Combining visual rating scales of medial temporal lobe atrophy and posterior atrophy to identify amnestic mild cognitive impairment from cognitively normal older adults: evidence based on two datasets

Can Sheng
Xuanwu Hospital
Yu Sun
Xuanwu Hospital
Min Wang
Shanghai University
Xiaoni Wang
Xuanwu Hospital
Yi Liu
Xuanwu Hospital
Dongqing Pang
Tsinghua University
Jiaqi Liu
Tsinghua University
Xiaoxia Bi
Tsinghua University
Wenying Du
Xuanwu Hospital
Mingyan Zhao
Xuanwu Hospital
Yuxia Li
Xuanwu Hospital
Xiaobo Li
New Jersey Institute of Technology
Jiehui Jiang
Shanghai University
Ying Han (hanying@xwh.ccmu.edu.cn)
Xuanwu Hospital https://orcid.org/0000-0003-0377-7424

Research

Keywords: Mild cognitive impairment, Visual rating scales, Magnetic resonance imaging, Medial temporal lobe atrophy, Posterior atrophy

Posted Date: December 10th, 2019

DOI: https://doi.org/10.21203/rs.2.18215/v1

Version of Record: A version of this preprint was published at Journal of Alzheimer's Disease on September 1st, 2020. See the published version at https://doi.org/10.3233/JAD-200016.

Abstract

Background: Structural magnetic resonance imaging (MRI)-based visual rating scales are considered as the primary method for rapid and cost-effective evaluation of regional brain atrophy in patients with Alzheimer's disease (AD) in routine clinical practice. Both medial temporal lobe atrophy (MTA) and posterior atrophy (PA) visual rating scales have been reported to be useful in AD diagnosis. However, very few existing studies have investigated the efficacy of combined MTA and PA for identification of amnestic mild cognitive impairment (aMCI) from cognitively normal elderly.

Methods: This study included T1-weighted MRI images acquired inXuanwu Hospital of Capital Medical University, Beijing, China from two different cohorts. In the first cohort, we recruited 73 patients with aMCI and 48 group-matched cognitively normal controls for the training and validation. Visual assessments of MTA and PA, including left MTA, right MTA, mean MTA, left PA, right PA and mean PA, were carried out from each participant. Global gray matter (GM) volume and density was estimated using voxel-based morphometry analysis as the objective reference. Based on the receiver operating characteristic (ROC) analysis, we investigated the discriminative power of single visual rating scale and the combination of MTA and PA for successful classification between the two diagnostic groups, respectively, and then compared them to GM measures. The second cohort, consisted of 33 aMCI patients and 45 NCs, was used to verify the reliability of the discriminative power of visual assessments.

Results: In the first cohort, visual rating scales of MTA and PA showed different potential to distinguish aMCI from controls. Moreover, the combination of MTA and PA exhibited the best discriminative power, with the AUC of 0.818 ± 0.041, which was similar to the diagnostic accuracy of GM volumetric measures (0.857 ± 0.034). The discriminative power was verified in the second cohort when combining MTA and PA visual rating scales, with the AUC of 0.824 ± 0.058.

Conclusion: The combined visual rating scales of MTA and PA demonstrated practical diagnostic values for distinguishing aMCI patients from controls, suggesting its potential to serve as a convenient and reproducible method to assess the degree of atrophy in the clinical setting.

Background

Alzheimer's disease (AD) is the most common cause of neurodegenerative disorders leading to dementia. Amnestic mild cognitive impairment (aMCI), with the 2.5-year overall conversion rate of 48.7%, has been regarded as a high risk population for developing AD [1]. Given the lack of effective treatment strategies for AD dementia, early identification and intervention of aMCI may offer the opportunity for therapeutic success [2].

With the advances of neuroimaging techniques, valuable biomarkers supportive of AD pathology are becoming available [3–5]. Structural magnetic resonance imaging (MRI), characterized by its relative noninvasiveness and feasibility, has been widely employed to reveal brain global and regional morphological changes in individuals with AD and aMCI [6–8]. Current diagnostic criteria have also

recommended the utility of structural MRI for assisting in the early detection of AD [7]. Using structural MRI, researchers have reported a significant gray matter (GM) volume reduction of the medial temporal cortices (e.g., hippocampus and entorhinal cortex) in aMCI patients, which is strongly correlated with memory loss [9, 10]. However, due to the long processing time and dependency on specific algorithms in MRI quantitative analysis, brain morphological measurements have not yet been widely applied in routine clinical work [11].

Alternatively, visual rating scales may serve as a useful and cost-effective diagnostic tool in clinical settings. In 1992, Scheltens et al. first reported the diagnostic value of visual medial temporal lobe atrophy (MTA) for AD patients [12]. Using this semi-quantitative method, AD showed a significantly higher degree of MTA than controls. Subsequently, other studies also confirmed the feasibility of MTA in discriminating AD and MCI from healthy elderly controls [13–15]. For instance, Westman et al. reported a prediction accuracy of 81% for distinguishing AD from controls based on subjectively assessed MTA, which was similar to predications based on the volume of manually segmented hippocampus [15]. However, the disadvantage of visual rating scales based on sole MTA is that the atrophy of MTA also could be observed in frontotemporal lobe degeneration (FTLD) and vascular dementia (VaD), suggesting its limitation of distinguishing AD from other types of dementia [16, 17]. Moreover, it is reported that besides MTA, AD patients also present significant posterior cingulate gyrus and temporoparietal cortex atrophy, especially in early-onset individuals [18–21]. A visual rating scale of posterior atrophy (PA), proposed by Koedam et al. in 2011, was developed and referred as a valuable tool for daily assessment of dementia [18]. Möller C et al. further confirmed that PA is quantitatively validated and reflects GM atrophy in parietal regions [22]. Therefore, previous studies have demonstrated improved accuracy of the combination of MTA and PA in identifying patients with AD from normal aging [23]. Nevertheless, very few existing studies assessed the discriminative power of combined visual MTA and PA for identifying aMCI from controls.

Nowadays, two major methods were proposed to define aMCI: the conventional Petersen/Winblad criteria and the actuarial neuropsychological Jak/Bondi criteria [24]. Thus, in this study, we aimed to investigate the effectiveness of combined visual rating scales of MTA and PA to discriminate aMCI patients from normal controls (NCs). To prove the robustness of visual rating scales, we used two sources of datasets with above two criteria for aMCI diagnosis. Patients with aMCI in the first cohort were diagnosed according to the Petersen criteria, while participants with aMCI in the second cohort were identified through the Jak/Bondi criteria. First, based on cohort A, we proposed a hypothesis that compared with sole visual measure, the combination of visual MTA and PA measures could increase the identification of aMCI patients from NCs. Then, we tested the reliability and consistency of discriminative power of combined MTA and PA in cohort B. Above discriminative results were also compared to GM volumetric measurement results. Finally, the correlation of visual rating scores to specific neuropsychological tests was also evaluated.

Methods

Participants

We recruited 73 patients with aMCI diagnosis and 48 group-matched NCs in the first cohort (cohort A). Both patients and controls were recruited from the Memory Clinic of the Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing, China during Sept. 2009 to Dec. 2015. Patients with aMCI were diagnosed according to the criteria proposed by Petersen et al. [25, 26], which included: (1) memory loss complaint confirmed by an informant; (2) objective cognitive impairment in single or multiple domains, adjusted for age and education; (3) preserved general cognitive function; (4) failure to meet the criteria for dementia; (5) the clinical dementia rating (CDR) score is 0.5. The inclusion criteria for the control group included: (1) no complaint of memory loss; (2) CDR score is 0; (3) no severe visual or auditory impairment. The general exclusion criteria for both groups included: (1) a history of stroke; (2) major depression, with Hamilton Depression Rating Scale (HAMD) score > 24 points; (3) other central nervous system diseases that may cause cognitive impairment, such as Parkinson's disease, tumors, encephalitis and epilepsy; (4) cognitive impairment caused by traumatic brain injury; (5) systemic diseases, such as thyroid dysfunction, syphilis and HIV; (6) a history of psychosis or congenital mental developmental delay.

Participants in cohort B, including 33 aMCI patients and 45 cognitively normal controls, were recruited mainly through standardized public advertisements and referrals from general physicians, or memory clinics, or informants during Mar. 2017 to Apr. 2019. The definition of MCI was in accordance with the criteria proposed by Jak and Bondi in 2014, which was mainly based on regular neuropsychological tests [24]. Participants will be diagnosed as MCI if they meet any one of the following three criteria and fail to meet the criteria for dementia: (1) having impaired scores (defined as > 1 SD below the age-corrected normative means) on both measures in at least one cognitive domain (memory, language, or speed/executive function); (2) having impaired scores in each of the three cognitive domains sampled (memory, language, or speed/executive function); (3) the Functional Activities Questionnaire (FAQ) ≥ 9. Furthermore, in this study, individuals with memory complaints were considered as aMCI patients. Group-matched cognitively normal older adults were included in the control group. The general exclusion criteria were in consistence with that in the first cohort.

Research activities involved in this study have been conducted in accordance with the ethical standards of the Helsinki Declaration, and have been approved by the Medical Research Ethics Committee and Institutional Review Board of Xuanwu Hospital in the Capital Medical University (ClinicalTrials.gov Identifier: NCT02353884 and NCT03370744). All participants were in voluntarily participated in the study and provided written informed consents.

Neuropsychological assessment

Participants in cohort A carried on regular neuropsychological tests, including Montreal cognitive assessment (MoCA) (Beijing version), Auditory Verbal Learning Test (AVLT), CDR. Based on the educational years, cut-off points of MoCA tests are: 13 (no formal education), 19 (1 to 6 years of education) and 24 (7 or more years of education) [27].

Participants in cohort B carried on following neuropsychological tests focusing on three cognitive domains: 1) Memory domain (Auditory Verbal Learning Test-HuaShan version [AVLT-H] [28]): AVLT-long delayed memory, with cut-off points as 5 (50–59 years old), 4 (60–69 years old), 3 (70–79 years old); AVLT-recognition, with cut-off points as 20 (50–59 years old), 19 (60–69 years old), 18 (70–79 years old). 2) Language domain: Animal Fluency Test (AFT) [29], with cut-off points as 12 (junior middle school), 13 (high school), 14 (college); 30-item Boston Naming Test (BNT) [30], with cut-off points as 20 (junior middle school), 21 (high school), 22 (college). 3) Speed/executive domain: Shape Trail Test Part A (STT-A), with cut-off points as 70 s (50–59 years old), 80 s (60–69 years old), 100 s (70–79 years old); Shape Trail Test Part B (STT-B) [31], with cut-off points as 180 s (50–59 years old), 200 s (60–69 years old), 240 s (70–79 years old). 4) MoCA-Basic (MoCA-B) [32]: the cut-off points are: 19 (no formal education or elementary school), 22 (junior middle school or high school) and 24 (college).

MRI data acquisition

Structural MRI scanning in cohort A was performed on the 3.0 T Siemens Trio scanner (Siemens, Erlangen, Germany) at Xuanwu Hospital of Capital Medical University. All the participants were examined using a standard dementia MRI protocol which included the following sequences: three dimensional (3D) magnetization-prepared rapid gradient echo (MPRAGE) T1-weighted sequence (parameters: TR = 1900 ms, TE = 2.2 ms, TI = 900 ms, FA = 9˚, FOV = 256 × 224mm2, number of slices = 176, slice thickness = 1 mm, voxel size = 1 × 1 × 1mm3, and matrix = 256 × 256); T2-weighted fast spin echo (FSE) sequence (parameters: TR = 5500 ms, TE = 94 ms, FA = 90˚, FOV = 256 × 208mm2, number of slices = 45, slice thickness = 3 mm, and matrix = 256 × 208). Multiplanar reconstructions (MPR) of 3D T1-weighted sequences were performed in sagittal (1 mm) and oblique–coronal orientations (1 mm slice perpendicular to the long axis of the hippocampus).

The second cohort of structural MRI was acquired using an integrated simultaneous 3.0 T TOF PET/MR (SIGNA PET/MR, GE Healthcare, Milwaukee, WI, USA) at Xuanwu Hospital of Capital Medical University. Parameters for T1-weighted 3D brain structural images are as follows: SPGR sequence, FOV = 256 × 256 mm^2, matrix = 256 × 256, slice thickness = 1 mm, gap = 0, slice number = 192, repetition time (TR) = 6.9 ms, echo time (TE) = 2.98 ms, inversion time (TI) = 450 ms, flip angle = 12°, voxel size = 1 × 1 × 1 mm^3.

Visual rating assessment

In this study, we selected T1-weighted structural MRI data for the visual rating assessments of both MTA and PA. Visual rating of the entire study sample in each cohort was performed by three trained raters with an average of 8 years clinical experience, separately (cohort A: rater 1 and 2: neurologists, rater 3: a rehabilitation physician; cohort B: rater 1, 2 and 3: all are neurologists). All raters were blind to any clinical diagnostic and behavioral evaluation information of the participants.

MTA for the left and right hemispheres were separately rated on the oblique-coronal MPR sections of the 3D T1-weighted sequence, using a 5-point rating scale (0–4) previously described by Scheltens et al [12]. The medial temporal region of each hemisphere was manually parcellated based on the height of the

hippocampal formation and the width of the choroid fissure and the temporal horn. MTA was defined as: 0 point for no atrophy; 1 point for minimal atrophy; 2 points for mild atrophy; 3 points for moderate atrophy; and 4 points for severe atrophy. PA was also assessed for the left and right hemispheres separately based on the sagittal, axial, and oblique-coronal MPR reconstruction of the T1-weighted sequence [18]. In the event of discrepant scores on different orientations (e.g. score 0 on axial orientation and score 1 on coronal orientation), the highest score was selected. Based on anatomical regions of the posterior cingulate sulcus, parieto-occipital sulcus, the cortex of parietal lobes and precuneus, PA was rated using a 4-point scale (0 point = absent; 1 point = mild widening of sulcus and mild atrophy; 2 points = substantial widening and substantial atrophy; and 3 points = evident widening of sulci and knife-blade atrophy). For both MTA and PA, mean values of the left and right side scores were also calculated. Figure 1 showed rough anatomical structure of MTA (A) and PA (B) regions.

Inter- and intra-rater agreement analysis

To quantify inter- and intra-rater agreements, we calculated intra-class correlation coefficient (ICC) values between any two raters and between the first and second session of three raters separately [33]. The reference point for the categorization ICC is: 1) < 0.4 (relatively low agreement); 2) 0.4–0.75 (moderate agreement); 3) > 0.75 (relatively high agreement). For evaluating the intra-rater agreement, we randomly selected twenty MRI images from all participants in each dataset and conducted random sampling with 500 times. Both cohort A and cohort B used the ICC method to quantify inter- and intra-rater agreements.

Voxel-based morphometry

Because not yet precise anatomical atlas can be used to extract MTA and PA regions same as visual assessments, we compared our results with objective GM volume and density measurements based on the whole brain. All T1-weighted structural MRI images were preprocessed using voxel-based morphometry (VBM) analysis in SPM12 (Statistical Parametric Mapping, Version 12, https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) and MATLAB 2014b. For variability in scanning parameters, MRI scans and Dartel imported scans were registered into the stereotaxic space by applying rigid-body transformations and Dartel nonlinear image registration procedure. GM, white matter (WM) and cerebrospinal fluid (CSF) tissue probability maps with priori tissue maps as reference were acquired by the unified segmentation algorithm. Subsequently, GM maps were normalized to the Montreal Neurological Institute (MNI) International Consortium for Brain Mapping 152 (ICBM152) template. Finally, the normalized GM images were smoothed with a 4 mm full-width half-maximum (FWHM) Gaussian kernel to reduce the spatial signal-to-noise ratio and the error caused by space normalization for individuals. The GM measures, including whole brain GM volume (absolute and relative volume) and density of each individual were calculated for further comparison.

ROC analysis

To determine whether visual rating measures have the potential to serve as a biomarker for distinguishing individuals with aMCI from controls, receiver operating characteristics (ROC) curves were employed. The

ROC curves were obtained based on the values of sensitivity and specificity for each of visual rating scales and GM volume measures. Then, the area under ROC curve (AUC) was used to quantitatively assess the discriminative power of these measures, and the confidence interval of AUC to clarify their significance. Furthermore, to estimate the discriminative power of combination of MTA and PA, a multivariate-based ROC analysis was employed. Using the Logistic regression equation models, new predicted probabilities were calculated via combining MTA and PA visual rating measures. The new predicted probabilities served as test variables for the ROC graph. AUC was also used to assess the discriminative power of this predicted probability.

Statistical analysis

For normally distributed data, two-sample t tests were performed to compare group differences. For non-normally distributed data, a Mann–Whitney U test was used. A Chi-squared test was used to compare group differences in gender. To facilitate analysis of group differences in visual rating measures, scores of a rater (rater 1) were used. To solve the possible defects of overfitting in the ROC analysis, standard permutation test was employed using 1000 random resamplings of data, and the results were averaged to produce a final classification performance with mean and standard deviation (SD) values. In addition, we accessed the correlation between visual rating scales and neuropsychological assessments using Spearman partial correlation analysis, with age, gender and years of education as covariates. All data processing and analyses were performed using SPSS 22.0 and R 3.5.1. $P < 0.05$ (two-tailed) was considered significant.

Results

Demographic and neuropsychological characteristics

There were no significant differences in age, gender, or years of education between aMCI patients and controls both in **cohort A** and **cohort B**. However, compared with controls, patients with aMCI in **cohort A** showed significantly lower scores of the MoCA, AVLT-immediate recall (AVLT-I), AVLT-delayed recall (AVLT-D) and AVLT-recognition (AVLT-R) ($P < 0.001$); and in **cohort B,** MCI patients exhibited significant decline on the MoCA-B, AVLT-D (long), AVLT-R, STT-A, STT-B, AFT and BNT. The detailed information was presented in **Table 1**.

The visual ratings of MTA and PA

Six visual rating measures, including left MTA, right MTA, mean MTA, left PA, right PA and mean PA, were assessed separately for each participant. Relative to controls, visual rating scores in above-mentioned measures were higher in patients with aMCI, suggesting more significant regional brain atrophy of the medial temporal lobe and parietal areas in the stage of aMCI (**Table 2**).

Inter- and intra-rater reliability

In **cohort A**, the value of inter-rater agreement ranged from 0.761 to 0.916, and intra-rater agreement for all visual rating measures ranged from 0.735 to 0.922, which indicate a relatively good consistency. In **cohort B**, inter-rater agreement for MTA was best between rater 1 and 2, with a value of 0.900, followed by a value of 0.833 between rater 1 and 3 and lowest between rater 2 and 3, with a value of 0.802. Intra-rater agreement for MTA varied between 0.719 and 0.884. Inter-rater agreement for PA was relatively high between rater 1 and 2, with a value of 0.845, whereas values for intra-rater agreement varied ranging from 0.709 to 0.832. The detailed information is shown in **Table 3**.

Voxel-based morphometry

For the GM images from all participants, we computed the whole brain GM volumetric measures using VBM analysis. The results revealed that aMCI patients clearly showed lower GM volumetric measures (GM volumes and density) relative to controls ($P < 0.001$ in **cohort A and B**) (**Figure 2**), and a negative correlation between whole brain GM measures and visual rating scores could be observed in MTA and PA regions (**Figure 3**). These phenomena indicated that the GM volume loss had happened in the aMCI stage and the visual assessment ratings in specific brain regions would help physicians to efficiently diagnose aMCI patients in clinical practice.

Visual rating and GM volume-based classification analysis

Using the ROC analysis approach, we first estimated the discriminative power of each of visual rating scales and GM volumetric measures in identifying patients with aMCI from controls (**Table 4**). The visual rating scales of MTA and PA exhibited the potential discriminative power, with the AUC of 0.776 ± 0.044 and 0.725 ± 0.045, respectively (**Figure 4A**). The GM relative volume showed a relatively good discriminative power, followed by the GM density, with the AUC of the ROC of 0.839 ± 0.034 and 0.783 ± 0.042, separately.

Furthermore, we calculated the discriminative power of combination of the MTA and PA, as well as the combined multiple GM volumetric measures for distinguishing aMCI from controls. The combination of the MTA and PA showed relatively higher classification accuracy compared with single visual rating scale, with the AUC of 0.818 ± 0.041(**Figure 4A**). The discriminative power of the combined GM measures was relatively excellent, with the AUC of 0.857 ± 0.034. These results suggested that the combination of the multiple visual rating scales were beneficial to optimize the classification accuracy of aMCI.

The similar findings were also demonstrated in **cohort B**. The visual rating scales of MTA exhibited best discriminative power with the AUC of 0.822 ± 0.053. When combining the MTA and PA, the discriminative power increased, with the AUC of 0.824 ± 0.058 (**Figure 4B**). Similarly, diagnostic efficacy of quantitative GM features was proximate to that of visual rating scales.

Correlation between cognitive scores and visual rating measures in aMCI patients

We implemented a spearman partial correlation analysis for all six visual rating measures and the cognitive assessments to investigated whether visual rating measures could reflect the cognitive decline in aMCI patients. The correlation results in **cohort A** and **cohort B** were summarized in **Table 5**. The results showed that left MTA had a significantly negative correlation with AVLT-I (R = -0.293, *P* = 0.015), AVLT-D (R = -0.265, *P* = 0.028) and AVLT-R (R = -0.248, *P* = 0.040), while mean MTA was negatively correlated with AVLT-I (R = -0.260, *P* = 0.031). The visual rating scale of PA had no correlation with cognitive scores. In **cohort B**, there was positive correlation between PA and STT-A, suggesting the executive dysfunction in patients with aMCI.

Discussion

In present study, we investigated the effectiveness of visual rating scales of MTA and PA in identifying patients with aMCI and found that: 1) Both visual assessments based on MTA and PA were effective, and the combination of visual rating scales of the MTA and PA achieved higher discriminative power between aMCI patients and NCs than sole visual rating scales; 2) the similar discriminative power has been verified in another cohort, indicating the repeatability and consistency of our results. Taken together, our findings demonstrated apparent regional brain atrophy both in medial temporal lobe and posterior areas among aMCI patients. Combining multiple visual rating scales appeared to be convenient and rapid in identifying aMCI with relatively higher diagnostic accuracy than single visual rating scale, which has the potential to be widespread application as a biomarker in clinical practice due to its convenience and speediness.

Visual rating characteristics of MTA and PA in aMCI

Medial temporal lobe, as a critical component of typical AD-related regions, has been revealed structural alterations both in patients with AD and aMCI [34]. Individuals with aMCI who ultimately converted to AD dementia initially presented GM volume loss in the medial temporal lobe, such as hippocampus and entorhinal cortex [9]. Therefore, atrophy in the medial temporal lobe is recommended as a topographical biomarker indicating the progression of MCI. In comparison to complicated GM volumetric measurements, evaluating regional brain atrophy using visual rating scale of MTA is considered as a quick and reproducible method in routine clinical practice [35]. As Shen Q and his colleagues reported, compared with hippocampus volumetric measures, visual MTA even provided better discriminatory power of distinguishing aMCI or AD from healthy controls [36]. In our study, we also found statistically significant higher degree of MTA in aMCI patients compared with control group in two different datasets, suggesting that medial temporal lobe was selectively attacked by AD-related pathology in the prodromal stage of AD. Furthermore, negative correlation between MTA and AVLT was also showed in this study, verifying the closely link between the disruption of medial temporal lobe and memory loss. Similarly,

previous structural MRI-based quantitative analysis also demonstrated the association between hippocampal atrophy and cognitive impairment [37, 38]. In a word, visual MTA assessment exhibits the similar effectiveness in mirroring the severity of memory decline.

As clinicians may be unaware of structural changes of the posterial regions in AD patients, PA may be under-recognized in clinical practice. It is increasingly acknowledged that the presence of MTA should not be restricted to patients with AD but also in other types of dementia, such as FTLD, VaD, and semantic dementia (SD), indicating that MTA may be less effective in distinguishing AD from other types of dementia [16, 17, 39, 40]. Additionally, not every AD patient presents with MTA and one of AD subtypes, called no-atrophy AD, has been reported in previous studies [19]. Furthermore, current studies have emphasized a prominent posterior (parietal) atrophy pattern in AD and about thirty percent of AD patients showed PA without MTA [11, 23]. Posterior cerebral atrophy, generally affecting the posterior cingulate gyrus, precuneus and parietal lobes, has been confirmed in a large cohort of patients with AD, particularly in younger individuals [41]. Similarly, in our study, aMCI patients showed higher level of PA compared with controls, suggesting significant atrophy of posterior areas in the stage of prodromal AD. PA is also considered to be associated with worse performance on visuospatial and executive function [42], which was confirmed that there was positive correlation between PA and STT-A in cohort B.

The reliability of combined visual rating scales

Previous studies have suggested that the pattern of GM loss in the aMCI subject scans is initially focused on the medial temporal lobes, and subsequently, extends into the posterior regions, including the parietal lobe and the temporoparietal association cortices [9]. This dynamically changing atrophy patterns reveal the importance of combining medial temporal areas and posterior cortices in objectively and comprehensively evaluating the structural changes in the stage of aMCI. Thus, owing to the complementary morphological information provided by different visual rating scales, the combination of MTA and PA appeared to achieve more effective diagnosis than single visual rating measure. Koedam et al. emphasized the increased diagnostic sensitivity of AD when combining MTA and PA [18]. One study also recommended that adding the PA to the MTA could further improve the discrimination of AD from controls (AUC 0.87), although discrimination abilities were good for both MTA and PA scales (AUC 0.80 and 0.74, respectively) [23]. As is shown in present study, we demonstrated the advantages of the integration of MTA and PA in identifying aMCI patients, with a significantly increased discriminative power (AUC 0.818 ± 0.041), which was approximate to the results of combined multiple GM measures. Our main findings were verified by the test dataset (cohort B) (AUC 0.824 ± 0.058), indicating the repeatability and consistency of the results in different clinical settings.

Various types of cortical morphological features, such as cortical thickness, GM volume, metric distortion, and GM densities, have shown promising capacities for the classification between aMCI and controls [43–45]. Our study revealed the relatively high classification accuracy by using each of parameters extracted from the whole brain (GM volume and density), which was in line with previous studies. For example, based on the cortical thickness, the classification accuracy in discriminating aMCI from controls was 78% in the left hemisphere and 60% in the right hemisphere [46]. Similarly, given that different

morphological features derived from structural MRI have unique neuropathological characteristics and various contributions in discriminating aMCI from cognitively healthy controls [47–49], integrations of multiple features may be beneficial to improve the diagnostic accuracy for aMCI [50, 51]. Li et al. extracted six cortical features for each aMCI subject and demonstrated the highest discriminative power (84%) by combining the metric distortion and cortical thickness features in the left hemisphere [46]. Xiao et al. also reported a relatively high classification accuracy of 86.11% for aMCI based on the combination of texture features and morphometric features [45], indicating that multi-feature combination was better than the single-feature method. Our study exhibited similar classification power via the combination of multiple GM measures, with the AUC of 0.857 ± 0.034.

However, the discrimination accuracy of combined GM measures in cohort B was significantly improved, with the mean AUC of 0.957 ± 0.038. This may be caused by the possible defect of overfitting in the ROC analysis, although random permutation was repeated 1000 times and results were averaged to produce a final classification performance.

In summary, visual assessments were clinically useful and yielded a diagnostic accuracy being close to the quantitative MRI measures. This study emphasized the improved effectiveness in the diagnostic work-up of aMCI patients when combining visual assessments of MTA and PA.

Limitations and future directions

Several limitations of this study need to be considered. First, although manifesting the convenient and fast advantages in evaluating brain morphological alterations, visual rating scales are still subjectively semi-quantitative method without the relatively objective accuracy as traditionally MRI quantitative analysis. This study provided the evidence of capability and possibility of visual assessments in clinical practice. Second, our study was based on the small clinical cohorts without brain autopsy confirmation. Though the reliability of our current findings had been confirmed by the test dataset, a larger sample size from multi centers might be essential to provide more evidence in the future study. Third, the different sources of participants in two datasets may signify different severity of the disease. In cohort A, participants were recruited from the memory clinics. Several studies have revealed the relationship between the severity of cognitive concern and the extent of AD-related pathology in aging cohorts of cognitively normal individuals [52, 53]. Participants in cohort A were active to seek medical help and advice due to their concerns about the cognitive decline, suggesting that they might be in the later stage of aMCI. Nevertheless, participants in cohort B were mainly from community-based advertisements and they might be in the earlier stage of MCI. As was showed in this study, visual rating scores of aMCI patients in cohort B were lower than that in cohort A. Thus, in the future, maintaining the consistent source of participants seems to be important to ensure patients in the same stage of the disease. Finally, the present study was a cross-sectional design and the longitudinal development of these aMCI participants is unknown. In the future, studies should focus on aMCI with the evidence of AD pathologies derived from the CSF or positron emission tomography (PET), which might provide higher reliability and accuracy in the diagnosis of prodromal AD.

Conclusions

The discriminative power of visual rating scales for identifying patients with aMCI from cognitively normal controls was preliminarily assessed in this study. Based on two datasets with different criteria for MCI diagnosis, the combination of visual rating scales of MTA and PA exhibited more effective discriminative power in discriminating aMCI from controls, suggesting its repeatability and diagnostic value as a neuroimaging biomarker in the clinical practice. Although semi-quantitative and subjective, visual rating scales remain the primary method for extracting diagnostically useful information in the clinical settings.

Declarations

Ethics approval and consent to participate

Research activities involved in this study have been conducted in accordance with the ethical standards of the Helsinki Declaration, and have been approved by the Medical Research Ethics Committee and Institutional Review Board of Xuanwu Hospital in the Capital Medical University.

Consent for publication

All participants voluntarily participated in the study and provided written informed consents.

Availability of data and materials

All data generated or analysed during this study are included in this published article. I can confirm I have included a statement regarding data and material availability in the declaration section of my manuscript.

Competing interests

There are no conflicts of interest including any financial, personal, or other relationships with people or organizations for any of the authors described in the article.

Funding

This article was supported by the National Key Research and Development Program of China (2016YFC1306300, 2018YFC1312001), National Natural Science Foundation of China (61633018, 81801052), Beijing Nature Science Foundation (7161009), China Postdoctoral Science Foundation

(2018M641414), Beijing Municipal Commission of Health and Family Planning (PXM2019_026283_000002).

Authors' contributions

CS, YH, JJ designed the study. YH, XL, YL, WD, MZ, JL supervised data collection, interpreted the data and offered significant comments on the manuscript. CS, YS, MW, XW, YL, XB, JL and DP analyzed the data and wrote the first draft of the manuscript. All authors reviewed the manuscript.

Acknowledgements

We would like to thank all the participants in this study. We also acknowledge Dr. Chunxiu Wang for her assistance in statistical analysis.

Authors' information

[1]Department of Neurology, Xuanwu Hospital of Capital Medical University, Beijing, China, 100053,

[2]Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai, China, 200444,

[3]Department of Neurology, the First Hospital of Tsinghua University, Beijing, China, 100016,

[4]Department of Biomedical Engineering, New Jersey Institute of Technology, NJ, USA,

[5]Center of Alzheimer's Disease, Beijing Institute for Brain Disorders, Beijing, China, 100053,

[6]National Clinical Research Center for Geriatric Disorders, Beijing, China, 100053.

Additional file

This manuscript has no additional file.

Abbreviations

Alzheimer's disease: AD; Medial temporal lobe atrophy: MTA; Posterior atrophy: PA; Amnestic mild cognitive impairment: aMCI; Normal controls: NCs; The receiver operating characteristic: ROC; Frontotemporal lobe degeneration: FTLD; Vascular dementia: VaD; Clinical dementia rating: CDR; Mini-mental state examination: MMSE; Montreal cognitive assessment: MoCA; Auditory Verbal Learning Test: AVLT; Auditory verbal learning test-immediate recall: AVLT-I; Auditory verbal learning test-long delayed

recall: AVLT-D (long); Auditory verbal learning test-recognition: AVLT-R; Montreal Cognitive Assessment-Basic: MoCA-B; Shape Trail Test Part A: STT-A; Shape Trail Test Part B: STT-B; Animal Fluency Test: AFT; Boston Naming Test: BNT；Three dimensional: 3D; Magnetization-prepared rapid gradient echo: MPRAGE; Fast spin echo: FSE; Multiplanar reconstructions: MPR; Gray matter: GM; Voxel-based morphometry: VBM; Intraclass correlation coefficient: ICC; Cerebrospinal fluid: CSF; Full-width half-maximum: FWHM; The area under ROC curve: AUC; Semantic dementia: SD; Dementia with Lewy bodies: DLB; Positron emission tomography: PET

References

1. Fischer P, Jungwirth S, Zehetmayer S, Weissgram S, Hoenigschnabl S, Gelpi E, et al. Conversion from subtypes of mild cognitive impairment to Alzheimer dementia. Neurology. 2007;68(4):288-91.
2. Sheng C, Huang Y, Han Y. Dissection of prodromal Alzheimer's disease. Front Biosci (Landmark Ed). 2018;23:1272-91.
3. Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. The Lancet Neurology. 2014;13(6):614-29.
4. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2016;12(3):292-323.
5. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr., Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011;7(3):263-9.
6. Verhagen MV, Guit GL, Hafkamp GJ, Kalisvaart K. The impact of MRI combined with visual rating scales on the clinical diagnosis of dementia: a prospective study. Eur Radiol. 2016;26(6):1716-22.
7. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011;7(3):280-92.
8. Fennema-Notestine C, McEvoy LK, Hagler DJ, Jr., Jacobson MW, Dale AM, The Alzheimer's Disease Neuroimaging I. Structural neuroimaging in the detection and prognosis of pre-clinical and early AD. Behav Neurol. 2009;21(1):3-12.
9. Whitwell JL, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, et al. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. Brain. 2007;130(Pt 7):1777-86.
10. Korf ESC, Wahlund L-O, Visser PJ, Scheltens P. Medial temporal lobe atrophy on MRI predicts dementia in patients with mild cognitive impairment. Neurology. 2004;63(1):94-100.

11. Harper L, Fumagalli GG, Barkhof F, Scheltens P, O'Brien JT, Bouwman F, et al. MRI visual rating scales in the diagnosis of dementia: evaluation in 184 post-mortem confirmed cases. Brain. 2016;139(Pt 4):1211-25.
12. Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, et al. Atrophy of medial temporal lobes on MRI in "probable" Alzheimer's disease and normal ageing: diagnostic value and neuropsychological correlates. J Neurol Neurosurg Psychiatry. 1992;55(10):967-72.
13. Duara R, Loewenstein DA, Shen Q, Barker W, Varon D, Greig MT, et al. The utility of age-specific cut-offs for visual rating of medial temporal atrophy in classifying Alzheimer's disease, MCI and cognitively normal elderly subjects. Frontiers in aging neuroscience. 2013;5:47-.
14. Tuokkola T, Koikkalainen J, Parkkola R, Karrasch M, Lötjönen J, Rinne JO. Visual rating method and tensor-based morphometry in the diagnosis of mild cognitive impairment and Alzheimer's disease: a comparative magnetic resonance imaging study. Acta Radiol. 2016;57(3):348-55.
15. Westman E, Cavallin L, Muehlboeck JS, Zhang Y, Mecocci P, Vellas B, et al. Sensitivity and specificity of medial temporal lobe visual ratings and multivariate regional MRI classification in Alzheimer's disease. PLoS One. 2011;6(7):e22506-e.
16. Pleizier CM, van der Vlies AE, Koedam E, Koene T, Barkhof F, van der Flier WM, et al. Episodic memory and the medial temporal lobe: not all it seems. Evidence from the temporal variants of frontotemporal dementia. J Neurol Neurosurg Psychiatry. 2012;83(12):1145-8.
17. Bastos-Leite AJ, van der Flier WM, van Straaten ECW, Staekenborg SS, Scheltens P, Barkhof F. The contribution of medial temporal lobe atrophy and vascular pathology to cognitive impairment in vascular dementia. Stroke. 2007;38(12):3182-5.
18. Koedam ELGE, Lehmann M, van der Flier WM, Scheltens P, Pijnenburg YAL, Fox N, et al. Visual assessment of posterior atrophy development of a MRI rating scale. Eur Radiol. 2011;21(12):2618-25.
19. Persson K, Eldholm RS, Barca ML, Cavallin L, Ferreira D, Knapskog A-B, et al. MRI-assessed atrophy subtypes in Alzheimer's disease and the cognitive reserve hypothesis. PLoS One. 2017;12(10):e0186595-e.
20. Frisoni GB, Pievani M, Testa C, Sabattoli F, Bresciani L, Bonetti M, et al. The topography of grey matter involvement in early and late onset Alzheimer's disease. Brain. 2007;130(Pt 3):720-30.
21. Frisoni GB, Testa C, Sabattoli F, Beltramello A, Soininen H, Laakso MP. Structural correlates of early and late onset Alzheimer's disease: voxel based morphometric study. J Neurol Neurosurg Psychiatry. 2005;76(1):112-4.
22. Möller C, van der Flier WM, Versteeg A, Benedictus MR, Wattjes MP, Koedam ELGM, et al. Quantitative regional validation of the visual rating scale for posterior cortical atrophy. Eur Radiol. 2014;24(2):397-404.
23. Lehmann M, Koedam ELGE, Barnes J, Bartlett JW, Ryan NS, Pijnenburg YAL, et al. Posterior cerebral atrophy in the absence of medial temporal lobe atrophy in pathologically-confirmed Alzheimer's disease. Neurobiol Aging. 2012;33(3):627.e1-.e12.

24. Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. Journal of Alzheimer's disease : JAD. 2014;42(1):275-89.
25. Petersen RC, Doody R, Kurz A, Mohs RC, Morris JC, Rabins PV, et al. Current concepts in mild cognitive impairment. Arch Neurol. 2001;58(12):1985-92.
26. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med. 2004;256(3):183-94.
27. Li Y, Wang X, Li Y, Sun Y, Sheng C, Li H, et al. Abnormal Resting-State Functional Connectivity Strength in Mild Cognitive Impairment and Its Conversion to Alzheimer's Disease. Neural Plast. 2016;2016:4680972-.
28. Zhao Q, Lv Y, Zhou Y, Hong Z, Guo Q. Short-term delayed recall of auditory verbal learning test is equivalent to long-term delayed recall for identifying amnestic mild cognitive impairment. PLoS One. 2012;7(12):e51157-e.
29. Guo Q, Jin L, Hong Z, Lv C. A Specific Phenomenon of Animal Fluency Test in Chinese Elderly. Chinese Mental Health Journal. 2007;21(9):622-5.
30. Guo Q, Hong Z, Shi W, Sun Y, Lv C. Boston naming test in Chinese elderly, patient with mild cognitive impairment and Alzheimer's dementia. Chinese Mental Health Journal. 2006;20(2):81-4.
31. Zhao Q, Guo Q, Li F, Zhou Y, Wang B, Hong Z. The Shape Trail Test: application of a new variant of the Trail making test. PLoS One. 2013;8(2):e57333-e.
32. Chen K-L, Xu Y, Chu A-Q, Ding D, Liang X-N, Nasreddine ZS, et al. Validation of the Chinese Version of Montreal Cognitive Assessment Basic for Screening Mild Cognitive Impairment. J Am Geriatr Soc. 2016;64(12):e285-e90.
33. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Rep. 1966;19(1):3-11.
34. Sheng C, Xia M, Yu H, Huang Y, Lu Y, Liu F, et al. Abnormal global functional network connectivity and its relationship to medial temporal atrophy in patients with amnestic mild cognitive impairment. PLoS One. 2017;12(6):e0179823-e.
35. Harper L, Barkhof F, Fox NC, Schott JM. Using visual rating to diagnose dementia: a critical evaluation of MRI atrophy scales. J Neurol Neurosurg Psychiatry. 2015;86(11):1225-33.
36. Shen Q, Loewenstein DA, Potter E, Zhao W, Appel J, Greig MT, et al. Volumetric and visual rating of magnetic resonance imaging scans in the diagnosis of amnestic mild cognitive impairment and Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2011;7(4):e101-e8.
37. Risacher SL, Anderson WH, Charil A, Castelluccio PF, Shcherbinin S, Saykin AJ, et al. Alzheimer disease brain atrophy subtypes are associated with cognition and rate of decline. Neurology. 2017;89(21):2176-86.
38. Cholerton B, Omidpanah A, Madhyastha TM, Grabowski TJ, Suchy-Dicey AM, Shibata DK, et al. Total Brain and Hippocampal Volumes and Cognition in Older American Indians: The Strong Heart Study. Alzheimer Dis Assoc Disord. 2017;31(2):94-100.

39. Staekenborg SS, van Straaten ECW, van der Flier WM, Lane R, Barkhof F, Scheltens P. Small vessel versus large vessel vascular dementia: risk factors and MRI findings. J Neurol. 2008;255(11):1644-814.
40. van de Pol LA, Hensel A, van der Flier WM, Visser PJ, Pijnenburg YAL, Barkhof F, et al. Hippocampal atrophy on MRI in frontotemporal lobar degeneration and Alzheimer's disease. J Neurol Neurosurg Psychiatry. 2006;77(4):439-42.
41. Smits LL, Tijms BM, Benedictus MR, Koedam ELGE, Koene T, Reuling IEW, et al. Regional atrophy is associated with impairment in distinct cognitive domains in Alzheimer's disease. Alzheimer's & dementia : the journal of the Alzheimer's Association. 2014;10(5 Suppl):S299-S305.
42. Yuan Z, Pan C, Xiao T, Liu M, Zhang W, Jiao B, et al. Multiple Visual Rating Scales Based on Structural MRI and a Novel Prediction Model Combining Visual Rating Scales and Age Stratification in the Diagnosis of Alzheimer's Disease in the Chinese Population. Front Neurol. 2019;10:93-.
43. Raamana PR, Wen W, Kochan NA, Brodaty H, Sachdev PS, Wang L, et al. The Sub-Classification of Amnestic Mild Cognitive Impairment Using MRI-Based Cortical Thickness Measures. Front Neurol. 2014;5:76-.
44. Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiol Aging. 2008;29(4):514-23.
45. Xiao Z, Ding Y, Lan T, Zhang C, Luo C, Qin Z. Brain MR Image Classification for Alzheimer's Disease Diagnosis Based on Multifeature Fusion. Comput Math Methods Med. 2017;2017:1952373-.
46. Li S, Yuan X, Pu F, Li D, Fan Y, Wu L, et al. Abnormal changes of multidimensional surface features using multivariate pattern classification in amnestic mild cognitive impairment patients. J Neurosci. 2014;34(32):10541-53.
47. Wang H, Golob E, Bert A, Nie K, Chu Y, Dick MB, et al. Alterations in regional brain volume and individual MRI-guided perfusion in normal control, stable mild cognitive impairment, and MCI-AD converter. J Geriatr Psychiatry Neurol. 2009;22(1):35-45.
48. Zhang D, Wang Y, Zhou L, Yuan H, Shen D, Alzheimer's Disease Neuroimaging I. Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage. 2011;55(3):856-67.
49. Xie Y, Cui Z, Zhang Z, Sun Y, Sheng C, Li K, et al. Identification of Amnestic Mild Cognitive Impairment Using Multi-Modal Brain Features: A Combined Structural MRI and Diffusion Tensor Imaging Study. Journal of Alzheimer's disease : JAD. 2015;47(2):509-22.
50. Wee C-Y, Yap P-T, Zhang D, Denny K, Browndyke JN, Potter GG, et al. Identification of MCI individuals using structural and functional connectivity networks. Neuroimage. 2012;59(3):2045-56.
51. Cui Y, Wen W, Lipnicki DM, Beg MF, Jin JS, Luo S, et al. Automated detection of amnestic mild cognitive impairment in community-dwelling elderly adults: a combined spatial atrophy and white matter alteration approach. Neuroimage. 2012;59(2):1209-17.
52. Verfaillie SCJ, Timmers T, Slot RER, van der Weijden CWJ, Wesselman LMP, Prins ND, et al. Amyloid-β Load Is Related to Worries, but Not to Severity of Cognitive Complaints in Individuals With Subjective Cognitive Decline: The SCIENCe Project. Frontiers in aging neuroscience. 2019;11:7-.

53. Buckley RF, Hanseeuw B, Schultz AP, Vannini P, Aghjayan SL, Properzi MJ, et al. Region-Specific Association of Subjective Cognitive Decline With Tauopathy Independent of Global β-Amyloid Burden. JAMA Neurol. 2017;74(12):1455-63.

Table

Table 1. Demographic and neuropsychological assessments for all participants.

	Cohort A			Cohort B		
	aMCI (*n* = 73)	NC (*n* = 48)	*P*-value	aMCI (*n* = 33)	NC (*n* = 45)	*P*-value
Age (years)	68.6 ± 9.24	65.7 ± 6.67	0.061[a]	64.94 ± 8.28	63.09 ± 4.80	0.218[a]
Gender (M/F)	33/40	19/29	0.541[b]	17/16	18/27	0.312[b]
Education (years)	9.36 ± 4.79	11.1 ± 4.84	0.059[a]	11.45 ± 3.15	12.22 ± 3.19	0.295[a]
MoCA/MoCA-B	18.2 ± 4.28	26.5 ± 2.44	<0.001[a]	21.21 ± 3.76	26.44 ± 2.38	<0.001[a]
AVLT-I	5.36 ± 1.46	9.24 ± 1.69	<0.001[a]	/	/	/
AVLT-D (long)	3 (0, 5)	10.1 ± 2.65	<0.001[c]	3.48 ± 1.94	7.29 ± 1.87	<0.001[a]
AVLT-R	6.51 ± 3.81	11.8 ± 2.39	<0.001[a]	18.67 ± 3.04	22.62 ± 1.28	<0.001[a]
STT-A	/	/	/	90.12 ± 34.99	60.10 ± 24.50	<0.001[a]
STT-B	/	/	/	238.09 ± 104.19	132.69 ± 41.73	<0.001[a]
AFT	/	/	/	15.00 ± 4.55	19.82 ± 3.58	<0.001[a]
BNT	/	/	/	21.61 ± 3.81	25.62 ± 2.87	<0.001[a]

For normally distributed data, they are presented as the mean ± SD;

For non-normally distributed data, they are presented as the median (IQR);

aMCI, amnestic mild cognitive impairment; NC, normal control; MoCA, Montreal Cognitive Assessment; AVLT-I, auditory verbal learning test-immediate recall; AVLT-D (long), auditory verbal learning test-long delayed recall; AVLT-R, auditory verbal learning test-recognition; MoCA-B, Montreal Cognitive Assessment-Basic; STT-A, Shape Trail Test Part A; STT-B, Shape Trail Test Part B; AFT, Animal Fluency Test; BNT, Boston Naming Test;

[a] Two-sample *t*-test; [b] Pearson chi-square test; [c] Mann-Whitney U test.

Table 2. Group differences of visual rating scales between aMCI and NC subjects.

	Cohort A			Cohort B		
	aMCI (*n* = 73)	NC (*n* = 48)	*P* value	aMCI (*n* = 33)	NC (*n* = 45)	*P*-value
left MTA	2 (1, 2)	1 (0, 1)	<0.001	1(0, 2)	0 (0, 1)	<0.001
right MTA	2 (1, 2.5)	1 (0, 1)	<0.001	1 (1, 2)	0 (0, 1)	<0.001
mean MTA	2 (1, 2)	0.75 (0, 1.38)	<0.001	1.5 (0.5, 1.75)	0.5 (0, 0.5)	<0.001
left PA	2 (1, 3)	1 (1, 2)	<0.001	1(1, 2)	1 (0, 1)	0.022
right PA	2 (1, 3)	1 (1, 2)	<0.001	1(1, 2)	1 (0, 1)	<0.001
mean PA	2 (1.25, 2.5)	1 (0.63, 2)	<0.001	1(1, 2)	0.5 (0, 1)	<0.001

Data are presented as the median (IQR), Group differences were compared using Mann-Whitney U test;

aMCI, amnestic mild cognitive impairment; NC, normal control; MTA, medial temporal lobe atrophy; PA, posterior atrophy.

Table 3. Inter- and intra-rater agreement for visual rating of MTA and PA in Data 1 and Data 2

	Cohort A						Cohort B					
	left MTA	right MTA	mean MTA	left PA	right PA	mean PA	left MTA	right MTA	mean MTA	left PA	right PA	mean PA
Rater couple 1–2	0.828	0.858	0.916	0.792	0.768	0.829	0.905	0.850	0.900	0.784	0.809	0.845
inter-rater agreement												
Rater couple 1–3	0.847	0.813	0.866	0.786	0.765	0.806	0.724	0.752	0.833	0.693	0.612	0.788
inter-rater agreement												
Rater couple 2–3	0.761	0.821	0.875	0.769	0.848	0.873	0.703	0.725	0.802	0.658	0.624	0.749
inter-rater agreement												
Rater 1	0.839	0.878	0.922	0.827	0.844	0.874	0.816	0.719	0.787	0.747	0.832	0.825
intra-rater agreement												
Rater 2	0.767	0.822	0.836	0.735	0.798	0.803	0.828	0.833	0.884	0.732	0.810	0.786
intra-rater agreement												
Rater 3	0.850	0.861	0.893	0.844	0.825	0.845	0.804	0.758	0.791	0.719	0.709	0.772
intra-rater agreement												

MTA, medial temporal lobe atrophy; PA, posterior atrophy.

Statistical significance was set at $P<0.05$.

Table 4. ROC analysis of visual rating scales and MRI image measures

Method	Measures	AUC (data 1)	AUC (data 2)
Visual rating scales	MTA	0.776 ± 0.044	0.822 ± 0.053
	PA	0.725 ± 0.045	0.721 ± 0.061
	MTA and PA	0.818 ± 0.041	0.824 ± 0.058
MRI Image	GM absolute volume	0.735 ± 0.043	0.785 ± 0.051
	GM relative volume	0.839 ± 0.034	0.863 ± 0.047
	GM density	0.783 ± 0.042	0.579 ± 0.059
	Combined GM measures	0.857 ± 0.034	0.957 ± 0.038

Combined GM volume measures means the combination of GM absolute volume, GM relative volume and GM density.

Table 5. The correlation between visual rating measures and cognitive assessments in aMCI.

		left MTA	right MTA	mean MTA	left PA	right PA	mean PA
Data 1							
MoCA	R	-0.207	-0.105	-0.166	-0.090	-0.064	-0.081
	P-value	0.087	0.392	0.173	0.462	0.599	0.509
AVLT-I	R	-0.293	-0.195	-0.260	-0.036	-0.131	-0.087
	P-value	**0.015***	0.108	**0.031***	0.769	0.283	0.476
AVLT-D	R	-0.265	-0.180	-0.237	-0.104	0.001	-0.054
	P-value	**0.028***	0.138	0.050	0.395	0.991	0.661
AVLT-R	R	-0.248	-0.157	-0.216	-0.075	0.035	-0.021
	P- value	**0.040***	0.197	0.075	0.541	0.776	0.864
Data 2							
MoCA-B	R	0.061	0.041	0.034	-0.085	-0.141	-0.124
	P-value	0.748	0.830	0.859	0.657	0.458	0.514
AVLT-D	R	-0.067	0.131	0.004	-0.127	0.027	-0.059
	P-value	0.723	0.491	0.982	0.505	0.887	0.758
AVLT-R	R	-0.271	-0.017	-0.234	-0.002	-0.110	-0.060
	P-value	0.148	0.930	0.213	0.992	0.561	0.752
STT-A	R	-0.118	0.103	-0.017	0.420	0.497	0.508
	P- value	0.535	0.588	0.928	**0.021***	**0.005***	**0.004***
STT-B	R	0.025	0.231	0.150	0.218	0.343	0.309
	P- value	0.897	0.219	0.430	0.248	0.063	0.097
AFT	R	0.123	0.140	0.139	0.186	-0.059	0.075
	P- value	0.517	0.459	0.463	0.326	0.755	0.693
BNT	R	0.178	0.109	0.129	-0.091	-0.224	-0.172
	P- value	0.348	0.566	0.497	0.631	0.235	0.363

*, Statistical significance was set at $P<0.05$

Figures

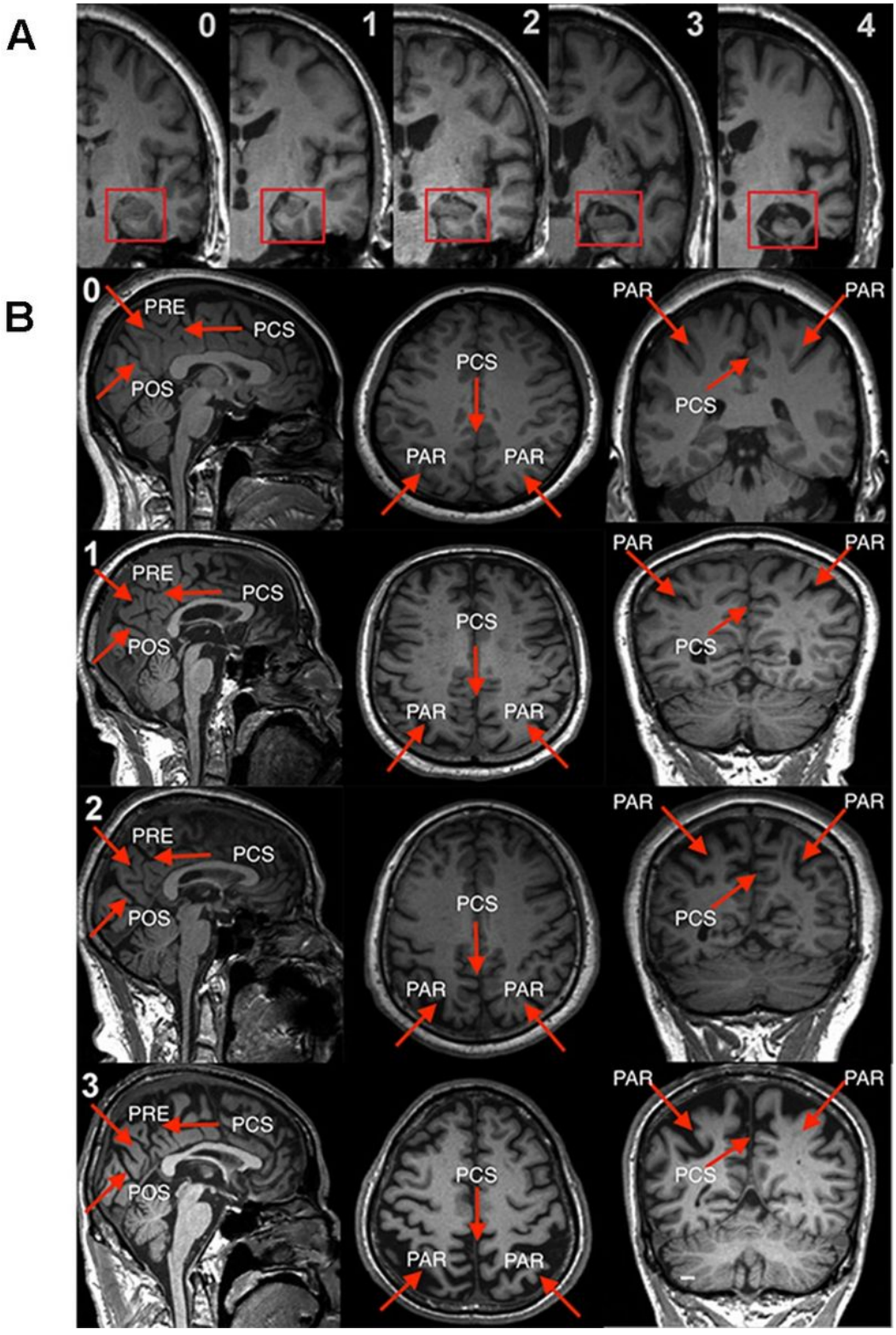

Figure 1

The rough anatomical structure of MTA (A) and PA (B) regions.

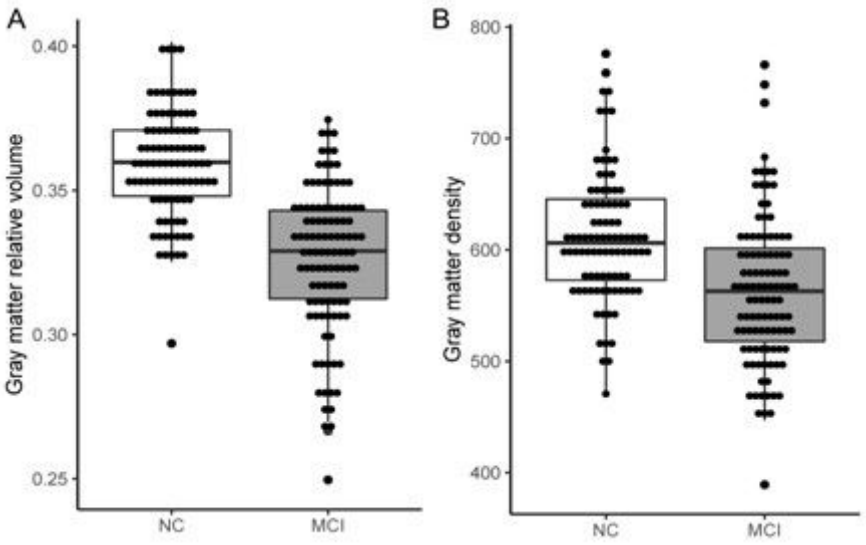

Figure 2

The lower GM volumetric measures in aMCI than controls in cohort A (P < 0.001). (A) GM relative volumes; (B) GM density.

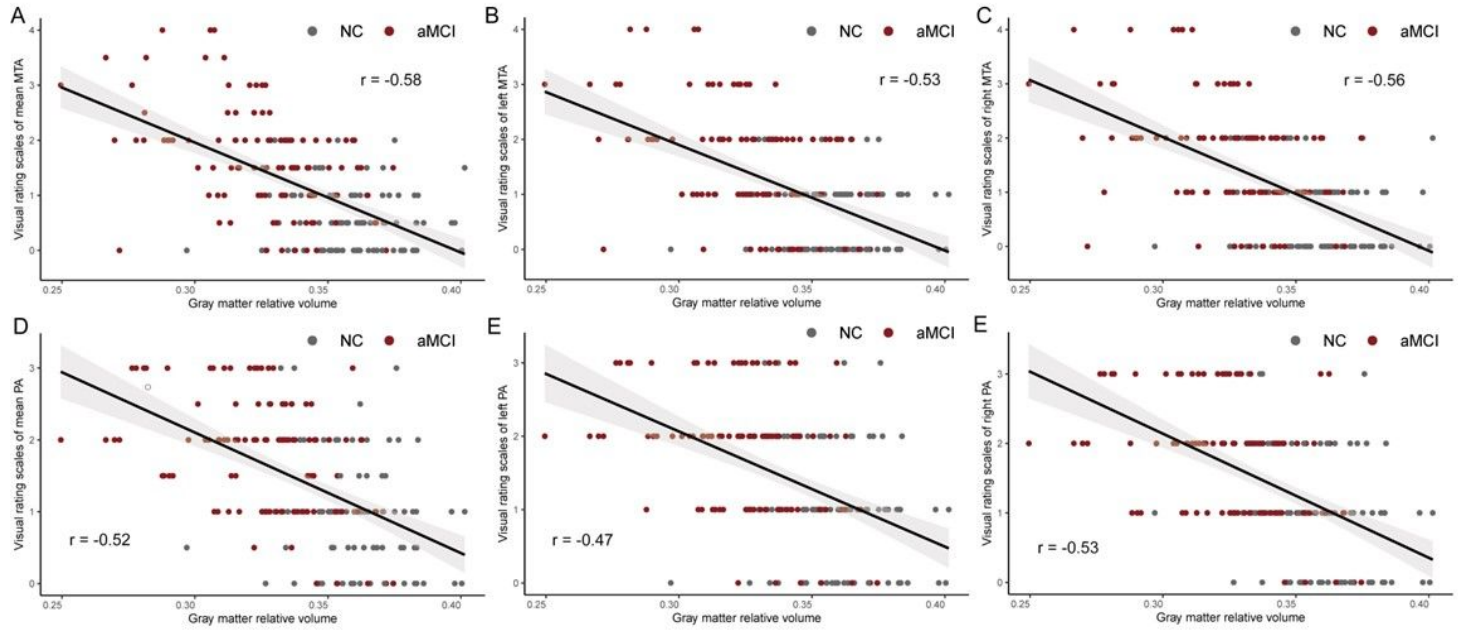

Figure 3

A negative correlation between whole brain GM measures and visual rating scores. (A) the negative correlation between GM relative volume and mean MTA; (B) the negative correlation between GM relative volume and left MTA; (C) the negative correlation between GM relative volume and right MTA; (D) the negative correlation between GM relative volume and mean PA; (E) the negative correlation between GM relative volume and left PA; (F) the negative correlation between GM relative volume and right PA.

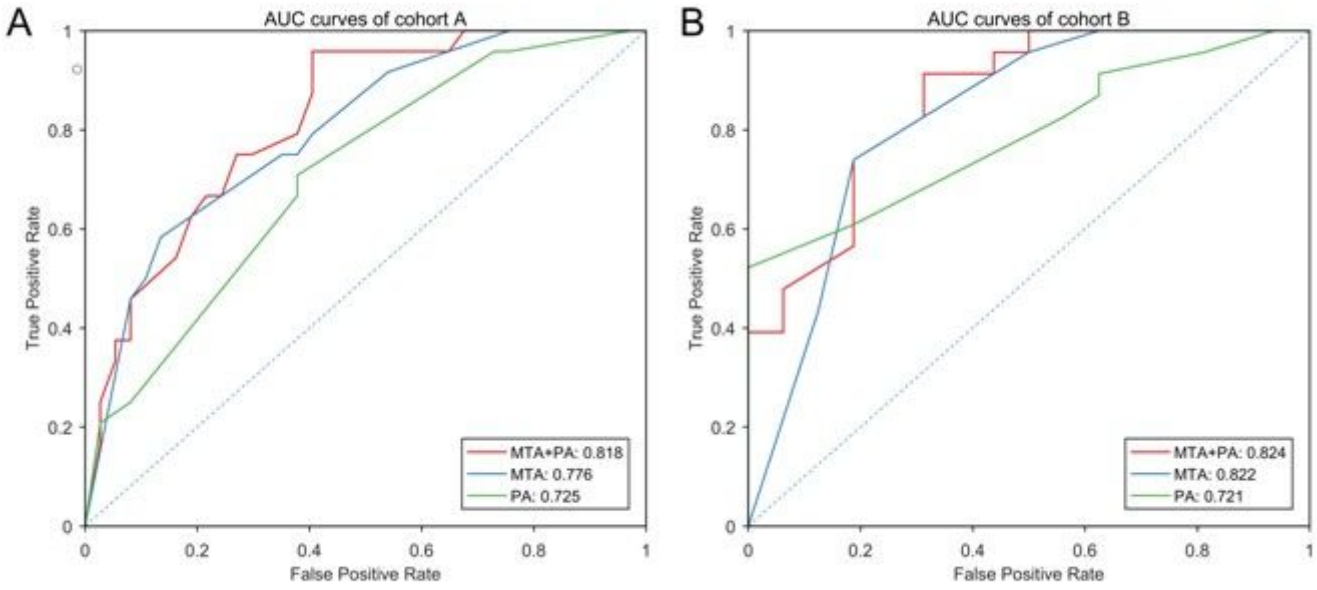

Figure 4

ROC of visual rating scales in cohort A and B. (A) ROC of single and combined visual rating scales in cohort A; (B) ROC of single and combined visual rating scales in cohort B.