

Predictions of Current and Future Episodic Memory Using Grey Matter Volume and Functional Connectome: A Longitudinal Study in Amnesic Mild Cognitive Impairment Patients

Zan Wang

Affiliated Zhongda Hospital of Southeast University

Hao Shu

Southeast University Zhongda Hospital

Duan Liu

Southeast University Zhongda Hospital

Fan Su

Southeast University Zhongda Hospital

Chunming Xie

Southeast University Zhongda Hospital

Zaixu Cui (✉ zaixucui@gmail.com)

<https://orcid.org/0000-0003-4385-8106>

Zhijun Zhang (✉ janemengzhang@vip.163.com)

Southeast University Zhongda Hospital

Research

Keywords: Alzheimer's disease, amnesic mild cognitive impairment, episodic memory, machine learning, magnetic resonance imaging, relevance vector regression

DOI: <https://doi.org/10.21203/rs.3.rs-31020/v2>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Amnesic mild cognitive impairment (aMCI) patients are considered an at-risk group for progression to Alzheimer's dementia and accurate prediction of aMCI progression could facilitate the optimal decision-making for both clinicians and patients. Based on the baseline whole-brain grey-matter volume (GMV) and resting-state functional connectivity (FC), we used relevance vector regression to predict the baseline and longitudinal Rey's Auditory Verbal Learning Test Delayed Recall (AVLT-DR) scores of individual aMCI patients.

Methods: Fifty aMCI patients completed baseline and 3-year follow-up visits. All patients underwent comprehensive neuropsychological assessments and multimodal brain MRI scans.

Results: We found that the GMV pattern predicted the baseline AVLT-DR score, while the pattern of FC predicted the longitudinal AVLT-DR score. In particular, GMV predicted the baseline AVLT-DR score with an accuracy of $r = 0.54$ ($P < 0.001$); the regions that contributed the most were within the default mode (e.g., the posterior cingulate gyrus, angular gyrus and middle temporal gyrus) and limbic systems (e.g., the hippocampus and parahippocampal gyrus). The FC predicted the longitudinal AVLT-DR score with an accuracy of $r = 0.50$ ($P < 0.001$), and the connections that contributed the most were the within- and between-system connectivity of the default mode and limbic systems. As a complement, we demonstrated that the GMV and FC patterns could also effectively predict the baseline and longitudinal composite episodic memory scores (calculated by averaging three well-known episodic memory test scores).

Conclusions: Our results demonstrated the multimodal brain features in the individualized prediction of aMCI patients' current and future episodic memory performance. These "neural fingerprints" have the potential to be biomarkers for aMCI patients and can help medical professionals optimize individual patient management and longitudinal evaluation.

Introduction

Alzheimer's disease (AD), a severe neurodegenerative disorder, is characterized by memory loss and a reduction in cognitive function due to the progressive impairment of neurons and their connections [1]. Amnesic mild cognitive impairment (aMCI) is associated with an increased risk of progression to a diagnosis of probable AD [2, 3]. However, rates of aMCI progression vary: some individuals with aMCI deteriorate rapidly, others remain stable for many years, and some revert to having a normal cognitive status. Thus, as part of efforts to better understand the disease, important concepts such as the way aMCI progresses and the identification of related pathological biomarkers for progression have recently received increasing attention.

Neuroimaging measures, in particular magnetic resonance imaging (MRI), are widely available, non-invasive, and exceedingly promising in the study of AD and its putative prodromal states (e.g., aMCI). Alterations of neuroimaging measures in grey-matter (GM), such as hippocampal and default mode

regions, and in white-matter (WM), such as the integrity of the uncinated fasciculus and parahippocampal cingulum and network connectivity breakdown, have been linked to AD/aMCI [4–6]. However, until now, there has been no consensus on neuroimaging-based biomarkers that can accurately reflect aMCI progression, and no studies have used a prediction/validation study design to establish the generalizability of such markers.

In an effort to increase the translational applicability of neuroimaging results, there has been a recent shift towards the use of multivariate machine learning techniques [7]. Recently, a number of pattern regression methods have been used to estimate AD-related clinical variables, e.g., the Mini-Mental State Examination (MMSE), based on neuroimaging data [8–10]. These studies demonstrated that integrating pattern regression methods and neuroimaging biomarkers might be extremely valuable for accurately estimating cognitive scores and helpful for tracking disease progression. However, notably, episodic memory deficit is a core symptom of aMCI [11]; therefore, it is of utmost importance to determine aMCI-relevant cognitive scores (e.g., Rey's Auditory Verbal Learning Test [AVLT] memory score [12], a well-known measure of episodic memory) and then to identify neuroimaging biomarkers associated with these scores so that essential pathways from brain structure or function to cognition can potentially be discovered. Furthermore, nearly all existing regression methods developed for the estimation of cognitive test scores have been based only on one imaging modality [9, 13, 14]; however, the neuroimaging biomarkers from different modalities could provide complementary information, e.g., structural brain atrophy patterns measured by structural MRI and network connectivity measured by resting functional MRI.

Inspired by the above problems, in this paper, we used a multivariate relevance vector regression (RVR) method [15], which has been found to be an effective regression method yielding robust estimation of continuous clinical variables with reasonable diagnostic accuracy and good generalization ability [9, 13, 16], and a longitudinal design to investigate the potential for neuroimaging to predict baseline and 3-year longitudinal episodic memory performance for individual aMCI patients. We reasoned that the episodic memory performance with the greatest predictive accuracy overall would closely reflect whole-brain structural changes and network connectivity breakdown associated with AD.

Materials And Methods

Participants

Patients with aMCI were recruited via a registry at the Affiliated ZhongDa Hospital of Southeast University [5, 17]. From this registry, 87 aMCI patients were included; all of them were Chinese Han and right-handed. All patients were recruited through a normal community health screening and newspaper advertisements, and they underwent a standardized clinical interview, a neuropsychological battery assessment, and multimodal brain MRI examinations. The detailed inclusion and exclusion criteria are described in our previous publications [4, 5]. The Research Ethics Committee of the Affiliated ZhongDa Hospital of

Southeast University approved this study, and written informed consent was obtained from all participants.

In the present study, we selected the data of 50 aMCI patients who completed visits at two time points (T1 and T2) with 3 years between them. Notably, at follow-up (i.e., the T2 time point, 3 years later), both the clinical/behaviour assessment and parameters of MRI scanning were identical to those conducted at baseline (i.e., the T1 time point). All aMCI patients included in this study had no excessive motion artifacts (i.e., exceeding 3 mm in translational movement or 3° in rotational movement) during MRI scanning or incomplete image coverage. Table 1 presents the aMCI patients' demographic information and neuropsychological performances covering the episodic memory domain.

Table 1
Demographic and clinical information.

	aMCI (N = 50)	
	Baseline	3-year Follow-up
Age (years)	68.0 ± 7.3	70.1 ± 7.3
Education (years)	11.9 ± 3.4	-
Gender (male/female)	30/20	-
MMSE	27.08 ± 2.03	26.30 ± 3.22
MDRS-2	133.74 ± 5.76	125.40 ± 20.65
AVLT-immediate recall	15.7 ± 3.5	14.4 ± 4.4
AVLT-recognition	19.6 ± 2.4	18.8 ± 4.2
Episodic Memory	6.53 ± 2.75	6.40 ± 3.35
AVLT-delayed recall	2.78 ± 1.54	2.46 ± 2.38
LMT-delayed recall	3.34 ± 2.26	2.60 ± 2.02
CFT-delayed recall	13.47 ± 6.03	14.14 ± 6.64
Data are presented as the mean ± standard deviation.		
Abbreviations: aMCI, amnesic mild cognitive impairment; MMSE, mini-mental state examination; MDRS-2, Mattis dementia rating scale-2; AVLT, auditory verbal learning test. LMT, logical memory test; CFT, Rey-Osterrieth complex figure test.		

Neuropsychological Examination

For all aMCI patients, we assessed their general cognitive function using the MMSE and Mattis Dementia Rating Scale-2 and administered a neuropsychological battery to evaluate their episodic memory

performance. This battery consisted of the AVLT-DR, the logical memory test with a 20 min delayed recall (LMT-DR), and the Rey-Osterrieth complex figure test with a 20 min delayed recall (CFT-DR). As in previous studies [18], a composite episodic memory score was also calculated by averaging the AVLT-DR, LMT-DR and CFT-DR scores for all patients.

Data Acquisition

MRI images were acquired in a 3.0 T Siemens Verio scanner with a 12-channel head coil at the Affiliated ZhongDa Hospital of Southeast University. All patients underwent both the high-resolution T1-weighted and resting-state functional MRI scanning. Participants lay supine with the head snugly fixed by a belt and foam pads to minimize head movement. High-resolution T1-weighted axial images covering the whole brain were acquired using a 3D magnetization prepared rapid gradient echo (MPRAGE) sequence as follows: repetition time (TR) = 1900 ms; echo time (TE) = 2.48 ms; flip angle (FA) = 9°; acquired matrix = 256 × 256; field of view (FOV) = 250 mm × 250 mm; thickness = 1.0 mm; gap = 0 mm; and number of slices = 176. Resting-state functional images were obtained for eight minutes with gradient-recalled echo-planar imaging (GRE-EPI) sequence: TR = 2000 ms; TE = 25 ms; FA = 90°; acquisition matrix = 64 × 64; FOV = 240 mm × 240 mm; thickness = 4.0 mm; gap = 0 mm; and number of slices = 36. Prior to the scan, all patients were instructed to keep their eyes closed, stay awake, relax their minds, and move as little as possible during data acquisition.

Data Processing

Structural MRI Processing

The VBM8 (<http://dbm.neuro.uni-jena.de/vbm/>) toolbox was used to perform the analysis of brain structural imaging. In this process, all images were spatially normalized using combinations of affine linear transform and nonlinear registration to the standard Montreal Neurological Institute (MNI) template and segmented into GM, WM, and cerebrospinal fluid (CSF). Segmented GM images were modulated to compensate for the volumetric effects of expansion or shrinking employed in spatial normalization by multiplying the voxel intensity with the Jacobian determinants reflecting the parameters for fitting a voxel in native space to the corresponding voxel in template space. The modulated images were then smoothed with a 10-mm full width half maximum (FWHM) isotropic Gaussian kernel and resampled to 10 mm isotropic voxels. These procedures created a whole-brain voxel-based GM volume (GMV) map for each individual. Then, we calculated the average GMV of each region in the Automated Anatomical Labelling (AAL) atlas, which includes 90 prior cortical and subcortical regions in total. These 90 regional average GMV values were used as a feature vector to perform the following prediction analyses.

Functional MRI Processing

Data pre-processing. Resting-state functional data were pre-processed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) and the DPARSF ([http://www.restfmri.net/ forum/dparsf](http://www.restfmri.net/forum/dparsf)) toolboxes.

The first ten functional volumes were discarded to minimize the effects of scanner stabilization and participant adaptation. The remaining images were corrected for timing differences and motion effects. The individual structural images were segmented into GM, WM and CSF using a unified segmentation algorithm. Using the transformation parameter estimate during unified segmentation, the motion-corrected functional volumes were spatially normalized to MNI space and resampled to 3 mm isotropic voxels. Further pre-processing included linear detrending and temporal bandpass filtering (0.01–0.1 Hz), which were applied to reduce the effects of low-frequency drift and high-frequency physiological noise. We regressed out several spurious effects of nuisance covariates, including six head motion parameters, mean global signal, WM signal and CSF signal.

Whole-brain resting-state functional connectivity (FC) analyses. To compute resting-state FC, the AAL atlas was applied to parcellate the entire GM into 90 cortical and subcortical regions. For each subject, a regional mean time series was calculated by averaging the time series over all voxels within this region, and a total of 90 regional mean time series were therefore yielded. The resting-state FC between each pair of regions was computed by using the Pearson correlation coefficient. We obtained one symmetric correlation matrix (i.e., 90×90) for each aMCI patient. Then, Fisher's z-transform was applied to improve the normality of the correlation coefficients. Finally, for each patient, we converted the FC matrix into a feature vector with 4,005 values.

Individualized Prediction of Current and Future Episodic Memory Performance by MRI-derived Features

Based on the MRI data at T1 time point, we applied multivariate RVR to predict both the current (i.e., T1 time point) and future (i.e., T2 time point, 3 years later) episodic memory performances (i.e., AVLT-DR and composite episodic memory scores) of unseen aMCI individuals. Particularly, both whole-brain GMV acquired from structural MRI and whole-brain FC acquired from functional MRI were used separately. We were interested in whether the structure and function performed equally well in the prediction of both current and future episodic memory performance.

RVR is performed in a probabilistic Bayesian learning framework and obtains sparse solutions of a multivariate regression model [15]. Under this framework, an explicit zero-mean Gaussian prior was applied to the model weights, and therefore, most weights were set to zero, resulting in only some samples, termed the 'relevance vector', being used to train the model. The maximum likelihood estimation was used to find the weights of these samples. The regression coefficients of all features were determined as the weighted sum of the feature vector of all "relevance vector" samples. This algorithm does not have an algorithm-specific free parameter and is computationally more efficient than other algorithms [19]. RVR has been widely used to predict age [20] and behaviours [21]. We used the codes from the PRoNTTo toolbox (<http://www.mlnl.cs.ucl.ac.uk/pronto/>) to implement RVR. All codes about prediction are publicly available on GitHub (https://github.com/ZaixuCui/Pattern_Regression_Clean).

Prediction Framework

To quantify prediction accuracy, we applied leave-one-out cross-validation (LOOCV) to estimate the out-of-sample generalizability of the models. Specifically, $N-1$ subjects (where N is the number of aMCI patients) were used as the training set, with the remaining individual used as the testing sample. During the training procedure, each feature was linearly scaled to a range of zero to one across the training set [19, 22], and then an RVR prediction model was constructed using this training set. During the testing procedure, each testing subject feature vector was scaled using the scaling parameter acquired during the training procedure. The training and testing procedures were repeated N times so that each patient was used once as the testing sample. The correlation coefficient r and mean absolute error (MAE) between the predicted and actual scores were used to quantify the prediction accuracy [19]. In the present study, we controlled for age, sex and years of education when calculating the correlation between predicted and actual scores.

Significance of Prediction Performance

Permutation tests were used to determine whether the coefficient r and MAE were significantly better than the results expected by chance. Specifically, the above prediction procedure was re-applied 1,000 times. For each time, we permuted the behavioural scores across the training samples without replacement. The P value of the mean correlation r was calculated by dividing the number of permutations that showed a higher value than the actual value for the real sample by the total number of permutations (i.e., 1,000). Similarly, the P value of the mean MAE was the portion of permutations that showed a lower value than the actual value for the real sample. To assess the specificity of the predictive models for current and future episodic memory performances in aMCI patients, we examined the correlation between the predicted AVLT-DR scores and their actual scores after adjusting for the effects of age, sex, and years of education.

Contributing Features and Corresponding Weights

To quantify the contribution of each feature to prediction, we constructed a new RVR model using all subjects. The absolute value of the RVR weight of each feature quantified its contribution to the model. A larger absolute value of weight indicated a greater contribution of the corresponding feature to the prediction in the context of all other features. For GMV predictions that included 90 features in total, the features were selected for visualization if the absolute value of their weight was in the top 10%. Given the large number of FC features (i.e., 4,005 features), we displayed the features with the highest absolute contribution weight in the top 1%. These thresholds, although arbitrary, eliminated noise components and enabled a better visualization of the most predictive features.

To better interpret our results, we used the standard 7-system template image provided by Yeo et al. [23] that was originally derived from a whole-brain clustering analysis, which yielded 7 large-scale functional networks. To define the priori network modules, network nodes obtained from the AAL atlas were each assigned to one of the 7 large-scale functional modules; subcortical nodes were assigned to an eighth subcortical module. Therefore, the primary modular partition defined by 90-node networks consisted of 8

brain networks: default mode, frontoparietal, ventral attention, dorsal attention, visual, sensorimotor, limbic, and subcortical systems.

Validation

We conducted two additional analyses to validate our results. First, a 10-fold cross-validation was applied to validate our prediction results. Specifically, all patients were divided into 10 subsets, of which nine were used as the training set, and the remaining one was used as the testing set. The training set was scaled and used to train an RVR prediction model, which was then used to predict the scores for the scaled testing data. The scaling of testing data used parameters acquired from training data. This procedure was repeated 10 times so that each subset was used as the testing set once. Finally, the correlation r and MAE between the actual and predicted scores were calculated across all patients. Since the full dataset was randomly divided into 10 subsets, performance might have depended on data division. Therefore, the 10-fold cross-validation was repeated 100 times, and the results were averaged to produce the final prediction performance. A permutation test was applied 1,000 times to test the significance of the prediction performance. Second, we further put T1 GMV and FC features into the RVR separately to predict the T1 and T2 composite episodic memory scores. Putatively, the predictive model would capture specific GMV and FC features for individualized predictions of current and future episodic performances in aMCI patients if the resultant correlation remained significant.

Results

Whole-Brain GMV Predicted Current Episodic Memory in aMCI Patients

Evaluated by LOOCV, the pattern of whole-brain GMV predicted the AVLT-DR scores at T1 time point (Fig. 1A, left panel). The partial correlation between the actual and predicted scores was $r = 0.54$ (permutation test, $P < 0.001$), controlling for age, sex and years of education. The regions that contributed the most to the prediction (i.e., the top 10% brain regions) included regions mainly located in the default mode (e.g., the bilateral posterior cingulate gyrus, left angular gyrus and middle temporal gyrus), limbic (e.g., the bilateral hippocampus and right parahippocampal gyrus), and dorsal attention (e.g., the left superior parietal gyrus and right inferior temporal gyrus) systems (Fig. 1A, right panel). In contrast, whole-brain GMV failed to predict the AVLT-DR scores at T2 time point (partial $r = -0.027$, $P_{perm} = 0.390$; Fig. 1B).

Table 2
Cortical and subcortical regions of interest (ROIs) included in AAL atlas.

Regions	Abbreviation	Regions	Abbreviation
Precentral gyrus	PreCG	Lingual gyrus	LING
Superior frontal gyrus, dorsolateral	SFGdor	Superior occipital gyrus	SOG
Superior frontal gyrus, orbital part	ORBsup	Middle occipital gyrus	MOG
Middle frontal gyrus	MFG	Inferior occipital gyrus	IOG
Middle frontal gyrus, orbital part	ORBmid	Fusiform gyrus	FFG
Inferior frontal gyrus, opercular part	IFGoperc	Postcentral gyrus	PoCG
Inferior frontal gyrus, triangular part	IFGtriang	Superior parietal gyrus	SPG
Inferior frontal gyrus, orbital part	ORBinf	Inferior parietal, but supramarginal and angular gyri	IPL
Rolandic operculum	ROL	Supramarginal gyrus	SMG
Supplemental motor area	SMA	Angular gyrus	ANG
Olfactory cortex	OLF	Precuneus	PCUN
Superior frontal gyrus, medial	SFGmed	Paracentral lobule	PCL
Superior frontal gyrus, medial orbital	ORBsupmed	Caudate nucleus	CAU
Gyrus rectus	REC	Lenticular nucleus, putamen	PUT
Insula	INS	Lenticular nucleus, pallidum	PAL
Anterior cingulate and paracingulate gyri	ACG	Thalamus	THA
Median cingulate and paracingulate gyri	DCG	Heschl gyrus	HES
Posterior cingulate gyrus	PCG	Superior temporal gyrus	STG

Regions	Abbreviation	Regions	Abbreviation
Hippocampus	HIP	Temporal pole: superior temporal gyrus	TPOsup
Parahippocampal gyrus	PHG	Middle temporal gyrus	MTG
Amygdala	AMYG	Temporal pole: middle temporal gyrus	TPOmid
Calcarine fissure and surrounding cortex	CAL	Inferior temporal gyrus	ITG
Cuneus	CUN		

Whole-Brain FC Predicted Future Episodic Memory in aMCI Patients

Evaluated by LOOCV, we found that the pattern of whole-brain FC failed to predict the AVLT-DR scores at T1 time point (partial $r = 0.17$, $P_{perm} = 0.082$; Fig. 2A). Interestingly, the FCs significantly predicted the AVLT-DR scores at T2 time point (Fig. 2B), with the partial correlation between actual and predicted scores being $r = 0.50$ (permutation test, $P < 0.001$). The connectivity that contributed the most to the prediction (i.e., the top 1% connectivity) included within-default mode connections, within-limbic connections and the connections between default mode and limbic systems (Fig. 2C). In particular, these connections with the highest absolute contribution weight mainly displayed long anatomical distances (i.e., Euclidean distance > 75 mm) (Fig. 2D).

Validation Results

Validation Using 10-Fold Cross-Validation

LOOCV was used to evaluate the performance of the model in the main results. Here, we applied 10-fold cross-validation to validate our results. Consistent with our main results, the pattern of whole-brain GMV significantly predicted the AVLT-DR scores at T1 time point (partial $r = 0.64$, $P_{perm} < 0.001$; MAE = 1.03, $P_{perm} < 0.001$; Fig. 3A) but failed to predict the AVLT-DR scores at T2 time point (partial $r = 0.33$, $P_{perm} = 0.019$; MAE = 2.00, $P_{perm} = 0.072$; Fig. 3B). In contrast, the pattern of whole-brain FC predicted the AVLT-DR scores at T1 time point (partial $r = 0.30$, $P_{perm} = 0.028$; MAE = 1.18, $P_{perm} = 0.004$; Fig. 3C) but more significantly predicted the AVLT-DR scores at T2 time point (partial $r = 0.59$, $P_{perm} < 0.001$; MAE = 1.61, $P_{perm} < 0.001$; Fig. 3D).

Validation Using the Composite Episodic Memory Scores

To validate the robustness of our results, we used the average AVLT-DR, LMT-DR and CFT-DR scores as the composite episodic memory performance. The results suggested that the pattern of whole-brain GMV significantly predicted the composite episodic memory scores at T1 time point (partial $r = 0.41$, $P_{perm} < 0.001$; Fig. 4A, left panel) but less accurately predicted the composite episodic memory scores at T2 time point (partial $r = 0.27$, $P_{perm} = 0.006$; Fig. 4B). The most important GMV features contributing to T1 composite episodic memory score prediction involved multiple regions within the default mode (e.g., the right precuneus, left angular gyrus and middle temporal gyrus), limbic systems (e.g., the bilateral hippocampus and amygdala, and right parahippocampal gyrus) (Fig. 4A, right panel). The pattern of whole-brain FC significantly predicted the composite episodic memory scores at T2 time point (partial $r = 0.37$, $P_{perm} = 0.002$; Fig. 4C, left panel) but failed to predict the composite episodic memory scores at T1 time point (partial $r = 0.18$, $P_{perm} = 0.071$; Fig. 4D); the top 1% of FCs for predicting T2 composite episodic memory scores mainly involved connections linking different functional modules (i.e., the default mode, limbic and subcortical systems) and connections within single modules (i.e., the default mode and limbic systems) (Fig. 4C, right panel).

Discussion

The early detection of probable dementia in aMCI patients and/or the prediction of the progression of aMCI to probable dementia are critical issues facing clinicians. Since the neurodegeneration of AD proceeds years before the onset of the disease and since therapeutic intervention is more effective at the early stage of the disease (e.g., the aMCI stage), there is an urgent need to (i) accurately predict the progression of the disease measured by cognitive scores, e.g., AVLT-DR, that would be more focused on functions related to aMCI, and (ii) identify a small set of biomarkers most predictive of the progression. In the current study, we employed baseline MRI-derived features (e.g., whole-brain GMV and network connectivity) in a machine learning framework to make continuously valued predictions on current and future episodic memory performance for aMCI patients at the individual level. We demonstrated that the whole-brain GMV feature-based multivariate models identified a wide range of important GM regions that better contributed to the prediction of the baseline AVLT-DR scores; these regions mainly covered the default mode and limbic systems. In contrast, the whole-brain FC feature-based prediction models identified a wide range of network connections that better contributed to the prediction of 3-year longitudinal AVLT-DR scores; these FCs mainly involved connections within and between the default mode and limbic systems. Overall, the presence of a relationship between baseline MRI features and current/future episodic memory performance indicates the ability to predict cognitive status, which would benefit physicians in the management of aMCI patients.

To identify neuroimaging-based biomarkers for predicting cognitive scores, it has been recently advocated to push the traditional correlational analysis across all samples to the individualized prediction that naturally evaluates whether the identified neuroimaging markers can be generalized and used in practice [24, 25]. Given that patients with aMCI are at high risk for conversion to AD, it is of great clinical interest to be able to identify a set of biomarkers most predictive of aMCI progression. However,

until now, there have been limited studies on the predictive ability of neuroimaging-based biomarkers for aMCI progression, and the accuracy of the prediction of cognitive change (i.e., MMSE changes) has been very modest [8]. Therefore, to improve the model's predictive accuracy, we employed the AVLT-DR test to evaluate aMCI patients' episodic memory deficit, which is known to be a core symptom of aMCI, and used a multimodal RVR method and longitudinal design to investigate the potential for neuroimaging to predict baseline and 3-year longitudinal AVLT-DR scores for individual aMCI patients.

In this study, we found that the baseline AVLT-DR scores were better predicted by GMV features in AD-related "hub" regions, e.g., the medial temporal lobe and default mode regions; however, the 3-year longitudinal AVLT-DR scores were better predicted by FC features within and between the limbic and default mode systems. This finding signifies the ability of MRI to detect morphological and functional changes that predict current and future episodic memory performance, putatively pointing to a pathological cause. Intriguingly, these findings correlate with the pathological process of AD. It is striking that early A β deposition occurs in a stereotypic set of heteromodal cortical regions, largely overlapping the neocortical regions of the default mode network (DMN) that are among the strongest "connection hubs", whereas tau accumulation begins in deep grey matter structures, prominently in the transentorhinal cortex and related structures in the medial temporal lobe that are functionally connected to the DMN [26].

Regarding AD, increasing evidence suggests that pathology may begin within key vulnerable "hubs", e.g., the medial temporal lobe and default mode regions, defined as central regions within the AD pathology-targeted network architecture [27]. These regions represent specific network "epicenters" whose connectivity serve as a template for the spatial patterning of disease [28]. Importantly, previous network findings in AD provide the strongest support for the "transneuronal spread model", in which AD pathology propagates along target network connections once the disease has spread throughout the "hub" regions [29, 30]. In particular, task-free fMRI studies have identified the consistent involvement of the limbic network and DMN in AD and aMCI [5, 31]. These networks show good topographic overlap with areas affected by pathology and neurodegeneration, that is, the hippocampus and parahippocampal gyrus, posterior cingulate and precuneus, and medial prefrontal cortex. Atrophy of these areas and aberrant functional connectivity within/between these networks correlated with disease severity [32, 33] and cognitive-behavioural scores [34]. Structural connectivity changes also showed good agreement with limbic system and DMN topography in AD patients; for example, significantly reduced fractional anisotropy and increased mean diffusivity values were observed in limbic white matter tracts connecting the medial temporal lobe with the posterior cingulate [35–37]. Moreover, these findings provided consistent evidence that limbic system or DMN damage corresponds to disease progression, as measured by cross-sectional studies at different disease stages [32, 34, 38] or longitudinal studies [39]. Therefore, our findings provide further evidence supporting the "transneuronal spread model" in AD and demonstrate that the integration of neuroimaging biomarkers from different modalities is more beneficial for revealing aMCI progression patterns and is helpful for accurately reflecting the underlying structural and functional changes associated with aMCI progression.

We finally sought to ascertain whether the whole-brain GMV and FC patterns could effectively predict the baseline and longitudinal composite episodic memory scores (i.e., calculated by averaging the AVLT-DR, LMT-DR and CFT-DR scores) for individual aMCI patients. In particular, our main findings were reproducible; the baseline composite episodic memory scores were better predicted by GMV features in AD-related “hub” regions (e.g., medial temporal lobe and default mode regions), and the longitudinal composite episodic memory scores were better predicted by FC features within and between the limbic, subcortical and default mode networks, supporting the models’ specificity for individualized predictions of aMCI patients’ current and future episodic memory performance. Additionally, studies of normal memory processes have indicated that subject demographics, especially age, have considerable effects on the AVLT cognitive test results in cognitively normal individuals [40], and at the same time, ageing changes the brain’s structure and function [41]. Intriguingly, the accuracies of estimated AVLT-DR scores and composite episodic memory scores were still statistically significant by run-wise applied permutations and by adding age-correction procedures to the predictive models, further supporting the models’ robustness and generalizability. Therefore, the combination of whole-brain GMV and FC feature-based models possesses potential clinical significance to predict aMCI progression at the individual level.

Several limitations should be considered when generalizing and extending the results in this study. First, our proposed method performs prediction based on the longitudinal and multimodality data and thus requires each aMCI patient to have the corresponding modality data and episodic memory scores (i.e., the AVLT-DR, LMT-DR and CFT-DR tests), which limits the size of the aMCI patient population that can be used for study. Therefore, generalization of the current findings requires further validation using a larger independent sample and other cross-validation methods. Second, recent studies have indicated that biomarkers from different modalities (e.g., blood, CSF, structural and functional MRI) provide complementary information and that genetic information could predict changes in cognitive examination performance. Therefore, further investigations are encouraged to achieve a better and more reliable prediction performance by combining MRI-derived features and other modality data (e.g., APOE genotype, AD pathology derived from blood and CSF). Finally, the study lacked amyloid imaging and histopathological data suggestive of AD pathology in aMCI patients and thus relied on clinical diagnosis only. Therefore, a considerable amount of clinical and biological heterogeneity existed in the present sample of aMCI subjects. Notably, the present study only recruited aMCI patients at baseline, who were considered to have a high risk of conversion to AD dementia. More importantly, as shown in our previous studies using the same dataset, aMCI patients showed significantly lower cortical thickness in the hippocampus and parahippocampal gyrus [4], lower plasma A β 40 and A β 42 levels [42], and higher plasma neurofilament light levels [42] than healthy controls. Therefore, the aMCI patients recruited in this study may have been relatively homogeneous.

Conclusion

In the present study, we carried out both baseline and longitudinal multimodal data analyses for the computer-aided prediction of aMCI progression. We demonstrated that baseline whole-brain GMV and FCs within the limbic and default mode systems effectively predict aMCI patients’ episodic memory

performance (i.e., both at baseline and 3-year follow-up) at the individual level. These findings revealed that individual differences in baseline GM and FCs within the limbic and default mode systems contribute to variability in current and future episodic memory performance. As such, these “neural fingerprints” may be appropriate biomarkers for aMCI patients to optimize individual patient management and longitudinal evaluation in a timely fashion.

Abbreviations

AAL: automated anatomical labelling; AD: Alzheimer’s disease; aMCI: Amnesic mild cognitive impairment; AVLT-DR: rey’s auditory verbal learning test delayed recall; CFT-DR: rey-osterrieth complex figure test with a 20 min delayed recall; DMN: default mode network; FC: functional connectivity; GMV: grey-matter volume; LMT-DR: logical memory test with a 20 min delayed recall; LOOCV: leave-one-out cross-validation; MAE: mean absolute error; MMSE: mini-mental state examination; RVR: relevance vector regression; WM: white-matter.

Declarations

Acknowledgements

We thank all the patients and volunteers for participating in this study.

Author Contributions

Z.W., Z.C. and Z.Z. conceived and designed the research. Z.W. performed the experiments. H.S., D.L., F.S. and L.G. collected the data. Z.C. and C.X. provided technical assistance. Z.W. wrote the manuscript. All authors critically evaluated and approved the manuscript.

Funding

This study was supported by the National Key Research and Development Plan of China (grant number: 2016YFC1306700), the Natural Science Foundation of China (grant number: 81801680, 81830040), the Science and Technology Program of Guangdong (grant number: 2018B030334001), the Program of Excellent Talents in Medical Science of Jiangsu Province (grant number: JCRCA2016006), and the Foundation of Jiangsu Commission of Health (grant number: Z2018023).

Availability of Data and Materials

The data supporting the conclusions of this article are available from the corresponding author upon request.

Ethics Approval and Consent to Participate

The Research Ethics Committee of the Affiliated ZhongDa Hospital of Southeast University approved this study, and written informed consent was obtained from all participants.

Consent for Publication

Not applicable.

Competing Interests

The authors declare that they have no competing interests.

References

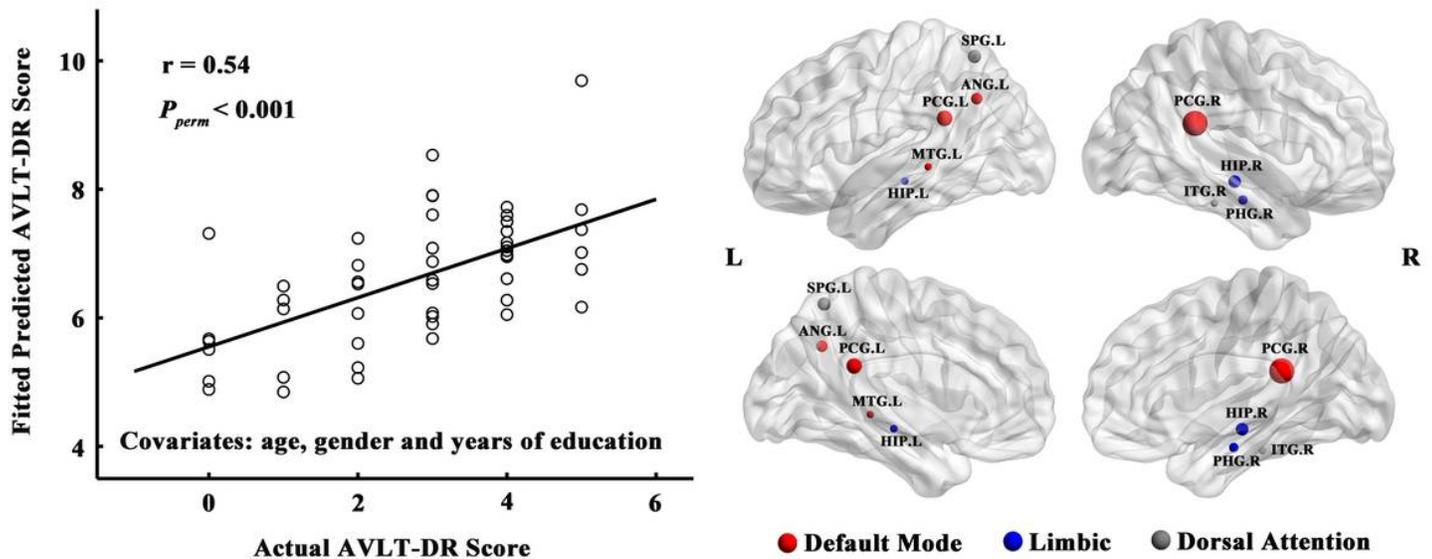
1. Khachaturian ZS Diagnosis of Alzheimer's disease. *Arch Neurol.* 1985; 42: 1097-1105.
2. Petersen RC, Bennett D Mild cognitive impairment: is it Alzheimer's disease or not? *J Alzheimers Dis.* 2005; 7: 241-245.
3. Dubois B, Albert ML Amnestic MCI or prodromal Alzheimer's disease? *Lancet Neurol.* 2004; 3: 246-248.
4. Wang Z, Dai Z, Shu H, et al. Cortical Thickness and Microstructural White Matter Changes Detect Amnestic Mild Cognitive Impairment. *J Alzheimers Dis.* 2017; 56: 415-428.
5. Wang Z, Dai Z, Shu H, et al. APOE Genotype Effects on Intrinsic Brain Network Connectivity in Patients with Amnestic Mild Cognitive Impairment. *Sci Rep.* 2017; 7: 397.
6. Bai F, Zhang Z, Watson DR, et al. Abnormal functional connectivity of hippocampus during episodic memory retrieval processing network in amnestic mild cognitive impairment. *Biol Psychiatry.* 2009; 65: 951-958.
7. Brammer M The role of neuroimaging in diagnosis and personalized medicine-current position and likely future directions. *Dialogues Clin Neurosci.* 2009; 11: 389.
8. Duchesne S, Caroli A, Geroldi C, et al. Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. *Neuroimage.* 2009; 47: 1363-1370.
9. Stonnington CM, Chu C, Kloppel S, et al. Predicting clinical scores from magnetic resonance scans in Alzheimer's disease. *Neuroimage.* 2010; 51: 1405-1413.
10. Huang L, Jin Y, Gao Y, et al. Longitudinal clinical score prediction in Alzheimer's disease with soft-split sparse regression based random forest. *Neurobiol Aging.* 2016; 46: 180-191.
11. Tromp D, Dufour A, Lithfous S, et al. Episodic memory in normal aging and Alzheimer disease: insights from imaging and behavioral studies. *Ageing Res Rev.* 2015; 24: 232-262.

12. Estévez-González A, Kulisevsky J, Boltes A, *et al.* Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging. *Int J Geriatr Psychiatry.* 2003; 18: 1021-1028.
13. Fan Y, Kaufer D, Shen D (2010) in *Proceedings of the 2010 IEEE international conference on Biomedical imaging: from nano to Macro.* IEEE Press. 852-855.
14. Moradi E, Hallikainen I, Hanninen T, *et al.* Rey's Auditory Verbal Learning Test scores can be predicted from whole brain MRI in Alzheimer's disease. *Neuroimage Clin.* 2017; 13: 415-427.
15. Tipping ME Sparse Bayesian learning and the relevance vector machine. *J Mach Learn Res.* 2001; 1: 211-244.
16. Wang Y, Fan Y, Bhatt P, *et al.* High-dimensional pattern regression using machine learning: from medical images to continuous clinical variables. *Neuroimage.* 2010; 50: 1519-1535.
17. Chen J, Shu H, Wang Z, *et al.* Convergent and divergent intranetwork and internetwork connectivity patterns in patients with remitted late-life depression and amnesic mild cognitive impairment. *Cortex.* 2016; 83: 194-211.
18. Doré V, Villemagne VL, Bourgeat P, *et al.* Cross-sectional and Longitudinal Analysis of the Relationship Between A β Deposition, Cortical Thickness, and Memory in Cognitively Unimpaired Individuals and in Alzheimer Disease. *JAMA Neurol.* 70: 903.
19. Cui Z, Gong G The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. *Neuroimage.* 2018; 178: 622-637.
20. Franke K, Ziegler G, Klöppel S, *et al.* Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. *Neuroimage.* 2010; 50: 883-892.
21. Feng C, Cui Z, Cheng D, *et al.* Individualized prediction of dispositional worry using white matter connectivity. *Psychol Med.* 2019; 49: 1999-2008.
22. Cui Z, Su M, Li L, *et al.* Individualized prediction of reading comprehension ability using gray matter volume. *Cereb Cortex.* 2018; 28: 1656-1672.
23. Thomas Yeo B, Krienen FM, Sepulcre J, *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol.* 2011; 106: 1125-1165.
24. Gabrieli JD, Ghosh SS, Whitfield-Gabrieli S Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. *Neuron.* 2015; 85: 11-26.
25. Dubois J, Adolphs R Building a science of individual differences from fMRI. *Trends Cogn Sci.* 2016; 20: 425-443.
26. Celone KA, Calhoun VD, Dickerson BC, *et al.* Alterations in memory networks in mild cognitive impairment and Alzheimer's disease: an independent component analysis. *J Neurosci.* 2006; 26: 10222-10231.

27. Buckner RL, Sepulcre J, Talukdar T, *et al.* Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease. *J Neurosci.* 2009; 29: 1860-1873.
28. Zhou J, Gennatas ED, Kramer JH, *et al.* Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron.* 2012; 73: 1216-1227.
29. Frost B, Diamond MI Prion-like mechanisms in neurodegenerative diseases. *Nat Rev Neurosci.* 2010; 11: 155-159.
30. Jucker M, Walker LC Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders. *Ann Neurol.* 2011; 70: 532-540.
31. Greicius MD, Srivastava G, Reiss AL, *et al.* Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc Natl Acad Sci.* 2004; 101: 4637-4642.
32. Petrella J, Sheldon F, Prince S, *et al.* Default mode network connectivity in stable vs progressive mild cognitive impairment. *Neurology.* 2011; 76: 511-517.
33. Brier MR, Thomas JB, Snyder AZ, *et al.* Loss of intranetwork and internetwork resting state functional connections with Alzheimer's disease progression. *J Neurosci.* 2012; 32: 8890-8899.
34. Binnewijzend MA, Schoonheim MM, Sanz-Arigita E, *et al.* Resting-state fMRI changes in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging.* 2012; 33: 2018-2028.
35. Pievani M, Agosta F, Pagani E, *et al.* Assessment of white matter tract damage in mild cognitive impairment and Alzheimer's disease. *Hum Brain Mapp.* 2010; 31: 1862-1875.
36. Acosta-Cabronero J, Alley S, Williams GB, *et al.* Diffusion tensor metrics as biomarkers in Alzheimer's disease. *PLoS One.* 2012; 7.
37. Douaud G, Menke RA, Gass A, *et al.* Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer's disease. *J Neurosci.* 2013; 33: 2147-2155.
38. Zhang H-Y, Wang S-J, Liu B, *et al.* Resting brain connectivity: changes during the progress of Alzheimer disease. *Radiology.* 2010; 256: 598-606.
39. Damoiseaux JS, Prater KE, Miller BL, *et al.* Functional connectivity tracks clinical deterioration in Alzheimer's disease. *Neurobiol Aging.* 2012; 33: 828. e819-828. e830.
40. Magalhães SS, Hamdan AC The Rey Auditory Verbal Learning Test: normative data for the Brazilian population and analysis of the influence of demographic variables. *Psychol Neurosci.* 2010; 3: 85-91.
41. Good CD, Johnsrude IS, Ashburner J, *et al.* A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage.* 2001; 14: 21-36.
42. Shi Y, Lu X, Zhang L, *et al.* Potential Value of Plasma Amyloid- β , Total Tau, and Neurofilament Light for Identification of Early Alzheimer's Disease. *ACS Chem Neurosci.* 2019; 10: 3479-3485.

Figures

A. T1 AVLT-DR Correlation: 0.54 (Permutation test, predicted by T1 whole-brain GMV)



B. T2 AVLT-DR Correlation: -0.027 (Permutation test, predicted by T1 whole-brain GMV)

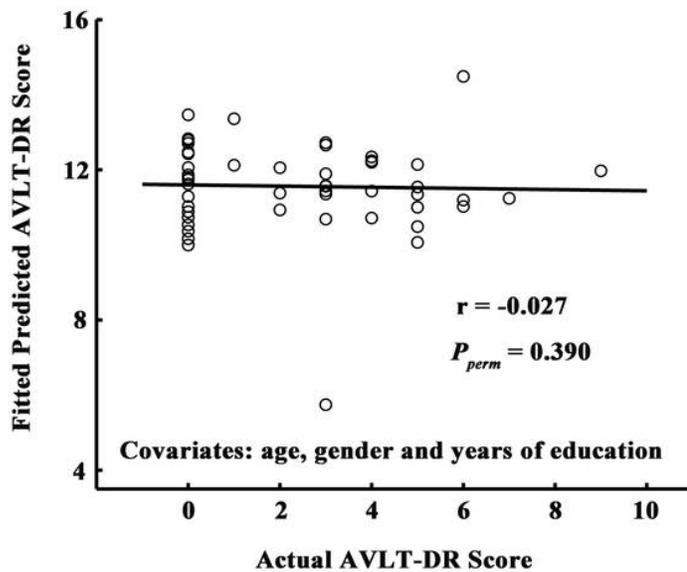


Figure 1

Prediction of the AVLT-DR score for each aMCI patient based on structural imaging. Evaluated by leave-one-out cross-validation (LOOCV), the pattern of T1 whole-brain GMV significantly predicted the T1 AVLT-DR scores (partial $r = 0.54$, $P < 0.001$; A, left panel) but failed to predict the T2 AVLT-DR scores (partial $r = -0.027$, $P = 0.390$; B). The most important T1 GMV features (A, right panel; i.e., the top 10% brain regions) for predicting the T1 AVLT-DR scores involved widespread regions that were mainly located in the default mode, limbic and dorsal attention systems. aMCI, amnesic mild cognitive impairment; AVLT-DR, Auditory Verbal Learning Test with a 20 min delayed recall; GMV, grey-matter volume; RVR, relevance vector

regression; T1, T1 time point (i.e., baseline); T2, T2 time point (i.e., 3-year follow-up). For the abbreviations of the brain regions, see Table 2.

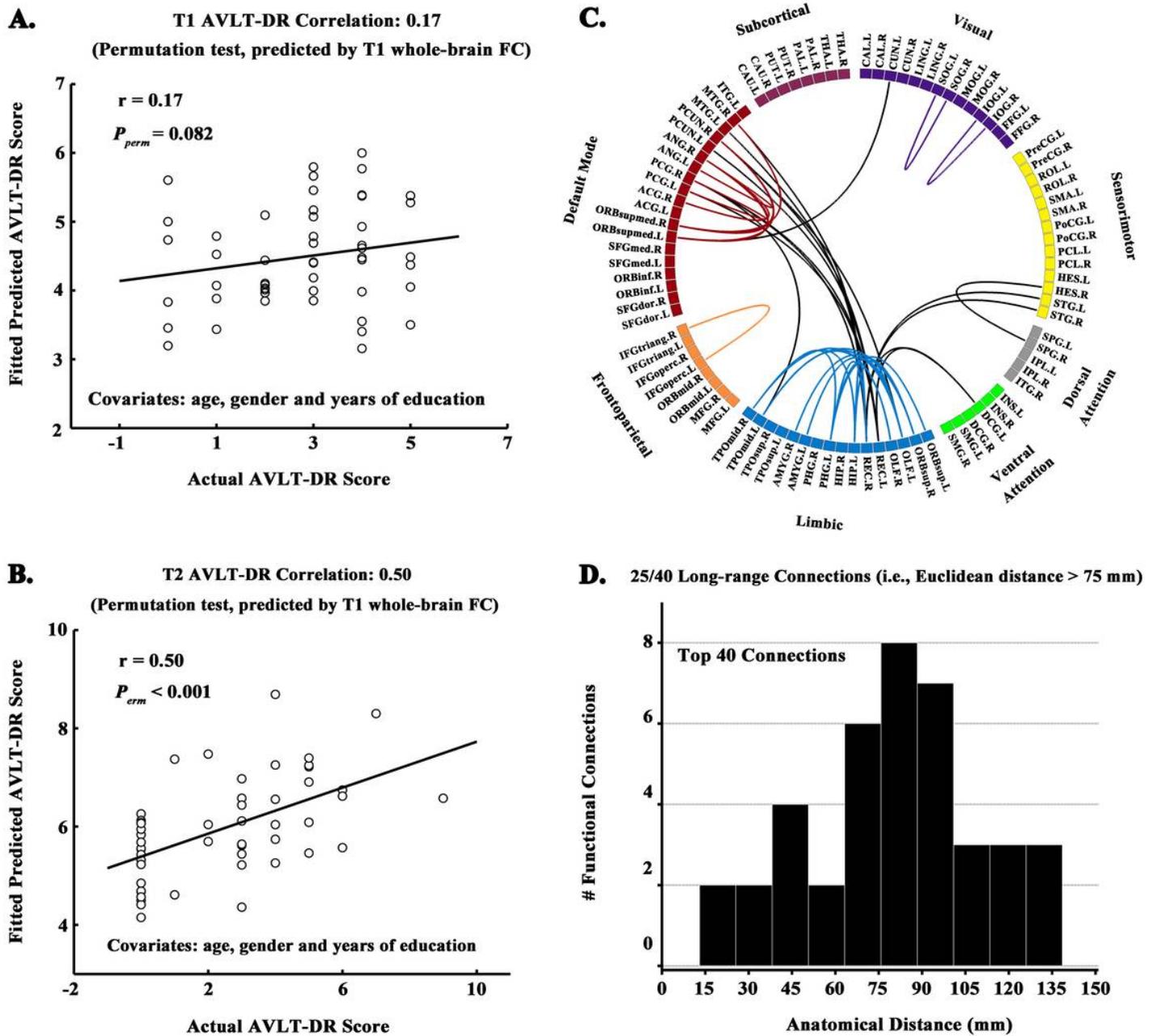


Figure 2

Prediction of AVLT-DR score for each aMCI patient derived from his/her T1 whole-brain FC. Evaluated by leave-one-out cross-validation (LOOCV), the pattern of T1 whole-brain FC failed to predict the T1 AVLT-DR scores (A; Partial $r = 0.17$, $P = 0.082$), but significantly predicted the T2 AVLT-DR scores (B; Partial $r = 0.50$, $P < 0.001$). (C) The connectivity that contributed the most to the T2 AVLT-DR scores prediction (i.e., the top 1% connectivity) included within-default mode connections, within-limbic connections and the connections between default mode and limbic systems. (D) These connections with the highest absolute contribution weight mainly displayed long anatomical distance (25/40 long-range connections). aMCI, amnesic mild cognitive impairment; AVLT-DR, auditory verbal learning test with a 20 min delayed recall;

FC, functional connectivity; T1, T1 time point (i.e., baseline); T2, T2 time point (i.e., 3-years follow-up). For the abbreviations of the brain regions, see Table 2.

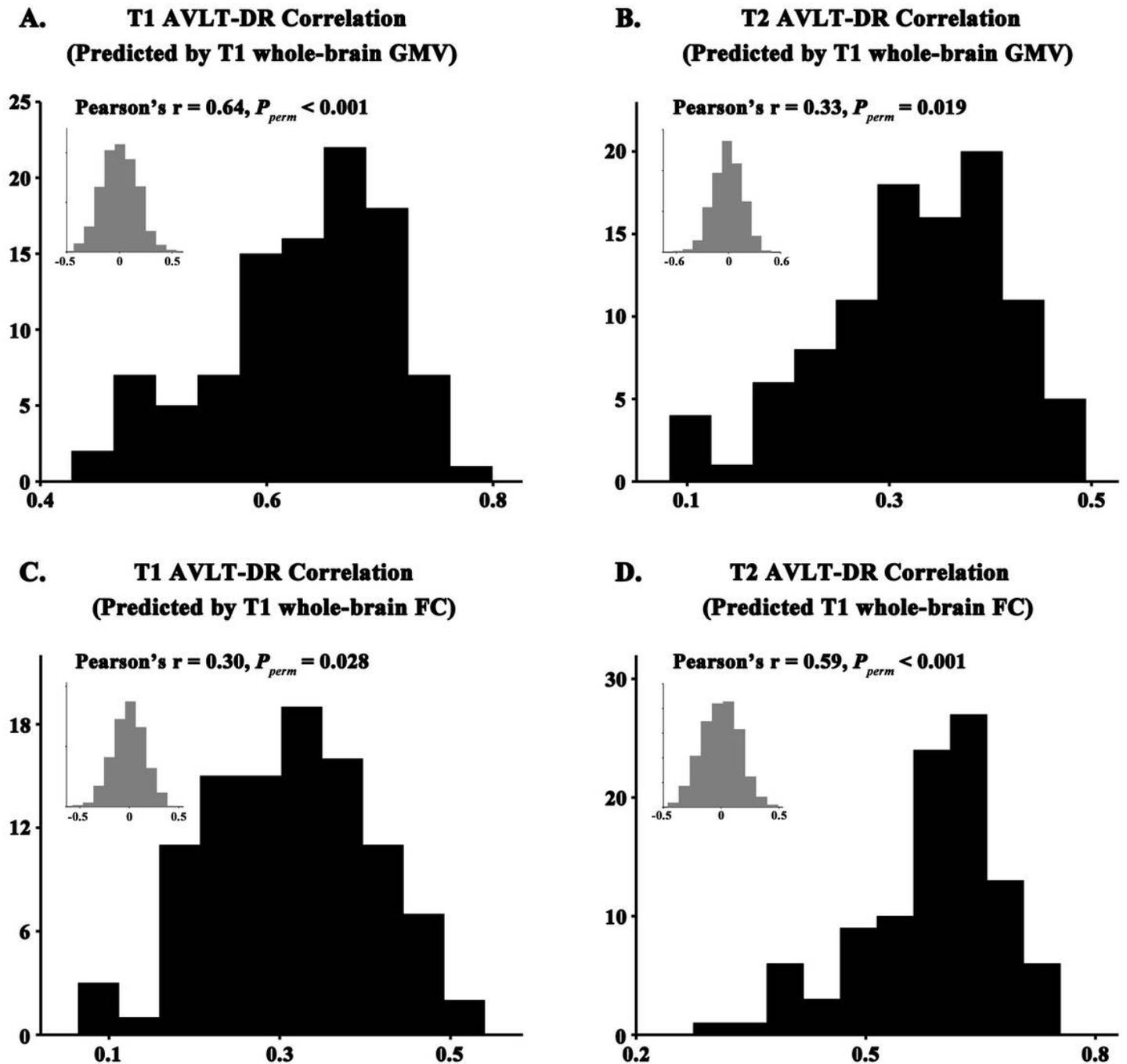


Figure 3

Validation using 10-fold cross-validation. The pattern of T1 whole-brain GMV significantly predicted the T1 AVLT-DR scores (A; partial $r = 0.64$, $P < 0.001$; mean MAE = 1.03, $P < 0.001$) but failed to predict the T2 AVLT-DR scores (B; partial $r = 0.33$, $P = 0.019$; mean MAE = 2.00, $P = 0.072$). In contrast, the pattern of T1 whole-brain FC less accurately predicted the T1 AVLT-DR scores (C; partial $r = 0.30$, $P = 0.028$; mean MAE = 1.18, $P = 0.004$) but more significantly predicted the T2 AVLT-DR scores (D; partial $r = 0.59$, $P < 0.001$; mean MAE = 1.61, $P < 0.001$). The distribution of permutation of the prediction r values is indicated by the grey bar chart, and the distribution of the prediction r values obtained using real scores is indicated by the black bar chart. AVLT-DR, Auditory Verbal Learning Test with a 20 min delayed recall; FC, functional

connectivity; GMV, grey-matter volume; MAE, mean absolute error; T1, T1 time point (i.e., baseline); T2, T2 time point (i.e., 3-year follow-up).

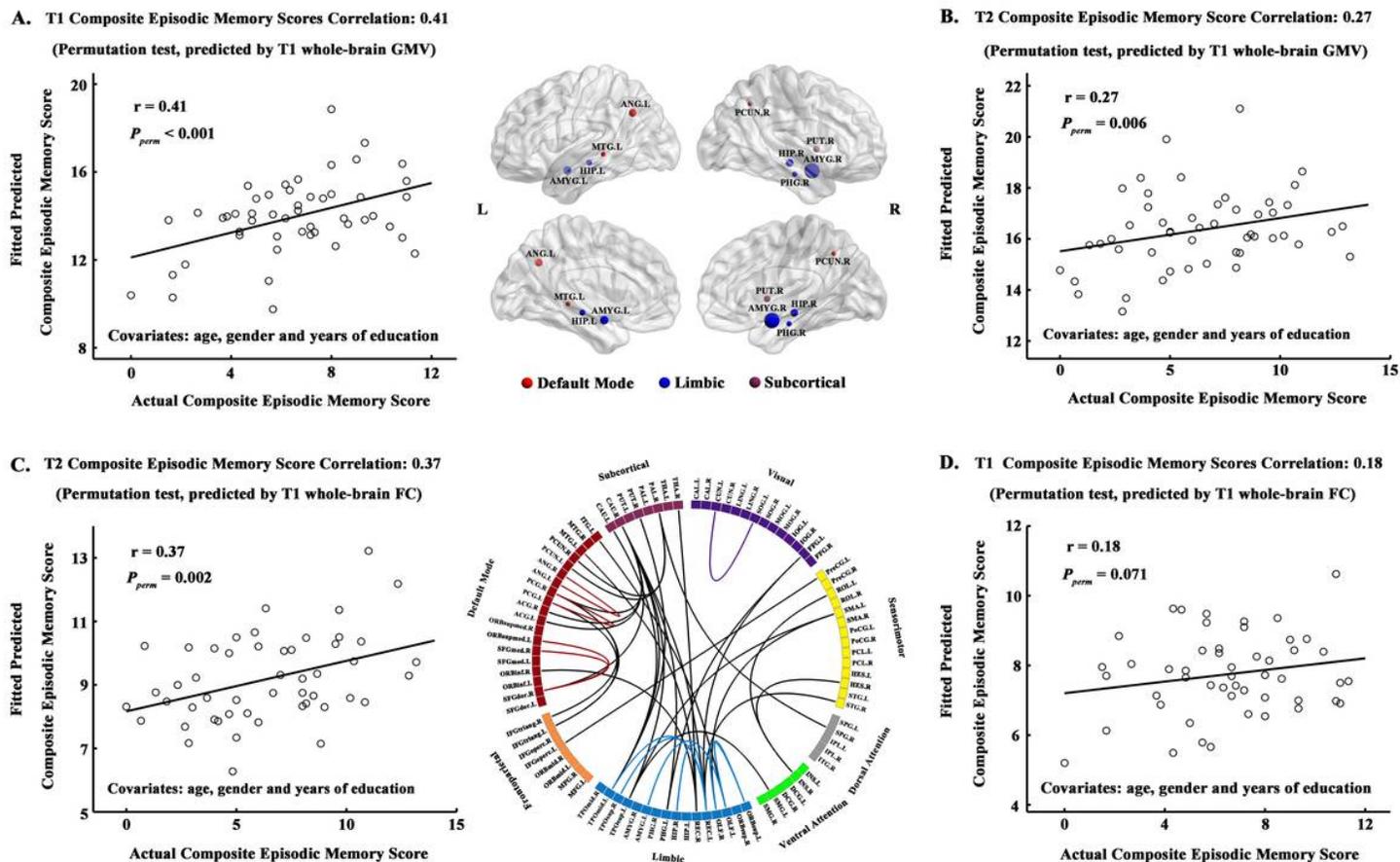


Figure 4

Validation using the composite episodic memory score. The pattern of T1 whole-brain GMV significantly predicted the T1 composite episodic memory scores (i.e., calculated by averaging the AVLT-DR, LMT-DR and CFT-DR scores) (A, left panel; partial $r = 0.41$, $P < 0.001$) but less accurately predicted the T2 composite episodic memory scores (B; partial $r = 0.27$, $P = 0.006$). (A, right panel) The most important GMV features contributing to T1 composite episodic memory score prediction involved multiple regions within the default mode and limbic systems. The pattern of T1 whole-brain FC significantly predicted the T2 composite episodic memory scores (C, left panel; partial $r = 0.37$, $P = 0.002$) but failed to predict the T1 composite episodic memory scores (D; partial $r = 0.18$, $P = 0.071$); and (C, right panel) the top 1% FCs for predicting T2 composite episodic memory scores mainly involved connections linking different functional modules (i.e., the default mode, limbic and subcortical systems) and connections within single modules (i.e., the default mode and limbic systems). AVLT-DR, Auditory Verbal Learning Test with a 20 min delayed recall; CFT-DR, the Rey-Osterrieth complex figure test with a 20 min delayed recall; GMV, grey-matter volume; FC, functional connectivity; LMT-DR, the logical memory test with a 20 min delayed recall; T1, T1 time point (i.e., baseline); T2, T2 time point (i.e., 3-year follow-up). For the abbreviations of the brain regions, see Table 2.