71.6 ± 6.0	0.234
Sex (Female/Male)	12/17	12/17	0.978	23/18	18/23	0.377
Education	16.5 ± 3.0	15.9 ± 2.6	0.406	16.9 ± 2.1	16.7 ± 2.4	0.735
MMSE-Baseline	28.3 ± 1.6	28.3 ± 1.6	0.933	27.0 ± 1.7	28.0 ± 1.7	0.007
Follow up*	ns		-	m24		0.003
				m36		<0.001
				m48		<0.001
MOCA-Baseline	21.9 ± 2.4	24.5 ± 2.7	<0.001	20.6 ± 2.5	23.6 ± 3.1	<0.001
Follow up*	m48		0.054	m24		0.008
				m36		<0.001
				m48		<0.001
ADAS-Baseline	17.0 ± 5.9	12.0 ± 5.0	<0.001	22.3 ± 5.8	16.4 ± 6.3	<0.001
Follow up*	m48		0.007*	m24		<0.001
				m36		<0.001
				m48		<0.001

* Significant interaction between follow-up and classification, with baseline as reference; m, month of follow-up.

Discussion

Given the correspondence between patterns of cerebral atrophy measured by MRI and the pathological processes of AD, it is theoretically possible to use these cost-effective and accessible measurements in the early detection of individuals with AD neuropathology. Prior studies have identified regional atrophy patterns in patients with AD (19, 61) subsequently used as features in machine learning models for patient classification (1). However, several prior studies have noted a range of factors that contribute to the accuracy of this approach and therefore limiting clinical generalizability. Although the control of confounding variables is reasonable and straightforward in linear statistical models, the generalization of these learned models is limited for classifying participants with wide ranging values and especially when this information is not explicitly modeled in the training data.

Age is one of the most consequential factors limiting generalizability as younger individuals may have brain structures that seem generally more ‘healthy’ even in the face of degeneration relative to older adults that may show typical and relatively benign patterns of atrophy (34, 62, 63). Moreover, age could partially mask other disease-related factors such as ApoE genotype, global cognitive impairment and sex (33). Thus, here we explicitly removed this age confound from the classifier using linear detrending of features to enhance classification performance.

There have been a range of methods proposed for estimating relationship between age and brain features (i.e., choice of imaging-derived features to use, and choice of supervised-learning approach). Generally, linear models have limitation when datasets do not have a Gaussian distribution (across all subjects) for dependent variables, and errors in measurement of predictors likely cause regression dilution (64). However, in our linear model framework, the chronological

age was used as predictor and assumed of no errors. Meanwhile, it does not generally matter what the distribution of the predictors is (here, age) (65). We applied linear detrending prior to SVM since our controls from HCP dataset are widely distributed on the age range and could be representative of population. Thus, the model may be more validly applied to both younger and older adults in clinics.

For simplicity, we used easily obtained measures from anatomical parcellations instead of voxel or vertex based features more specific to the pathologic patterns of AD or using whole voxel-wise patterns as in prior work (66). The features that greatly contributive to our classifier were similar to 'AD signature' (67). Moreover, it is suggested that MRI-based morphometric estimates (cortical thickness in 'AD signature' regions, hippocampal volume and global atrophy) parallel CSF neurofilament light chain protein in differentiating individuals across the AD continuum on neurodegeneration status (68). Although use of surface vertex based features could provide more information, we were able to achieve high accuracy using the more generally accessible automated features and it is possible that greatly expanding the feature set would result in overfitting in high-dimension models (69). Also, it is possible that more precise features would be less generalizable across samples and scanners. The parcellation based analysis also make application easier and more straightforward. We will explore the potential to expand the feature set without overfitting in future work.

Among features we included, the most prominent classification power was provided through volumes of hippocampus and amygdala. The results were consistent with observations from clinical practice and several prior research studies demonstrating early and profound atrophy in these regions with MCI and AD (9, 51, 70-72). In typical AD, medial temporal atrophy affects the amygdala and the hippocampus is usually accompanied by temporal horn enlargement and higher hazard ratio for disease progression (8). Stronger atrophy in the amygdala and hippocampus predicted conversion to AD, and the linear discriminant analysis on the principal component values of hippocampus, amygdala, and ventricular volume provided classification 86% - 88% for cognitive impairment (73). Age interacted with AD pathology to impact feature quantification. Before age correction, pronounced reduction in rates of atrophy are observed in individuals with AD with increasing age, while for cognitively unimpaired individuals, increasing age leads to increased rates of atrophy (74). We had cross-validation accuracy of 89.1% with only subcortical features included, however the accuracy increased with the addition of structural features. Cortical thinning provided essential information for AD diagnosis and thinning in vulnerable cortical regions relates to symptom severity even in the earliest stages of clinical symptoms. Furthermore, subtle thinning is present in asymptomatic older controls with brain amyloid binding as detected with amyloid imaging (11). Given the relative accessibility of the features used in this work, the inclusion of the full feature set did not limit general utilization of the procedures.

GWR was reported as age- and AD-related (6, 12, 25, 75), while its value for classification was unclear. We previously examined the utility of GWR in identification of MCI progression(15). To our knowledge, this is the first examination of the utility of the GWR measure in participant classification. Decreased contrast between gray and white matter in the bilateral medial temporal gyrus and poles provided the most information among the GWR features. In the present study, the GWR measures improved accuracy compared to the structural features alone and the specificity increased in the classifier with only GWRs, while overall accuracy and sensitivity decreased dramatically in the case. Compared with cortical thickness and subcortical volumes, the GWRs also

had much lower F-scores. In all, we suggested GWRs less sensitive than other structural measures, but still helpful due to its specificity for AD classification. A limitation of this work is that the images from ADNI as well as the additional dataset examined were acquired with a range of imaging protocols that would be expected to have a major impact on the image signal and contrast properties. It is therefore possible that the GWR metrics would be more powerful in the context of a single imaging platform.

White matter lesions measured as WMSA on MRI are commonly seen in normal aging and AD (76), and histopathological studies have indicated a mix of heterogeneous findings that correlate with WMSA including demyelination and gliosis (77). Individuals with AD have a higher total and greater regional volumes of WMSA regardless of age (2, 3, 5) and white matter lesions are associated with more rapid cognitive decline across time in MCI and AD(78). In the present study, we found that the deviation of total WMSA from age-regression line only had a minor impact on the SVM classifier (F-score 0.023). This may be due to the fact that that WMSA volume is highly correlated with other prominent features of AD such as hippocampal volume(4).

In the SVM, we set the deviations from the best fit lines as potential features as input. Thereby, age-related effects on MRI data were minimized and we focused on how far a person deviated from age-predicted measures. Although relatively successful for age correction, reviews of prior work demonstrated SVM model accuracy ranging from 85 to 91% (1, 33, 34). Cortical thickness in AD signature regions, hippocampus volume and global atrophy discriminated between AD and controls with sensitivity from 80-90%, specificity from 85-90%(68). Using age-detrended features, we had an increase in model accuracy of 2-8% (cross-validation accuracy: 93.1%, total accuracy: 97.03%, sensitivity: 94.1%, specificity: 98.5%). Critically, using a novel dataset with substantial differences from the training dataset (differing in clinical sample, racial composition, and scanner type and imaging protocols), a biomarker confirmed accuracy of 84.8% was achieved demonstrating generalization at least across these conditions.

A second validation was provided through examination of longitudinal cognitive changes in individuals with MCI. Participants classified as AD in the early and late MCI samples (EMCI-AD and LMCI-AD) were older than patients classified as controls. This may still be expected given the increase in prevalence of AD with increasing age and given potential age interactions with this condition (79). We conservatively performed an additional matching of the MCI groups and still observed significantly greater cognitive decline in both EMCI-AD and LMCI-AD compared to matched MCI classified as more control-like. Moreover, the LMCI-AD showed longitudinal differences earlier than EMCI-AD supporting their later stage of impairment.

Several limitations of the current work are being explored in ongoing research. First, we corrected age effect based on the cross-sectional dataset and the relationship between structural features and age could be ambiguous due to subject variability. Although linear regression is widely used in age-brain analysis, non-linear detrending for age could also be considered with longitudinal datasets. Second, we used the parcellations instead of surface features more specific to the effects. With careful check of overfitting as discussed before, we could potentially enhance performance using specific age-related surface features. Third, we examined performance of the SVM to differentiate a control group from a patient group. In the true clinical setting, classification would need to be performed among a range of distinct conditions that could contribute to impairment. We are exploring the potential of the procedures described here to differentiate among differing neurodegenerative diseases. Similarly, we will explore these procedures in the

identification of subgroups of AD, such as posterior cerebral atrophy that could be decoded based on spatial patterns. The differentiation of MCI that differ in cognitive trajectories provide some confidence that subgroups can be identified within the clinical sample. Finally, improvement for the proposed method could involve the inclusion of clinical variables that are also associated with structural measures in the linear detrending model prior to SVM.

Conclusions

The major conclusions from current work are as follows: (a) Linear detrending for age and followed SVM for combined structural measures provided better performance in classification, as well as classifier generalization. (b) Compared to other structural measures, the volumes of hippocampus and amygdala, as well as the cortical thickness in medial and superior temporal gyrus, had more important influence on classification accuracy after removal of age effect. (c) The classifier could help recognition of AD, as well as cognitive decline in MCI. Future work will examine the degree to which subgroups can be determined from the classified AD individuals based on detrending with addition of spatial parameters.

List of abbreviations

AD, Alzheimer's disease; PET, positron emission tomography; MRI, magnetic resonance imaging; GWR, gray to white matter signal intensity ratio; EMCI, early mild cognitive impairment; LMCI, late mild cognitive impairment; WMSA, white matter signal abnormalities; SVM, support vector machine; ADAS-13, Alzheimer’s Disease Assessment Scale Cognitive Subscale; RAVLT, Rey’s auditory verbal learning test; TMT, trail making test; eTIV, estimated total intracranial volume (eTIV)

Additional file

File name: Additional file 1
File format: .docx
Title of data: contribution of each cortical and subcortical regions in the classifier
Description of data: The tables list deviations from age-regression best fit line and F-scores of cortical thickness, gray/white matter ratio, subcortical volumes in each region of interest.

Declarations

Ethics approval and consent to participate

For HCP and ADNI data, we comply with institutional rules and regulations. For data3, its collection and analysis was approved by the Ethics Committee, Shanghai Jiao Tong University affiliated Ruijin Hospital, China.

Consent for publication

Not applicable

Availability of data and materials

The data used in this study were partially obtained from the ADNI and are available from the ADNI database upon registration and compliance with the data usage agreement. Data from the HCP-A cohort are in the process of being made freely available online and D3 are available upon appropriate request to the corresponding author.

Competing interests

The authors declare that they have no competing interests

Funding

D1 reported in this publication was supported by grants U01AG052564 and U01AG052564-S1 and by the 14 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research, by the McDonnell Center for Systems Neuroscience at Washington University, by the Office of the Provost at Washington University, and by the University of Minnesota Medical School. D3 was supported by the Science and Technology Commission of Shanghai Municipality (18JC1420303).

Author's contributions

ByL analyzed and interpreted all imaging data and drafted the manuscript. MZ performed PET-MR for D3 and was a major contributor in writing the manuscript. JR commented in the study. KMY collected HCP data. BL and JL helped D3 collection. DHS supervised, reviewed and modified the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank our anonymous reviewers for thoughtful comments on the manuscript, and Stephen M. Smith for help with methodology. We are also grateful to all research participants.

Data collection and sharing for D2 and MCI in 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.

Figure 1 Age-feature regression

Plots showing the best-fit regression coefficients between age and cortical thickness (A), ratio of gray to white matter signal intensity (GWR) (B), subcortical volumes (C) from 272 normal participants of Human Connectome Project. The name of 148 cortical regions for cortical thickness

and GWRs, as well as 16 subcortical regions, are listed in supplementary materials. CI, confidential intervals.

Figure 2 Different deviations from regression in AD and controls

Left: In each heatmap, columns represented participants, and rows represented regional difference between the actual measure and predicted one from age regression line (each row rescaled by max-min normalization across two groups, with the maximum = 1 and minimum = 0). Darker color showed relatively more negative deviation, while lighter color suggested more positive deviation when rows were compared between AD and controls from D2.

Right: Averaged difference between value of each row in two groups from the left heatmap (normal controls - AD). The regions that the rows represented were marked on the right. AD, Alzheimer's disease; GWR, ratio of gray to white matter signal intensity.

Figure 3 Mapping of most contributive features for classification

Maps of subcortical volumes and cortical thickness with F-score larger than 0.3, in the SVM classifier from the controls and AD groups in D2.

Figure 4 Longitudinal changes in classified MCI

Boxplot of cognitive performance at each visit for MCI. MMSE, mini-mental state examination; MoCA, Montreal cognitive assessment; ADAS: Alzheimer's Disease Assessment Scale Cognitive Subscale; RAVLT, Rey's auditory verbal learning test; TMT-B, trail making test-B; EMCI, early mild cognitive impairment; LMCI, late mild cognitive impairment; CN, control; bl, baseline; m, month.

Reference

1. Belathur Suresh M, Fischl B, Salat DH, Alzheimer's Disease Neuroimaging I. Factors influencing accuracy of cortical thickness in the diagnosis of Alzheimer's disease. Human brain mapping. 2018;39(4):1500-15.
2. Lindemer ER, Greve DN, Fischl B, Augustinack JC, Salat DH, Alzheimer's Disease Neuroimaging I. Differential Regional Distribution of Juxtacortical White Matter Signal Abnormalities in Aging and Alzheimer's Disease. J Alzheimers Dis. 2017;57(1):293-303.
3. Lindemer ER, Greve DN, Fischl BR, Augustinack JC, Salat DH. Regional staging of white matter signal abnormalities in aging and Alzheimer's disease. Neuroimage Clin. 2017;14:156-65.
4. Coutu JP, Goldblatt A, Rosas HD, Salat DH, Alzheimer's Disease Neuroimaging I. White Matter Changes are Associated with Ventricular Expansion in Aging, Mild Cognitive Impairment, and Alzheimer's Disease. J Alzheimers Dis. 2016;49(2):329-42.
5. Lindemer ER, Salat DH, Smith EE, Nguyen K, Fischl B, Greve DN, et al. White matter signal abnormality quality differentiates mild cognitive impairment that converts to Alzheimer's disease from nonconverters. Neurobiol Aging. 2015;36(9):2447-57.
6. Salat DH, Chen JJ, van der Kouwe AJ, Greve DN, Fischl B, Rosas HD. Hippocampal degeneration is associated with temporal and limbic gray matter/white matter tissue contrast in Alzheimer's disease. NeuroImage. 2011;54(3):1795-802.
7. Desikan RS, Sabuncu MR, Schmansky NJ, Reuter M, Cabral HJ, Hess CP, et al. Selective disruption of the cerebral neocortex in Alzheimer's disease. PLoS One. 2010;5(9):e12853.

8. Desikan RS, Cabral HJ, Settecase F, Hess CP, Dillon WP, Glastonbury CM, et al. Automated MRI measures predict progression to Alzheimer's disease. Neurobiology of aging. 2010;31(8):1364-74.
9. Desikan RS, Cabral HJ, Hess CP, Dillon WP, Glastonbury CM, Weiner MW, et al. Automated MRI measures identify individuals with mild cognitive impairment and Alzheimer's disease. Brain. 2009;132(Pt 8):2048-57.
10. Salat DH, Greve DN, Pacheco JL, Quinn BT, Helmer KG, Buckner RL, et al. Regional white matter volume differences in nondemented aging and Alzheimer's disease. Neuroimage. 2009;44(4):1247-58.
11. Dickerson BC, Bakkour A, Salat DH, Feczko E, Pacheco J, Greve DN, et al. The cortical signature of Alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild AD dementia and is detectable in asymptomatic amyloid-positive individuals. Cereb Cortex. 2009;19(3):497-510.
12. Salat DH, Lee SY, van der Kouwe AJ, Greve DN, Fischl B, Rosas HD. Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. NeuroImage. 2009;48(1):21-8.
13. Walhovd KB, Westlye LT, Amlien I, Espeseth T, Reinvang I, Raz N, et al. Consistent neuroanatomical age-related volume differences across multiple samples. Neurobiol Aging. 2011;32(5):916-32.
14. Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, et al. Thinning of the cerebral cortex in aging. Cereb Cortex. 2004;14(7):721-30.
15. Jefferson AL, Gifford KA, Damon S, Chapman GWt, Liu D, Sparling J, et al. Gray & white matter tissue contrast differentiates Mild Cognitive Impairment converters from non-converters. Brain Imaging Behav. 2015;9(2):141-8.
16. Chetelat G, Baron JC. Early diagnosis of Alzheimer's disease: contribution of structural neuroimaging. NeuroImage. 2003;18(2):525-41.
17. Park M, Moon WJ. Structural MR Imaging in the Diagnosis of Alzheimer's Disease and Other Neurodegenerative Dementia: Current Imaging Approach and Future Perspectives. Korean J Radiol. 2016;17(6):827-45.
18. Teipel SJ, Grothe M, Lista S, Toschi N, Garaci FG, Hampel H. Relevance of magnetic resonance imaging for early detection and diagnosis of Alzheimer disease. Med Clin North Am. 2013;97(3):399-424.
19. de Vos F, Schouten TM, Hafkemeijer A, Dopper EG, van Swieten JC, de Rooij M, et al. Combining multiple anatomical MRI measures improves Alzheimer's disease classification. Human brain mapping. 2016;37(5):1920-9.
20. Eskildsen SF, Coupe P, Garcia-Lorenzo D, Fonov V, Pruessner JC, Collins DL. Prediction of Alzheimer's disease in subjects with mild cognitive impairment from the ADNI cohort using patterns of cortical thinning. NeuroImage. 2013;65:511-21.
21. Raamana PR, Weiner MW, Wang L, Beg MF, Alzheimer's Disease Neuroimaging I. Thickness network features for prognostic applications in dementia. Neurobiology of aging. 2015;36 Suppl 1:S91-S102.
22. Jack CR, Jr., Petersen RC, Xu Y, O'Brien PC, Smith GE, Ivnik RJ, et al. Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. Neurology. 2000;55(4):484-89.
23. Gosche KM, Mortimer JA, Smith CD, Markesbery WR, Snowdon DA. Hippocampal volume as an index of Alzheimer neuropathology: findings from the Nun Study. Neurology. 2002;58(10):1476-82.
24. Kwak K, Yun HJ, Park G, Lee JM, Alzheimer's Disease Neuroimaging I. Multi-Modality Sparse

Representation for Alzheimer's Disease Classification. J Alzheimers Dis. 2018;65(3):807-17.
25. Westlye LT, Walhovd KB, Dale AM, Espeseth T, Reinvang I, Raz N, et al. Increased sensitivity to effects of normal aging and Alzheimer's disease on cortical thickness by adjustment for local variability in gray/white contrast: a multi-sample MRI study. NeuroImage. 2009;47(4):1545-57.
26. Salat DH. Imaging small vessel-associated white matter changes in aging. Neuroscience. 2014;276:174-86.
27. Lindemer ER, Greve DN, Fischl B, Salat DH, Gomez-Isla T. White matter abnormalities and cognition in patients with conflicting diagnoses and CSF profiles. Neurology. 2018;90(17):e1461-e9.
28. Ruiz E, Ramirez J, Gorriz JM, Casillas J, Alzheimer's Disease Neuroimaging I. Alzheimer's Disease Computer-Aided Diagnosis: Histogram-Based Analysis of Regional MRI Volumes for Feature Selection and Classification. J Alzheimers Dis. 2018;65(3):819-42.
29. Samper-Gonzalez J, Burgos N, Bottani S, Fontanella S, Lu P, Marcoux A, et al. Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. NeuroImage. 2018;183:504-21.
30. Aguilar C, Westman E, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, et al. Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. Psychiatry Res. 2013;212(2):89-98.
31. Cuingnet R, Gerardin E, Tessieras J, Auzias G, Lehericy S, Habert MO, et al. Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. NeuroImage. 2011;56(2):766-81.
32. Salvatore C, Battista P, Castiglioni I. Frontiers for the Early Diagnosis of AD by Means of MRI Brain Imaging and Support Vector Machines. Curr Alzheimer Res. 2016;13(5):509-33.
33. Falahati F, Ferreira D, Soininen H, Mecocci P, Vellas B, Tsolaki M, et al. The Effect of Age Correction on Multivariate Classification in Alzheimer's Disease, with a Focus on the Characteristics of Incorrectly and Correctly Classified Subjects. Brain Topogr. 2016;29(2):296-307.
34. Dukart J, Schroeter ML, Mueller K, Alzheimer's Disease Neuroimaging I. Age correction in dementia--matching to a healthy brain. PLoS One. 2011;6(7):e22193.
35. Jack CR, Jr., Therneau TM, Weigand SD, Wiste HJ, Knopman DS, Vemuri P, et al. Prevalence of Biologically vs Clinically Defined Alzheimer Spectrum Entities Using the National Institute on Aging-Alzheimer's Association Research Framework. JAMA Neurol. 2019:e191971.
36. Harms MP, Somerville LH, Ances BM, Andersson J, Barch DM, Bastiani M, et al. Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects. NeuroImage. 2018;183:972-84.
37. Bookheimer SY, Salat DH, Terpstra M, Ances BM, Barch DM, Buckner RL, et al. The Lifespan Human Connectome Project in Aging: An overview. Neuroimage. 2019;185:335-48.
38. Pascoal TA, Mathotaarachchi S, Shin M, Park AY, Mohades S, Benedet AL, et al. Amyloid and tau signatures of brain metabolic decline in preclinical Alzheimer's disease. European journal of nuclear medicine and molecular imaging. 2018;45(6):1021-30.
39. Katzman R, Zhang MY, Ouang Ya Q, Wang ZY, Liu WT, Yu E, et al. A Chinese version of the Mini-Mental State Examination; impact of illiteracy in a Shanghai dementia survey. Journal of clinical epidemiology. 1988;41(10):971-8.
40. Fang R, Wang G, Huang Y, Zhuang JP, Tang HD, Wang Y, et al. Validation of the Chinese version of Addenbrooke's cognitive examination-revised for screening mild Alzheimer's disease and mild cognitive impairment. Dementia and geriatric cognitive disorders. 2014;37(3-4):223-31.

41. Jack CR, Jr., Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement. 2018;14(4):535-62.
42. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord. 1997;11 Suppl 2:S13-21.
43. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci U S A. 2000;97(20):11050-5.
44. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage. 1999;9(2):195-207.
45. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage. 1999;9(2):179-94.
46. Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. Human brain mapping. 1999;8(4):272-84.
47. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, et al. Automatically parcellating the human cerebral cortex. Cerebral cortex. 2004;14(1):11-22.
48. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. NeuroImage. 2006;31(3):968-80.
49. Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, et al. Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology. 2002;58(5):695-701.
50. Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, et al. Regionally localized thinning of the cerebral cortex in schizophrenia. Archives of general psychiatry. 2003;60(9):878-88.
51. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002;33(3):341-55.
52. Salat DH, Tuch DS, van der Kouwe AJ, Greve DN, Pappu V, Lee SY, et al. White matter pathology isolates the hippocampal formation in Alzheimer's disease. Neurobiology of aging. 2010;31(2):244-56.
53. Salat DH, Williams VJ, Leritz EC, Schnyer DM, Rudolph JL, Lipsitz LA, et al. Inter-individual variation in blood pressure is associated with regional white matter integrity in generally healthy older adults. Neuroimage. 2012;59(1):181-92.
54. Leritz EC, Shepel J, Williams VJ, Lipsitz LA, McGlinchey RE, Milberg WP, et al. Associations between T1 white matter lesion volume and regional white matter microstructure in aging. Hum Brain Mapp. 2014;35(3):1085-100.
55. Riphagen JM, Gronenschild E, Salat DH, Freeze WM, Ivanov D, Clerx L, et al. Shades of white: diffusion properties of T1- and FLAIR-defined white matter signal abnormalities differ in stages from cognitively normal to dementia. Neurobiol Aging. 2018;68:48-58.
56. Dadar M, Maranzano J, Ducharme S, Collins DL, Alzheimer's Disease Neuroimaging I. White matter in different regions evolves differently during progression to dementia. Neurobiol Aging. 2019;76:71-9.
57. Destrieux C, Fischl B, Dale A, Halgren E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage. 2010;53(1):1-15.
58. Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST). 2011.

59. Chen F-L, Li F-C. Combination of feature selection approaches with SVM in credit scoring. Expert systems with applications. 2010;37(7):4902-9.
60. Polat K, Güneş S. A new feature selection method on classification of medical datasets: Kernel F-score feature selection. Expert Systems with Applications. 2009;36(7):10367-73.
61. Falahati F, Westman E, Simmons A. Multivariate data analysis and machine learning in Alzheimer's disease with a focus on structural magnetic resonance imaging. J Alzheimers Dis. 2014;41(3):685-708.
62. Koikkalainen J, Polonen H, Mattila J, van Gils M, Soininen H, Lotjonen J, et al. Improved classification of Alzheimer's disease data via removal of nuisance variability. PLoS One. 2012;7(2):e31112.
63. Li L, Rakitsch B, Borgwardt K. ccSVM: correcting Support Vector Machines for confounding factors in biological data classification. Bioinformatics. 2011;27(13):i342-8.
64. Le TT, Kuplicki RT, McKinney BA, Yeh HW, Thompson WK, Paulus MP. A Nonlinear Simulation Framework Supports Adjusting for Age When Analyzing BrainAGE. Front Aging Neurosci. 2018;10:317.
65. Smith SM, Vidaurre D, Alfaro-Almagro F, Nichols TE, Miller KL. Estimation of brain age delta from brain imaging. Neuroimage. 2019;200:528-39.
66. Kloppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain. 2008;131(Pt 3):681-9.
67. Schwarz CG, Gunter JL, Wiste HJ, Przybelski SA, Weigand SD, Ward CP, et al. A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease severity. Neuroimage Clin. 2016;11:802-12.
68. Allison SL, Koscik RL, Cary RP, Jonaitis EM, Rowley HA, Chin NA, et al. Comparison of different MRI-based morphometric estimates for defining neurodegeneration across the Alzheimer's disease continuum. Neuroimage Clin. 2019;23:101895.
69. Baumes LA, Serra JM, Serna P, Corma A. Support vector machines for predictive modeling in heterogeneous catalysis: a comprehensive introduction and overfitting investigation based on two real applications. J Comb Chem. 2006;8(4):583-96.
70. Tentolouris-Piperas V, Ryan NS, Thomas DL, Kinnunen KM. Brain imaging evidence of early involvement of subcortical regions in familial and sporadic Alzheimer's disease. Brain Res. 2017;1655:23-32.
71. Sengoku R. Aging and Alzheimer's disease pathology. Neuropathology. 2019.
72. Wachinger C, Salat DH, Weiner M, Reuter M, Alzheimer's Disease Neuroimaging I. Whole-brain analysis reveals increased neuroanatomical asymmetries in dementia for hippocampus and amygdala. Brain. 2016;139(Pt 12):3253-66.
73. Tang X, Holland D, Dale AM, Younes L, Miller MI, Alzheimer's Disease Neuroimaging I. Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: detecting, quantifying, and predicting. Human brain mapping. 2014;35(8):3701-25.
74. Holland D, Desikan RS, Dale AM, McEvoy LK, Alzheimer's Disease Neuroimaging I. Rates of decline in Alzheimer disease decrease with age. PLoS One. 2012;7(8):e42325.
75. Grydeland H, Westlye LT, Walhovd KB, Fjell AM. Improved prediction of Alzheimer's disease with longitudinal white matter/gray matter contrast changes. Human brain mapping. 2013;34(11):2775-85.
76. de Leeuw FE, de Groot JC, Achten E, Oudkerk M, Ramos LM, Heijboer R, et al. Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study. The Rotterdam Scan Study. J Neurol Neurosurg Psychiatry. 2001;70(1):9-14.
77. Gouw AA, Seewann A, van der Flier WM, Barkhof F, Rozemuller AM, Scheltens P, et al.

Heterogeneity of small vessel disease: a systematic review of MRI and histopathology correlations. J Neurol Neurosurg Psychiatry. 2011;82(2):126-35.
78. Coutu JP, Lindemer ER, Konukoglu E, Salat DH, Alzheimer's Disease Neuroimaging I. Two distinct classes of degenerative change are independently linked to clinical progression in mild cognitive impairment. Neurobiol Aging. 2017;54:1-9.
79. Yao Z, Hu B, Liang C, Zhao L, Jackson M, Alzheimer's Disease Neuroimaging I. A longitudinal study of atrophy in amnestic mild cognitive impairment and normal aging revealed by cortical thickness. PLoS One. 2012;7(11):e48973.

Figures

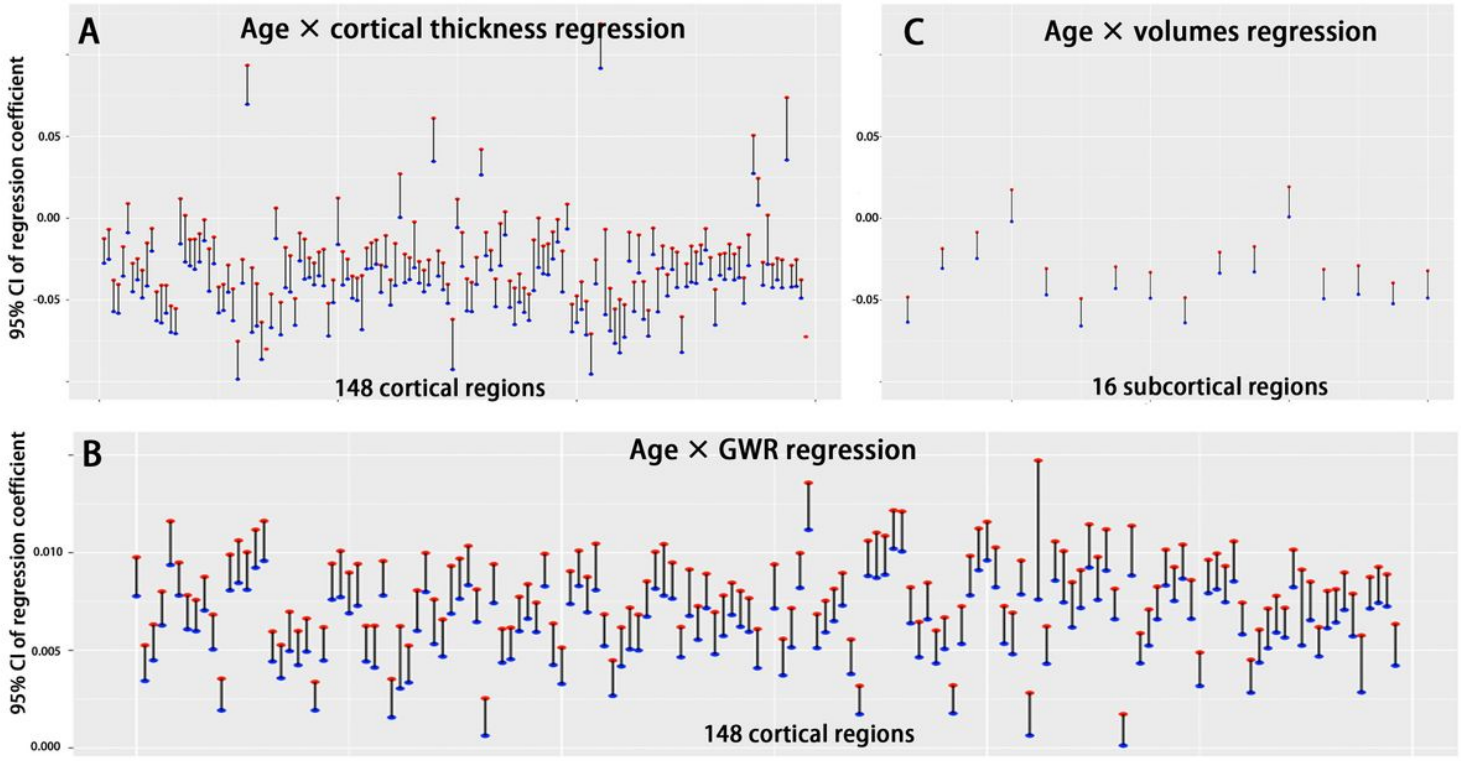

Figure 1

Age-feature regression Plots showing the best-fit regression coefficients between age and cortical thickness (A), ratio of gray to white matter signal intensity (GWR) (B), subcortical volumes (C) from 272 normal participants of Human Connectome Project. The name of 148 cortical regions for cortical thickness and GWRs, as well as 16 subcortical regions, are listed in supplementary materials. CI, confidential intervals.

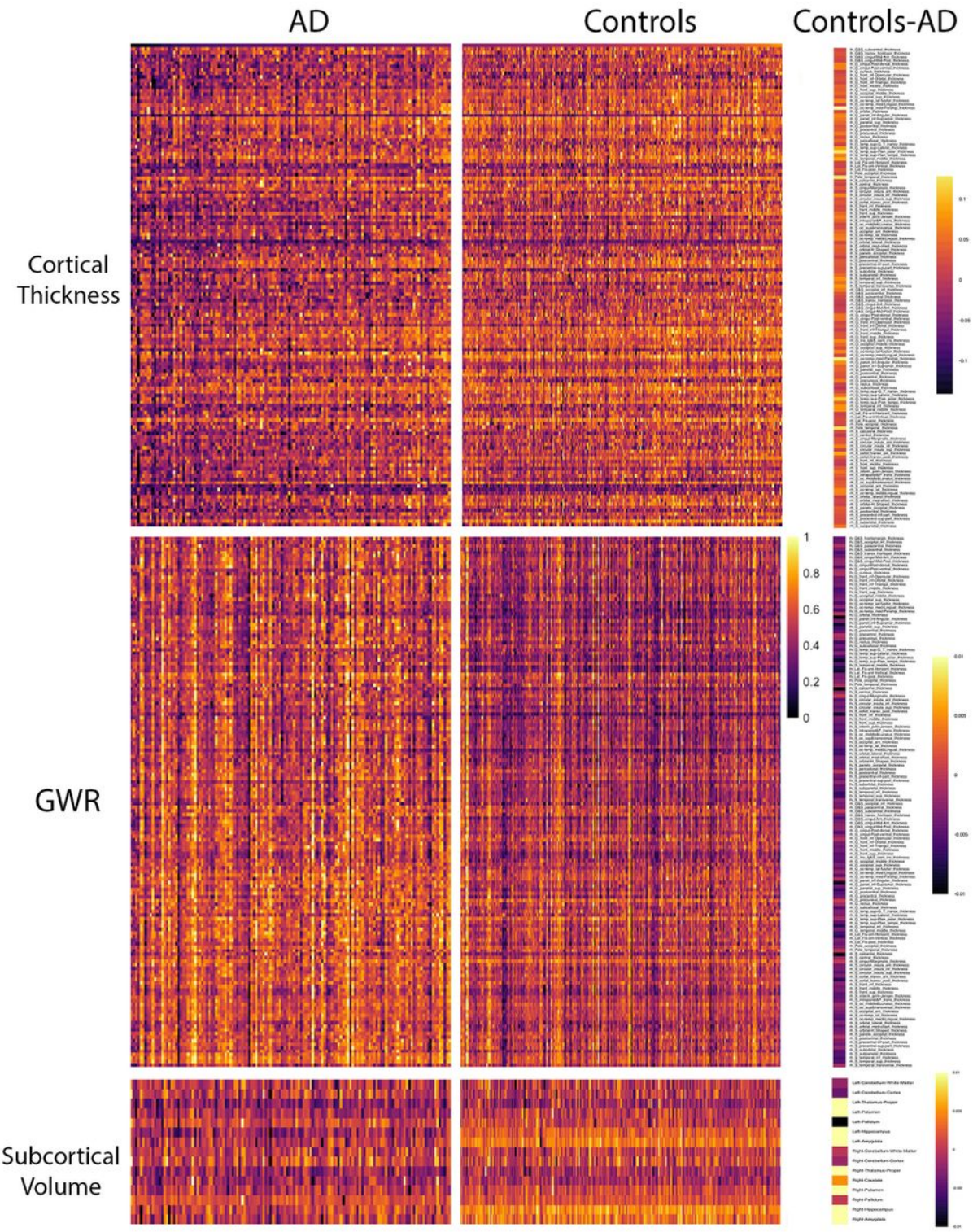

Figure 2

Different deviations from regression in AD and controls Left: In each heatmap, columns represented participants, and rows represented regional difference between the actual measure and predicted one from age regression line (each row rescaled by max-min normalization across two groups, with the maximum = 1 and minimum = 0). Darker color showed relatively more negative deviation, while lighter color suggested more positive deviation when rows were compared between AD and controls from D2.

Right: Averaged difference between value of each row in two groups from the left heatmap (normal controls - AD). The regions that the rows represented were marked on the right. AD, Alzheimer's disease; GWR, ratio of gray to white matter signal intensity.

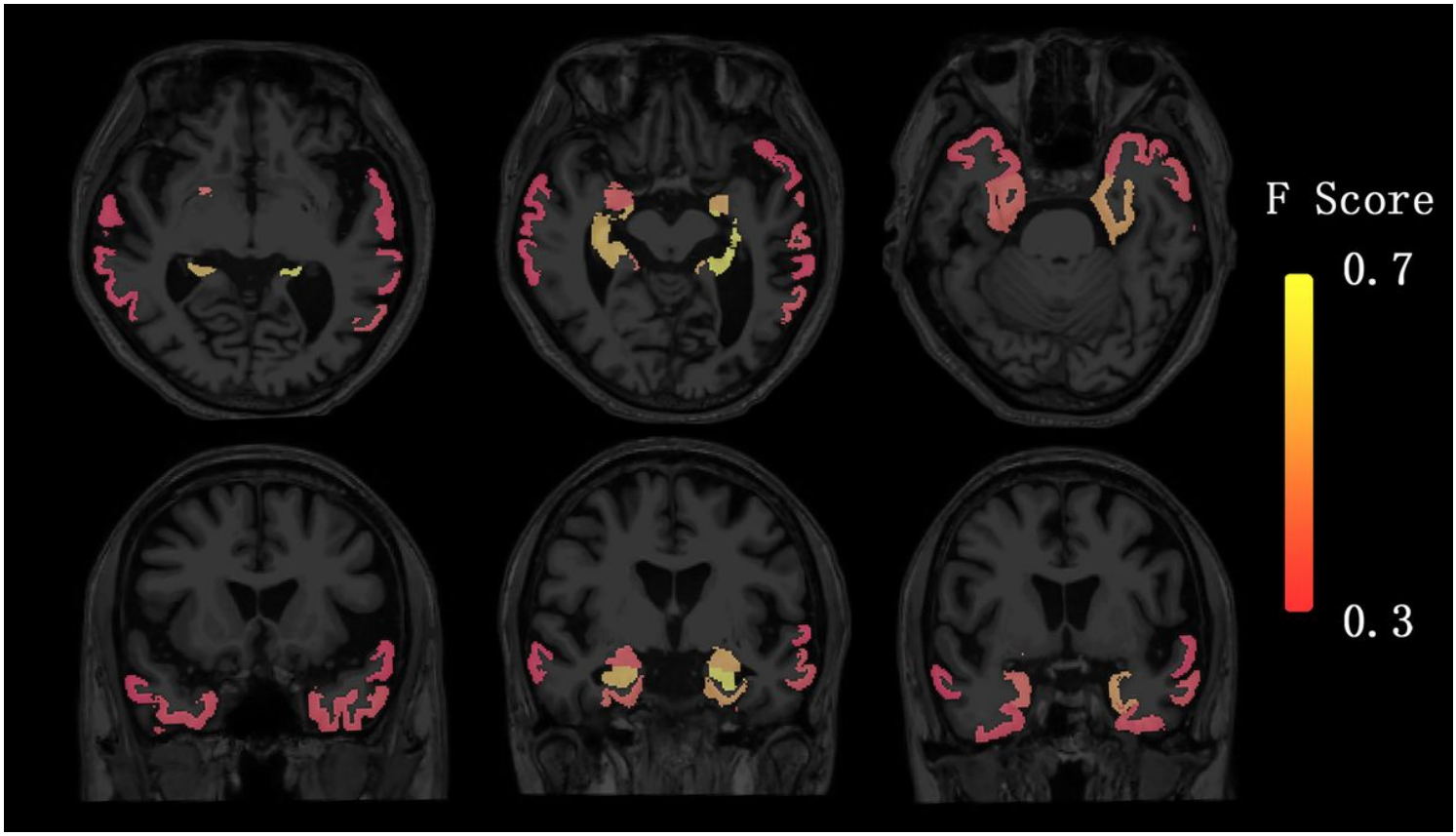

Figure 3

Mapping of most contributive features for classification Maps of subcortical volumes and cortical thickness with F-score larger than 0.3, in the SVM classifier from the controls and AD groups in D2.

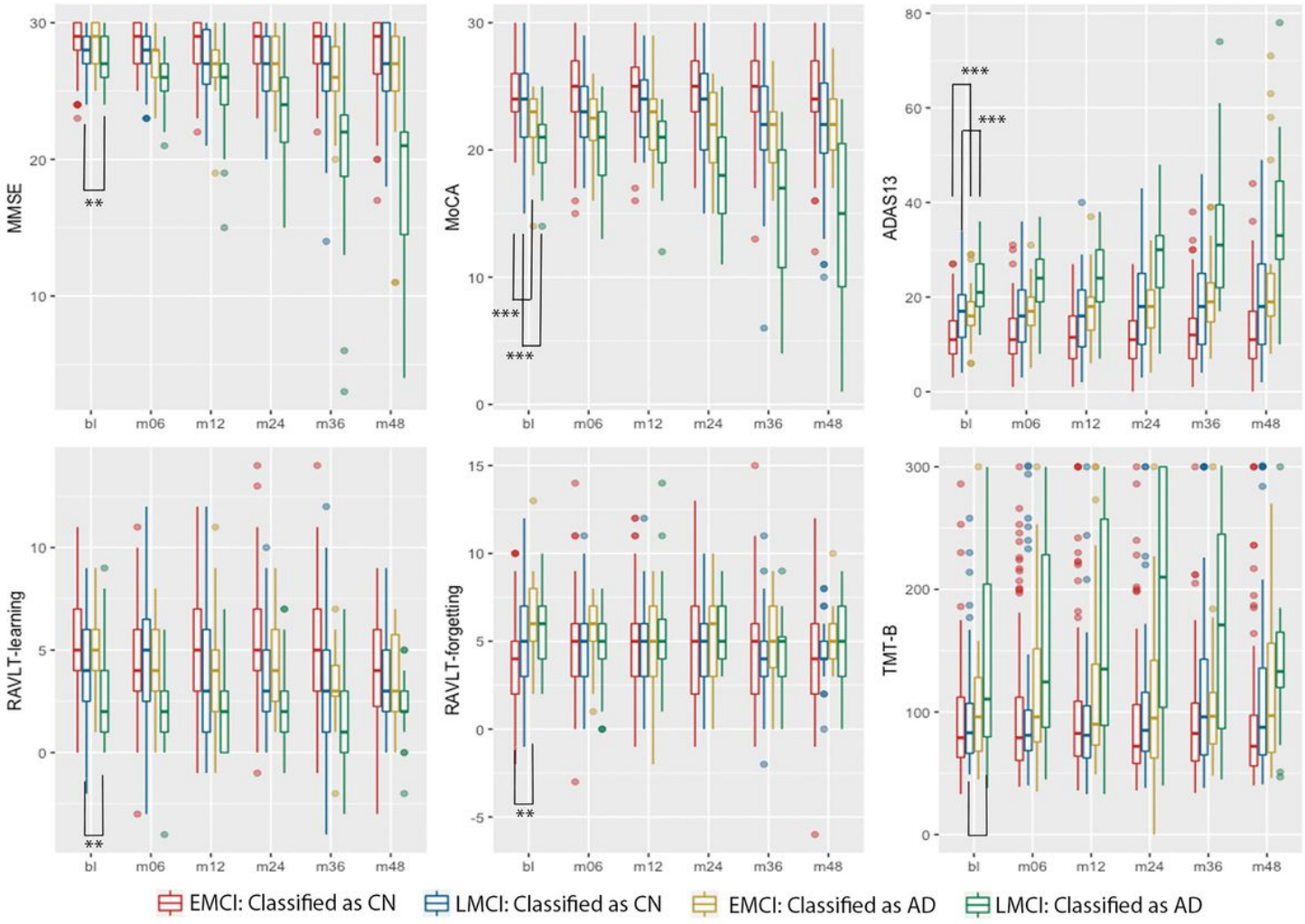

Figure 4

Longitudinal changes in classified MCI Boxplot of cognitive performance at each visit for MCI. MMSE, mini-mental state examination; MoCA, Montreal cognitive assessment; ADAS: Alzheimer's Disease Assessment Scale Cognitive Subscale; RAVLT, Rey's auditory verbal learning test; TMT-B, trail making test-B; EMCI, early mild cognitive impairment; LMCI, late mild cognitive impairment; CN, control; bl, baseline; m, month.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.docx