

Automated Brain Volumetric Program Measuring Regional Brain Atrophy in Diagnosis of Mild Cognitive Impairment and Alzheimer's Disease Dementia

Dong-Woo Ryu

Catholic University of Korea School of Medicine <https://orcid.org/0000-0003-3390-4607>

Yun Jeong Hong

Catholic University of Korea School of Medicine

Jung Hee Cho

Catholic University of Korea School of Medicine

Kichang Kwak

Hanyang University Department of Biomedical Engineering

Jong-Min Lee

Hanyang University Department of Biomedical Engineering

Yong S Shim

Catholic University of Korea School of Medicine

Young Chul Youn

Chung-Ang University Hospital: Chung Ang University Hospital

Dong-Won Yang ([✉ neuroman@catholic.ac.kr](mailto:neuroman@catholic.ac.kr))

Catholic University of Korea School of Medicine

Research Article

Keywords: Alzheimer's Disease, Dementia, Cognitive Dysfunction, Magnetic Resonance Imaging, Computer-Assisted Image Processing

Posted Date: July 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-640508/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

A quantitative analysis of brain volume can assist in diagnosis of Alzheimer's disease (AD) usually accompanied by brain atrophy. With an automated analysis program Quick Brain Volumetry (Q BraVo) developed for volumetric measurements, we measured regional volumes and ratios to evaluate their performance in discriminating AD dementia (ADD) and mild cognitive impairment (MCI) patients from normal controls (NC). Validation of Q BraVo was based on intra-rater and inter-rater reliability with a manual measurement. The regional volumes and ratios to total intracranial volume (TIV) and to total brain volume (TBV) or total cerebrospinal fluid volume (TCV) were compared among subjects. The regional volume to total cerebellar volume ratio named Standardized Atrophy Volume Ratio (SAVR) was calculated to compare brain atrophy. Diagnostic performances to distinguish among NC, MCI, and ADD were compared between MMSE, SAVR, and the predictive model. In total, 56 NCs, 44 MCI, and 45 ADD patients were enrolled. The average run time of Q BraVo was 5 minutes 36 seconds. Intra-rater reliability was 0.999. Inter-rater reliability were high for TBV, TCV, and TIV ($R = 0.97, 0.89$ and 0.93 , respectively). The medial temporal SAVR showed the highest performance for discriminating ADD from NC ($AUC = 0.808$, diagnostic accuracy = 80.2%). The predictive model using both MMSE and medial temporal SAVR improved the diagnostic performance for MCI in NC ($AUC = 0.844$, diagnostic accuracy = 79%). Our results demonstrated Q BraVo as a fast and accurate method to measure brain volume. The regional volume calculated as SAVR could help to diagnose ADD and MCI and increase diagnostic accuracy for MCI.

Introduction

Alzheimer's disease (AD) is a neurodegenerative disease characterized by progressive decline of cognition caused by deposition of amyloid- β protein. Repeated failures of a clinical trial for disease-modifying therapy in AD emphasize the need for precise diagnosis in the early phase (Mangialasche et al. 2010). Genetic testing and pathological biomarkers for AD have advanced, but there are many limitations for clinical application. The centerpieces of AD diagnosis are clinical access, cognitive function tests, laboratory tests, and brain imaging.

Magnetic resonance (MR) brain imaging is a useful tool for diagnosing AD by identifying cerebral degeneration or other structural abnormalities that can cause a cognitive decline. The most common finding of brain imaging in AD is parieto-temporal lobe dominant cortical atrophy, which usually is accompanied by hippocampal atrophy. However, these signs can be unremarkable in early AD. In the preclinical stage, brain atrophy in specific regions can be detected before the presence of AD symptoms (Tan et al. 2014), but brain atrophy in AD can vary and be combined with other damage (Poulakis et al. 2018). Neurodegeneration due to mixed pathology interferes with a precise diagnosis of AD by imaging. A quantitative analysis of cortical thickness, connectivity, and brain volume can help to assess dementia and discriminate AD from other types of dementia (Verhooij et al. 2018). Although there are methodical difficulties and some applicatory limitations by individual, the availability of these analyses is increasing through the spread of high-resolution MR and the technical advancement of imaging analysis.

An easy, fast, and accurate measurement of brain volume can help clinicians to make an early diagnosis of AD or other degenerative diseases. Especially, regional brain volumes and their ratios can provide a clinical clue for detecting mild cognitive impairment (MCI) and AD dementia (ADD) (Risacher et al. 2017). We developed an automated program to measure regional volumes and their ratios among normal control (NC), MCI, and ADD patients using brain MRI imaging. The aim of the present study was to validate the automated program by comparisons with previous volumetric methods and to compare the regional volumes and their ratios including volume ratio to cerebellar volume among NC, MCI, and ADD groups. Finally, we evaluated the diagnostic performance of regional volume ratio alone or combined with MMSE for discrimination of MCI and ADD from NC.

Methods

Patients

We conducted a retrospective observational study from patients with mild to moderate ADD, MCI, and NC. The participants were enrolled at the dementia clinics of Seoul St. Mary's Hospital, Yeouido St. Mary's Hospital, and Chung-Ang University Hospital between January 2011 and December 2018. A diagnosis of ADD was based on the criteria of the National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association or the Diagnostic and Statistical Manual of Mental Disorder, fifth edition (DSM-5). MCI was diagnosed on the basis of the Petersen amnestic MCI criteria (Petersen 2004). Subjects with depression documented on the basis of a score greater than 7 on the short form of the Geriatric Depression Scale (GDS) were excluded.

NCs were cognitively and functionally normal, independent, and fulfilled the health-screening exclusion criteria of Christensen et al. (Christensen et al. 1990). They had MMSE scores above 1.5 SD of age- and education-adjusted norms.

Demographic information on age, sex, and educational history were collected. All enrolled subjects were evaluated with the MMSE, CDR, and GDS.

Magnetic resonance imaging

The imaging data were collected used T1-weighted (T1W) three-dimensional (3D) MR images acquired using a 1.5T MR scanner (Signa HDxt, GE Medical Systems, Milwaukee, WI, USA) or a 3T MR scanner (Philips Intera Achieva, Amsterdam, Netherlands) equipped with an 8-channel sensitivity-encoding head coil. An axial MPRAGE sequence was used to acquire T1W 3D images with the following MR parameters: TR = 1,780 ms, TE = 2.2 ms, FA = 9°, FOV = 256 × 256 × 256 mm, voxel size 1 × 1 × 1 mm, and thickness = 1.0 mm.

Automated volumetry program

The automated software for volumetric analysis of brain MR images is named Quick Brain Volumetry (QBravo) and was programmed by J.M.L. and K.C.K. QBravo, based on the 8th version of the Statistical Parametric Map (SPM8) package (Wellcome Trust Centre for Human Neuroimaging, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>) and MATLAB (The Mathworks, Inc., Natick, MA, USA), provides a simple interface to input data and quickly measure a volume of segmented brain in regional divisions.

QBravo registers native MR images into a standardized stereotaxic space using a 12-parameter transformation process called normalization. The normalized images undergo the unified segmentation procedure implemented in SPM8. In brief, the unified algorithm combines voxel intensity, imaging noise, and non-linear registration in tissue probabilistic models, being used as a priori information for the tissue classes. Then, the segmentation procedure automatically divides the structural image into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). However, the classification is probabilistic in the sense that a probability value of belonging to each of the tissue classes is assigned to each voxel in the output images. For regional division of the brain, the regional masks were manually mapped onto the MNI template. Based on neuroanatomical boundaries, the regional masks display border lines dividing the brain bilaterally into 6 frontal (orbital, anterior, anterior medial, dorsolateral, inferior, and posterior medial), 3 temporal (anterior, medial, and lateral), and 2 parietal (medial and lateral) areas as well as occipital, central, cerebellar, and ventricular areas (Fig. 1). Anatomical subdivisions of the brain have the advantage of easy access over functional sub-divisions. The volume of regional brain is the sum of tissue proportions for each region in all voxels multiplied by voxel size. Finally, QBravo calculates volumes of the 26 brain regions (GM plus WM), total brain volume (TBV), total cerebrospinal fluid volume (TCV), ventricular volume, and total intracranial volume (TIV). TBV is the sum of all GM and WM volumes, and TIV is the sum of TBV and TCV.

Though QBravo uses SPM8-based algorithms for measurements of brain volume, several factors were advanced for improved reliability and usability in clinics. A more accurate template (avgT1_Dartel_Ixi550_MNI152.nii) replaced the previous template (icbm_avg_152_t1_tal_lin.nii) used for normalization in SPM8. QBravo integrated multiple processes into an automated analytic course. Furthermore, QBravo used anatomic division of the brain that might have advantages in accuracy for regional volume analysis due to larger areas and clear borders.

Validation

Assessments of consistency and accuracy were based on intra- and inter-rater reliability of QBravo. First, the intra-rater reliability of QBravo was measured using MR images for all enrolled subjects for validation. Agreement of measured TBVs between repeated analyses for each subject was evaluated using test-retest reliability. Second, QBravo was validated by inter-rater reliability with manual volumetry, which has been considered the gold standard of volume measurements, by the Intraclass Correlation Coefficient (ICC). The manual measurements of brain volume were conducted by a skilled researcher with expertise in MR image analysis (J. H. C.). The TBV, TCV, and TIV quantities measured using QBravo were compared with those obtained using manual volumetry. For comparisons with other volumetric programs, ICCs of TBV, TCV, and TIV were measured with manual volumetry using SPM8 and FreeSurfer (The General Hospital Corporation, Boston, MA, USA).

Regional volume analysis

The regional volumes and Volume to TIV ratio (V:TIV) measured using QBravo were compared among ADD, MCI, and NC groups. The V:TIV represents atrophic degree in the region. Regional brain Volume to TBV ratio (V:TBV) and regional CSF Volume to TCV ratio (V:TCV) also were compared. Comparisons of V:TBV or V:TCV among the groups demonstrate the relative atrophic degree to total brain atrophy and can reveal regional severity of degeneration in AD. We found that the volume ratio between the most and the least

degenerated regions could help to diagnose ADD and MCI, and the standardized atrophy volume ratio (SAVR), which is a regional volume divided by that of the least degenerated region, was designed to discriminate among the groups. In this study, SAVR was used to compare regional volume with that of the cerebellum for discrimination of ADD and MCI from NC, and this calculation was similar to that for the standardized uptake value ratio (SUVR) in amyloid PET. The SAVR of the most degenerated brain region was analyzed using MMSE, which is the most commonly used screening test for general cognitive function, for diagnosis MCI or ADD. Using a combined predictive model with MMSE and SAVR, the effectiveness of SAVR was evaluated to assist in diagnosis MCI or ADD.

Statistical analysis

All statistical analyses were conducted with SPSS software version 24.0 for Windows (SPSS Incorporated, Chicago, IL, USA) and the R Statistical Environment (R Foundation for Statistical Computing, Vienna, Austria). ICCs were calculated to determine the degrees of intra-rater and inter-rater reliability for validation of QBravo. One-way analysis of variance (ANOVA) was used to compare mean differences among the 3 groups. For multiple test correction, Tukey's range test was used in subgroup comparison. Linear by linear association tests were used to compare frequencies for categorical variables. Diagnostic performance was assessed using AUC, sensitivity, specificity, and diagnostic accuracy in ROC curves. Combined predictive models for ADD vs. NC, MCI vs. NC, and ADD vs. MCI were created using a logistic regression that included MMSE and SAVR of the most degenerated region. The ROC curves were compared using the Delong method. Statistical significance was set as $p < 0.05$.

Standard Protocol Approvals, Registrations, and Patient Consents

The study protocol was approved by the institutional review board and the ethical standard committee at our institution. The ethics board determined that participant consent was not required for the retrospective observations.

Results

Patients

During the study period, 56 NCs, 44 MCI patients, and 45 ADD patients were enrolled. Clinical characteristics of the subjects are summarized in Table 1. Patients with ADD and MCI were significantly older than the NC subjects (71.8 ± 8.7 vs. 71.5 ± 6.3 vs. 67.9 ± 5.9 ; ADD vs. NC, $p = .017$; MCI vs. NC, $p = .032$). However, differences in sex distribution and educational duration were not significant among the groups.

Table 1
Clinical characteristics and measured brain volumes of subjects

	NC (n = 56)	MCI (n = 44)	ADD (n = 45)	p-value
Age, year	67.9 ± 5.9	71.5 ± 6.3	71.8 ± 8.7	A [†] B [†]
Education, year	10.5 ± 4.1	11.1 ± 4.6	11.0 ± 4.5	0.77
Sex (Male/Female)	27/29	17/27	14/31	0.08
MMSE	28.2 ± 1.6	26.2 ± 2.3	21.5 ± 5.3	A ^{††} B ^{††} C ^{††}
CDR	0.02 ± 0.09	0.48 ± 0.15	0.84 ± 0.55	A ^{††} B ^{††} C ^{††}
TIV (cm ³)	1344.3 ± 107.2	1387.7 ± 111.9	1349.6 ± 129.8	0.15
TBV (cm ³)	1134.7 ± 114.7	1158.9 ± 89.8	1109.5 ± 110.4	0.09
TBV to TIV ratio	84.3 ± 3.8	83.6 ± 3.5	82.2 ± 3.1	B ^{††}
TCV (cm ³)	209.6 ± 49.1	228.8 ± 57.6	240.1 ± 48.9	B ^{††}
TCV to TIV ratio	15.7 ± 3.8	16.4 ± 3.5	17.8 ± 3.8	B ^{††}

ADD: Alzheimer's disease dementia, MCI: mild cognitive impairment, NC: normal control, MMSE: Mini Mental Status Examination, CDR: Clinical Dementia Rating, TBV: total brain volume, TCV: total cerebrospinal fluid volume, TIV: total intracranial volume

A: MCI vs. NC, B: ADD vs. NC, C: ADD vs MCI, [†] p < 0.05, ^{††} p < 0.01

The MMSE total score were significantly different among ADD, MCI, and NC (21.5 ± 5.3, 26.2 ± 2.3, and 28.2 ± 1.6, respectively; *p* < .001). Global CDR also significantly differed among groups (ADD = 0.84 ± 0.55, MCI = 0.48 ± 0.15, NC = 0.02 ± 0.09; *p* < .001). Though patients with mild to moderate ADD were included, their mean CDR was relatively low. Among total ADD subjects, the numbers of subjects with CDR 0.5, CDR 1, and CDR 2 were 26 (57.8%), 14 (31.1%), and 5 (11.1%), respectively.

The TIV calculated by QBraVo showed no significant difference among ADD, MCI, and NC groups (1349.6 ± 129.8, 1387.7 ± 111.9, and 1344.33 ± 3.8, respectively; *p* = .15). Though TBV did not significantly differ among them (ADD: 1109.5 ± 110.4, MCI: 1158.9 ± 89.8, NC: 1134.7 ± 114.7; *p* = .09), TBV to TIV ratio was statistically smaller in ADD than in NC (ADD: 82.2 ± 3.1, MCI: 83.6 ± 3.5, NC: 84.3 ± 3.8; ADD vs. NC: *p* = .009). In addition, TCV and TCV to TIV ratio were larger in ADD than in NC.

Validation of QBraVo

The average QBraVo run time for a single data analysis was 5 minutes 36 seconds, producing quick analytic volumetric results for optimal patient care. The ICC for TBV of QBraVo to measure intra-rater reliability was 0.999 by test-retest reliability. ICCs for volumetric measurement methods were compared with the manual measurements as shown in Table 2. The ICC of QBraVo was superior to those of SPM8 and FreeSurfer in all brain locations. The inter-rater reliability of QBraVo was high, with ICCs of 0.97 in TBV, 0.89 in TCV, and 0.93 in TIV. The ICCs for SPM8 were 0.94 in TBV, 0.37 in TCV, and 0.73 in TIV, while the ICCs of FreeSurfer were 0.79 in TBV, 0.39 in TCV, and 0.70 in TIV.

Table 2

Inter-rater reliability of automated volumetric methods compared with manual measurements of brain volume

	QBraVo	SPM8	FreeSurfer
ICC_Manual in TBV	R = 0.97 [†]	R = 0.94 [†]	R = 0.79 [†]
ICC_Manual in TCV	R = 0.89 [†]	R = 0.37 [†]	R = 0.39 [†]
ICC_Manual in TIV	R = 0.93 [†]	R = 0.73 [†]	R = 0.70 [†]
ICC: Intra-class correlation coefficient, TBV: Total Brain Volume, TCV: Total Cerebrospinal Fluid Volume, TIV: Total Intracranial Volume			
[†] P < 0.001 by intraclass correlation coefficient			

Regional volume analysis

Comparisons of the regional volumes and their ratios among NC, MCI, and ADD are listed in Table 3. Medial temporal and anterior temporal volumes were significantly smaller in ADD than in NC. Ventricular volume was larger in ADD and MCI than in NC. Anterior frontal, orbital frontal, and medial temporal volume were smaller in ADD than in MCI.

Table 3
Comparisons of regional brain volumes and ratios

	NC (n = 56)			MCI (n = 44)			ADD (n = 45)			P-value		
	Volume (cm ³)	V:TIV (%)	V:TBV / V:TCV (%)	Volume (cm ³)	V:TIV (%)	V:TBV / V:TCV (%)	Volume (cm ³)	V:TIV (%)	V:TBV / V:TCV (%)	Volume	V:TIV	V:TBV / V:TCV
Anterior F	51.8 ± 6.5	3.8 ± 0.3	4.6 ± 0.2	52.9 ± 5.1	3.8 ± 0.2	4.6 ± 0.2	49.9 ± 6.1	3.7 ± 0.3	4.5 ± 0.3	C [†]	B ^{††}	0.269
Anterior medial F	79.4 ± 9.0	5.9 ± 0.4	7.0 ± 0.3	80.1 ± 7.9	5.8 ± 0.4	6.9 ± 0.3	77.0 ± 9.1	5.7 ± 0.3	6.9 ± 0.3	0.203	B [†]	0.368
Dorsal lateral F	69.2 ± 8.2	5.1 ± 0.3	6.1 ± 0.3	69.6 ± 5.6	5.0 ± 0.3	6.0 ± 0.3	67.0 ± 8.2	5.0 ± 0.3	6.0 ± 0.3	0.209	B ^{††}	0.279
Inferior F	58.6 ± 6.4	4.4 ± 0.3	5.2 ± 0.2	60.2 ± 4.4	4.3 ± 0.2	5.2 ± 0.2	57.3 ± 6.2	4.3 ± 0.3	5.2 ± 0.3	0.069	0.111	0.732
Posterior medial F	54.8 ± 6.2	4.1 ± 0.3	4.8 ± 0.2	55.5 ± 5.4	4.0 ± 0.2	4.8 ± 0.3	54.9 ± 6.2	4.1 ± 0.2	4.9 ± 0.2	0.838	0.268	B [†] C ^{††}
Orbital F	32.9 ± 4.7	2.4 ± 0.2	2.9 ± 0.2	33.3 ± 3.5	2.4 ± 0.2	2.9 ± 0.2	31.1 ± 3.6	2.3 ± 0.2	2.8 ± 0.2	C [†]	B ^{††}	0.091
Anterior T	30.7 ± 6.1	2.3 ± 0.3	2.7 ± 0.4	29.7 ± 3.8	2.1 ± 0.2	2.6 ± 0.3	28.0 ± 4.1	2.1 ± 0.2	2.5 ± 0.3	B [†]	B ^{††}	B [†]
Medial T	30.8 ± 3.3	2.3 ± 0.1	2.7 ± 0.1	30.1 ± 3.5	2.2 ± 0.2	2.6 ± 0.2	28.2 ± 3.9	2.1 ± 0.2	2.5 ± 0.2	B ^{††} C [†]	A ^{††} B ^{††}	A ^{††} B ^{††}
Lateral T	134.9 ± 13.6	10.0 ± 0.5	11.9 ± 0.5	136.6 ± 12.1	9.9 ± 0.5	11.8 ± 0.5	130.4 ± 15.0	9.7 ± 0.5	11.7 ± 0.6	0.084	B ^{††}	0.334
Lateral P	120.2 ± 11.7	8.9 ± 0.5	10.6 ± 0.4	121.9 ± 11.1	8.8 ± 0.6	10.5 ± 0.5	116.4 ± 14.7	8.6 ± 0.7	10.5 ± 0.8	0.110	B [†]	0.561
Medial P	66.7 ± 7.2	5.0 ± 0.3	5.9 ± 0.3	69.2 ± 13.1	5.0 ± 0.9	6.0 ± 0.9	65.3 ± 7.3	4.8 ± 0.3	5.9 ± 0.3	0.138	0.379	0.739
Occipital lobe	135.2 ± 16.0	10.0 ± 0.7	11.9 ± 0.5	140.3 ± 13.0	10.1 ± 0.7	12.1 ± 0.6	134.6 ± 15.5	10.0 ± 0.6	12.1 ± 0.6	0.138	0.583	0.073
Central lobe	124.6 ± 11.8	9.3 ± 0.4	11.0 ± 0.3	126.5 ± 10.8	9.1 ± 0.3	10.9 ± 0.4	121.4 ± 11.3	9.0 ± 0.4	11.0 ± 0.5	0.105	B ^{††}	0.589
Cerebellum	144.9 ± 14.4	10.8 ± 0.6	12.8 ± 0.6	153.0 ± 14.5	11.0 ± 1.0	13.2 ± 0.9	148.0 ± 15.9	11.0 ± 0.9	13.4 ± 0.8	A [†]	0.195	A [†] B ^{††}
Ventricle	66.3 ± 15.8	4.9 ± 1.2	31.9 ± 3.9	79.0 ± 23.6	5.6 ± 1.4	34.4 ± 4.5	82.9 ± 23.2	6.1 ± 1.5	34.3 ± 5.1	A ^{††} B ^{††}	A [†] B ^{††}	A [†] B [†]
Subarachnoid space	143.2 ± 36.0	10.7 ± 2.8	68.1 ± 3.9	149.8 ± 37.7	10.8 ± 2.4	65.6 ± 4.5	157.2 ± 32.2	11.7 ± 2.2	65.7 ± 5.1	0.146	0.125	A [†] B [†]

ADD: Alzheimer's disease dementia, MCI: mild cognitive impairment, NC: normal control, V:TIV: Volume to Total Intracranial volume Ratio, V:TBV: Volume to Total Brain volume Ratio, V:TCV: Volume to Total Cerebrospinal fluid volume Ratio, F: Frontal lobe, T: Temporal lobe, P: Parietal lobe; A: MCI vs. NC, B: ADD vs. NC, C: ADD vs MCI, [†] p < 0.05, ^{††} p < 0.01

Differences between the groups were more significant for V:TIV than for raw regional volume. Except for the inferior frontal, posterior medial frontal, medial parietal, occipital area, and cerebellum, V:TIVs were smaller in ADD than in NC. Medial temporal V:TIV was smaller and ventricular volume was larger in MCI than in NC.

Through comparisons of V:TBV or V:TCV, medial temporal lobe and ventricle were more degenerated than other brain regions in MCI or ADD compared to NC. Cerebellum was the least degenerated region in MCI or ADD.

Table 4 shows the performances of regional brain volumes and volume ratios, which are represented as AUCs with sensitivity and specificity at the optimal cut-off, for discriminating ADD or MCI from other groups. Although the raw regional volumes were better than V:TIV in some regions for distinguishing ADD from MCI, the discriminating performance was generally better for V:TIV than for raw regional volume. Medial temporal V:TIV had the best performance for discriminating ADD or MCI from NC (ADD vs. NC: AUC = 0.771, sensitivity = 64.4%, specificity = 83.9%; MCI vs. NC: AUC = 0.684, sensitivity = 36.4%, specificity = 94.6%). To distinguish ADD from MCI, orbital frontal lobe showed the highest performance (volume, AUC = 0.696, sensitivity = 60.0%, specificity = 75.0%).

Table 4
Discriminating performance of regional volumes and ratios among ADD, MCI, and NC

Participants	AUC	Sensitivity	Specificity	Brain region	AUC	Sensitivity	Specificity
NC vs. MCI							
Medial temporal volume	0.537	22.7%	94.6%	Ventriele volume	0.640	36.4%	94.6%
Medial temporal V:TIV	0.684	36.4%	94.6%	Ventriele V:TIV	0.628	38.6%	85.7%
Medial temporal V:TBV	0.675	45.5%	94.6%	Ventriele V:TCV	0.658	61.4%	69.6%
Anterior temporal volume	0.513	100%	16.1%	Dorsolateral frontal volume	0.456	90.9%	23.2%
Anterior temporal V:TIV	0.608	59.1%	62.5%	Dorsolateral frontal V:TIV	0.629	84.1%	48.2%
Anterior temporal V:TBV	0.609	43.2%	80.4%	Dorsolateral frontal V:TBV	0.591	50.0%	69.6%
NC vs. ADD							
Medial temporal volume	0.714	51.1%	82.1%	Lateral temporal volume	0.608	60.0%	62.5%
Medial temporal V:TIV	0.771	64.4%	83.9%	Lateral temporal V:TIV	0.687	88.9%	39.3%
Medial temporal V:TBV	0.752	53.3%	98.2%	Lateral temporal V:TBV	0.593	31.1%	91.1%
Ventriele volume	0.701	42.2%	96.4%	Anterior temporal volume	0.632	68.9%	58.9%
Ventriele V:TIV	0.714	77.8%	55.4%	Anterior temporal V:TIV	0.688	66.7%	67.9%
Ventriele V:TCV	0.633	62.2%	69.6%	Anterior temporal V:TBV	0.671	53.3%	78.6%
MCI vs. ADD							
Orbito-frontal volume	0.696	60.0%	75.0%	Sub-arachnoid space volume	0.588	75.6%	50.0%
Orbito-frontal V:TIV	0.665	64.4%	68.2%	Sub-arachnoid space V:TIV	0.634	80.0%	59.1%
Orbito-frontal V:TBV	0.620	40.0%	84.1%	Sub-arachnoid space V:TCV	0.522	66.7%	47.7%
Anterior frontal volume	0.665	75.6%	54.5%	Interior frontal volume	0.665	42.2%	88.6%
Anterior frontal V:TIV	0.635	48.9%	77.3%	Interior frontal V:TIV	0.578	24.4%	97.7%
Anterior frontal V:TBV	0.579	35.6%	81.8%	Interior frontal V:TBV	0.529	28.9%	90.9%
ADD: Alzheimer's Disease Dementia, MCI: Mild Cognitive Impairment, NC: Normal Control, V:TIV: Volume to Total Intracranial Volume ratio, V:TBV: Volume to Total Brain Volume ratio, V:TCV: Volume to Total Cerebrospinal fluid Volume ratio, AUC: Area Under the Curve, by receiver operating characteristic curves							

The discriminating performance of V:TBV or V:TCV was lower than that of V:TIV in most brain regions. However, cerebellar V:TBV had better performance than V:TIV and increased in opposite order to V:TIV, with NC < MCI < ADD. These results suggest that the cerebellum was less degenerated than other regions (cerebellar V:TBV for ADD vs. NC: AUC = 0.717, cerebellar V:TIV for ADD vs. NC: AUC = 0.431).

Comparisons between MMSE, SAVR, and the combined model with MMSE and SAVR are shown in Table 5. Though MMSE had a high diagnostic performance for ADD vs. NC (AUC: 0.941, sensitivity: 82.2%, specificity: 90.1%, cut-off: 25.5), it did not show high

performance for MCI vs. NC (AUC: 0.771, sensitivity: 65.9%, specificity: 73.2%, cut-off: 27.5) or ADD vs. MCI (AUC: 0.817, sensitivity: 82.2%, specificity: 65.9%, cut-off: 25.5). For discriminating MCI from NC, medial temporal SAVR (AUC = 0.747, diagnostic accuracy = 73%) was not superior to MMSE by comparison of the ROC curves. However, diagnostic ability of the combined model (MMSE + SAVR) was significantly superior for discriminating MCI from NC (AUC: 0.844, sensitivity: 65.9%, specificity: 94.6%) (Fig. 2.A). For ADD vs. NC, medial temporal SAVR showed lower performance than MMSE (AUC: 0.808, sensitivity: 60.0%, specificity: 96.4%). Diagnostic performance of the MMSE + SAVR combined model was the best for ADD vs. NC (AUC: 0.950, sensitivity: 86.7%, specificity: 98.1%). Although diagnostic performance of the combined model was not statistically superior to that of MMSE, its AUC and diagnostic accuracy were better than those of MMSE (Fig. 2.B). The performances for discriminating ADD from MCI were not significantly different among MMSE, orbital frontal SAVR, and the combined model. The combined model for ADD vs. MCI showed AUC: 0.855, sensitivity: 82.2%, and specificity: 81.0%.

Table 5

Comparisons of discrimination performances among MMSE, SAVR, and the combined predictive models of the most degenerated brain regions .

MMSE (A)		SAVR (B)		MMSE + SAVR (C)		Delong test	
	Diagnostic accuracy	AUC	Diagnostic accuracy	AUC	Diagnostic accuracy	AUC	
MCI vs. NC	70%	0.771	73%	0.747	79%	0.844	A/ B < C
ADD vs. NC	90%	0.941	80.2%	0.808	92.9%	0.950	B < A/C
ADD vs. MCI	74.2%	0.817	68.5%	0.703	81.6%	0.855	Not diff.

MMSE: Mini Mental Status Examination, SAVR: Standardized Atrophy Volume Ratio, ADD: Alzheimer's Disease Dementia, MCI: Mild Cognitive Impairment, NC: Normal Control, AUC: Area Under the Curve

Discussion

Brain volume has clinical implications beyond the individual characteristics of ethnicity or sex. Brain volume is closely related to aging and can reflect cognitive function or brain activity(Kiraly et al. 2016; Qing and Gong 2016). In dementia risk analysis, brain volume has been used as an indicator of cognitive reserve (van Loenhoud et al. 2018). Due to the correlation of lower brain volume with cognitive decline (Vibha et al. 2018), brain atrophy enables suspicion of dementia. Furthermore, it is hypothesized that a decrease in brain volume indicates neuropathological progression in AD (Reiter et al. 2017). Accurate longitudinal measurements of brain volume can reveal the status of a degenerating brain. In this study, we measured regional brain volume using an automated program that we developed for quick and accurate measurements and to effectively discriminate ADD and MCI from NC using medial temporal volume normalization with cerebellar volume.

It has been demonstrated that AD can be classified into four atrophy subtypes of medial temporal dominant, parieto-occipital dominant, diffuse cortical, and mild atrophy (Ten Kate et al. 2018). The hippocampal sparing variant in AD, which displays parieto-occipital dominant or mild atrophy, has been suspected to have limitation in assessment of hippocampal atrophy (Murray et al. 2011). However, hippocampal sparing AD also demonstrates medial temporal atrophy including that of the hippocampus, though its volume loss can be less than those of other subtypes (Ten Kate et al. 2018). In the present study, medial temporal volume and SAVR showed high accuracy to discriminate ADD from NC. However ,they exhibited lower performance to distinguish between MCI and NC because hippocampal sparing variants are more frequent in MCI than in AD (Ten Kate et al. 2018). However, measurement of medial temporal volume decline increased the accuracy of diagnosis of MCI patients in our study.

Medial temporal lobe atrophy and hippocampal atrophy have been reported as diagnostic markers specific to ADD (Hodson 2018). The results of this study using QBravo correspond well with those observations. The medial temporal V:TIV had the highest accuracy for diagnosing ADD, with a sensitivity of 64.4% and specificity of 83.9%. It can be hypothesized that QBravo can not only measure accurate brain volume, but also precisely conduct regional discrimination. The previous automated programs had low accuracy of measurement in the hippocampus (Guenette et al. 2018). Measurement of medial temporal volume was effectively substituted for hippocampal volume in the current study.

In this study, medial temporal lobe was the most degenerative region relative to total brain atrophy in ADD and MCI. On the contrary, the least degenerative region in ADD and MCI was the cerebellum, which showed AD pathology in the final stages in an earlier autopsy study (Braak and Braak 1991). Progression of deposition of β -amyloid occurs from the neocortex and hippocampus in the early phase to the

cerebellum in the late phase. Despite this, the cerebellum might be involved in cognition and presented reduced activation on functional imaging (Jacobs et al. 2018). A previous study has reported cerebellar GM volume to be lower in MCI than in NC (Möller et al. 2013). Further evaluations of cerebellar volume change are needed to prove the role of the cerebellum in AD.

In the validation of QBraVo, it took about 5 to 6 minutes to obtain results after input of preprocessed MR images. The runtimes of QBraVo were considerably shorter than those of the other methods. Manual volumetry commonly takes several days with a skilled analyst. FreeSurfer requires approximately 20 to 40 hours to process both hemispheres. About 8 hours is required for volume measurements using Inbrain (MIDAS Information Technology Corporation, Seongnam, Republic of Korea), which has been recently commercialized. We demonstrated that QBraVo is faster than previously used methods for volume measurements. Furthermore, QBraVo showed excellent reproducibility and relatively high accuracy that were confirmed by a significant correlation with the results of manual volumetry in the present study. This suggests QBraVo as an easy and rapid tool to measure brain volume compared to the traditional volumetric methods.

Compared with manual volumetry, the ICC was higher for TBV than for TCV or TIV. This is because the analysis of QBraVo is based on T1W images, which are limited in differentiating surface CSF from skull bone. This is similar to the results of an SPM validation in an earlier study (Heinen et al. 2016). Overestimation by SPM is caused by probabilistic segmentation that can include tissues outside the subarachnoid space in CSF (Nordenskjöld et al. 2013). Because the regional mask was created with strict limitations on the subarachnoid space, QBraVo showed improved reliability not only in TBV, but also in TCV compared to SPM and FreeSurfer. Although volume measurements in regions including surface CSF have relatively lower accuracy, QBraVo showed excellent ICCs in TCV and TIV.

In the present study, although subjects with ADD or MCI were older than the NCs, other clinical characteristics did not differ among them. Though old age can affect brain volume decrements in ADD and MCI, the effects of aging were not statistically significant by multiple regression analysis.

In the analysis of regional brain volume, discrimination performance by V:TIV was generally better than that by raw regional volume, which implies that brain volume normalized by TIV could adjust for head size variation among subjects. However, raw regional volume showed better performance than V:TIV in some regions for distinguishing between ADD and MCI, possibly because brain volume in MCI was larger than in ADD.

The V:TBV or V:TCV did not show higher performance than V:TIV for discriminating among the groups. However, performances of V:TBV/V:TCV in more or less degenerated regions were similar to that of V:TIV. In MCI and ADD, the more degenerated regions were medial temporal, anterior temporal lobe, and ventricle, and the less degenerated regions were posterior medial frontal lobe and cerebellum.

Ventricular volume and ratios also showed good performance to distinguish ADD and MCI. It has been documented that WM changes are associated with ventricular enlargement in AD (Coutu et al. 2016). Researchers have theorized that subcortical WM changes occur relatively quickly in AD. However, further investigations are needed to verify this hypothesis.

The SAVR of the most degenerated regions showed larger AUC and higher diagnostic accuracy than did V:TIV for differentiating among the groups. In comparison of ROC curves, medial temporal SAVR was inferior to MMSE for ADD vs. NC but was not inferior to MMSE for MCI vs. NC. The difference of performance between orbital frontal SAVR and MMSE was not significant for ADD vs. MCI. Accordingly, it was posited that SAVR can clarify the difference of regional atrophy between groups. Especially, SAVR was effective for discrimination of ADD and MCI from NC, which suggests that normalization of regional brain volume by cerebellar volume is more sensitive than normalization by TIV in AD spectrum neurodegenerative cognitive disorders. In an earlier study, hippocampal volume to neocortical volume ratio was a predictor of cognition and subtype in AD (Risacher et al. 2017). Volume comparisons with the cerebellum have not been used previously. The combined model using MMSE and medial temporal SAVR had higher performance for discriminating MCI from NC than did MMSE alone, which suggests that SAVR is an advantageous method to assess AD in the early phase.

Quantitative information about brain volume can improve diagnostic capabilities in a clinic. However, AD diagnosis using volumetric data alone shows decreased accuracy. Volumetric data should be added to visual assessment by radiologists, clinical information, and a neurocognitive evaluation for better results.

Although the brain volume measurements using the automated analytic program were precise in previous validations, those results could have been affected by the imaging analytic program and the MR sequence parameters (Haller et al. 2016). To overcome these

biases, strict standardization should be applied to all analysis processes. Analysis in specific regions, which can be conducted with insufficient precision, are necessary to be corrected in a regular range.

The limitations of our study are as follows. First, the number of enrolled subjects was relatively small to obtain high statistical power; therefore, the results of this study should be carefully interpreted. Second, the enrolled patients with ADD could be heterogeneous because their diagnoses were clinical without pathologic confirmation. though some subjects with ADD or MCI were confirmed with F¹⁸-florbetaben PET. Third, while subjects with a brain lesion were excluded, vascular damage was not evaluated in this study. Vascular lesions can affect not only brain volume, but also signal intensity on MR. Finally, we did not measure volume of GM, WM, or hippocampus, which can provide more specific information for AD. Further studies on these topics are needed.

Conclusions

Our results suggest that SAVR is an effective method for regional volume comparisons in AD. However, SAVR should be validated in other large populations and evaluated for diagnosis of other degenerative diseases including dementia with Lewy body and frontotemporal dementia.

In our present study, we confirmed that QBraVo was rapid, easy to use, and showed high reliability to measure brain volume. We suggest that medial temporal lobe to cerebellum volume ratio is an effective method to diagnosis ADD and MCI independently or in combination with the simple cognitive screening test MMSE.

Declarations

Author Contributions

DW Ryu: Conceptualization; Methodology; Validation; Formal analysis; Data curation, Writing - original draft preparation, review and editing; Approval of final manuscript, YJ Hong: Formal analysis; Data curation; Approval of final manuscript, JH Cho: Software; Formal analysis; Approval of final manuscript, KC Kwak: Software; Validation; Approval of final manuscript, JM Lee: Software; Approval of final manuscript, YS Shim: Formal analysis; Data curation; Approval of final manuscript, YC Youn: Formal analysis; Data curation; Approval of final manuscript, DW Yang: Conceptualization; Methodology; Software; Validation; Data curation; Writing - original draft preparation, review and editing; Approval of final manuscript.

Funding Sources

This study was supported by a grant from the Ministry of Health and Welfare (HI18C0530) and by the Health Fellowship Foundation.

Availability of data and material

The data used for the analyses is available on request.

Code availability

The software that we made for the brain volume analysis is available on request from the corresponding author.

Compliance with Ethical Standards

The study protocol was approved by the institutional review board and the ethical standard committee at our institution.

Conflicts of interest

Authors had confirmed that there is no conflict of interest to disclose.

Consent to participate

The ethics board determined that participant consent was not required for the retrospective observations.

Consent for publication

All the authors aggressed with the publication of this article.

References

1. Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, 82, 239-259.
2. Christensen, K. J., Multhaup, K. S., Nordstrom, S., & Voss, K. (1990). Cognitive test profile analysis for the identification of dementia of the Alzheimer type. *Alzheimer Disease and Associated Disorders*, 4, 96-109.
3. Coutu, J. P., Goldblatt, A., Rosas, H. D., & Salat, D. H. (2016). White Matter Changes are Associated with Ventricular Expansion in Aging, Mild Cognitive Impairment, and Alzheimer's Disease. *Journal of Alzheimer's Disease*, 49, 329-342.
4. Guenette, J. P., Stern, R. A., Tripodis, Y., Chua, A. S., Schultz, V., Sydnor, V. J., et al. (2018). Automated versus manual segmentation of brain region volumes in former football players. *Neuroimage. Clinical*, 18, 888-896.
5. Haller, S., Falkovskiy, P., Meuli, R., Thiran, J. P., Krueger, G., Lovblad, K. O., et al. (2016). Basic MR sequence parameters systematically bias automated brain volume estimation. *Neuroradiology*, 58, 1153-1160.
6. Heinen, R., Bouvy, W. H., Mendrik, A. M., Viergever, M. A., Biessels, G. J., & de Bresser, J. (2016). Robustness of Automated Methods for Brain Volume Measurements across Different MRI Field Strengths. *PLoS One*, 11, e0165719.
7. Hodson, R. (2018). Alzheimer's disease. *Nature*, 559, S1.
8. Jacobs, H. I. L., Hopkins, D. A., Mayrhofer, H. C., Bruner, E., van Leeuwen, F. W., Raaijmakers, W., et al. (2018). The cerebellum in Alzheimer's disease: evaluating its role in cognitive decline. *Brain*, 141, 37-47.
9. Kiraly, A., Szabo, N., Toth, E., Csete, G., Farago, P., Kocsis, K., et al. (2016). Male brain ages faster: the age and gender dependence of subcortical volumes. *Brain Imaging and Behavior*, 10, 901-910.
10. Mangialasche, F., Solomon, A., Winblad, B., Mecocci, P., & Kivipelto, M. (2010). Alzheimer's disease: clinical trials and drug development. *The Lancet. Neurology*, 9, 702-716.
11. Möller, C., Vrenken, H., Jiskoot, L., Versteeg, A., Barkhof, F., Scheltens, P., et al. (2013). Different patterns of gray matter atrophy in early- and late-onset Alzheimer's disease. *Neurobiology of Aging*, 34, 2014-2022.
12. Murray, M. E., Graff-Radford, N. R., Ross, O. A., Petersen, R. C., Duara, R., & Dickson, D. W. (2011). Neuropathologically defined subtypes of Alzheimer's disease with distinct clinical characteristics: a retrospective study. *The Lancet. Neurology*, 10, 785-796.
13. Nordenskjöld, R., Malmberg, F., Larsson, E. M., Simmons, A., Brooks, S. J., Lind, L., et al. (2013). Intracranial volume estimated with commonly used methods could introduce bias in studies including brain volume measurements. *Neuroimage*, 83, 355-360.
14. Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, 256, 183-194.
15. Poulakis, K., Pereira, J. B., Mecocci, P., Vellas, B., Tsolaki, M., Kloszewska, I., et al. (2018). Heterogeneous patterns of brain atrophy in Alzheimer's disease. *Neurobiology of Aging*, 65, 98-108.
16. Qing, Z., & Gong, G. (2016). Size matters to function: Brain volume correlates with intrinsic brain activity across healthy individuals. *Neuroimage*, 139, 271-278.
17. Reiter, K., Nielson, K. A., Durgerian, S., Woodard, J. L., Smith, J. C., Seidenberg, M., et al. (2017). Five-Year Longitudinal Brain Volume Change in Healthy Elders at Genetic Risk for Alzheimer's Disease. *Journal of Alzheimer's Disease*, 55, 1363-1377.
18. Risacher, S. L., Anderson, W. H., Charil, A., Castelluccio, P. F., Shcherbinin, S., Saykin, A. J., et al. (2017). Alzheimer disease brain atrophy subtypes are associated with cognition and rate of decline. *Neurology*, 89, 2176-2186.
19. Tan, C. C., Yu, J. T., & Tan, L. (2014). Biomarkers for preclinical Alzheimer's disease. *Journal of Alzheimer's Disease*, 42, 1051-1069.
20. Ten Kate, M., Dicks, E., Visser, P. J., van der Flier, W. M., Teunissen, C. E., Barkhof, F., et al. (2018). Atrophy subtypes in prodromal Alzheimer's disease are associated with cognitive decline. *Brain*, 141, 3443-3456.
21. van Loenhoud, A. C., Groot, C., Vogel, J. W., van der Flier, W. M., & Ossenkoppele, R. (2018). Is intracranial volume a suitable proxy for brain reserve? *Alzheimer's Research & Therapy*, 10, 91.
22. Vernooij, M. W., Jasperse, B., Steketee, R., Koek, M., Vroooman, H., Ikram, M. A., et al. (2018). Automatic normative quantification of brain tissue volume to support the diagnosis of dementia: A clinical evaluation of diagnostic accuracy. *Neuroimage. Clinical*, 20, 374-379.
23. Vibha, D., Tiemeier, H., Mirza, S. S., Adams, H. H. H., Niessen, W. J., Hofman, A., et al. (2018). Brain Volumes and Longitudinal Cognitive Change: A Population-based Study. *Alzheimer Disease and Associated Disorders*, 32, 43-49.

Figures

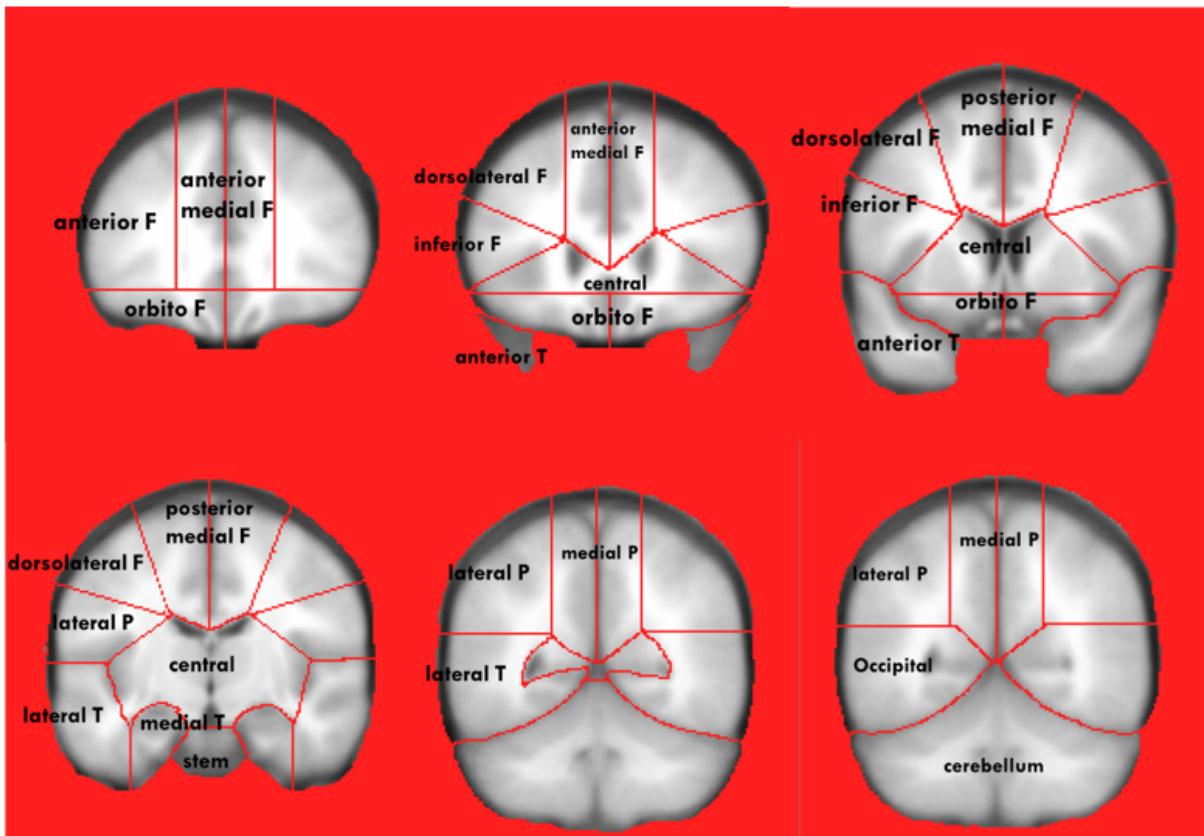


Figure 1

Regional mask with borderlines dividing the brain into 26 regions and CSF space F: Frontal lobe, T: Temporal lobe, P: Parietal lobe

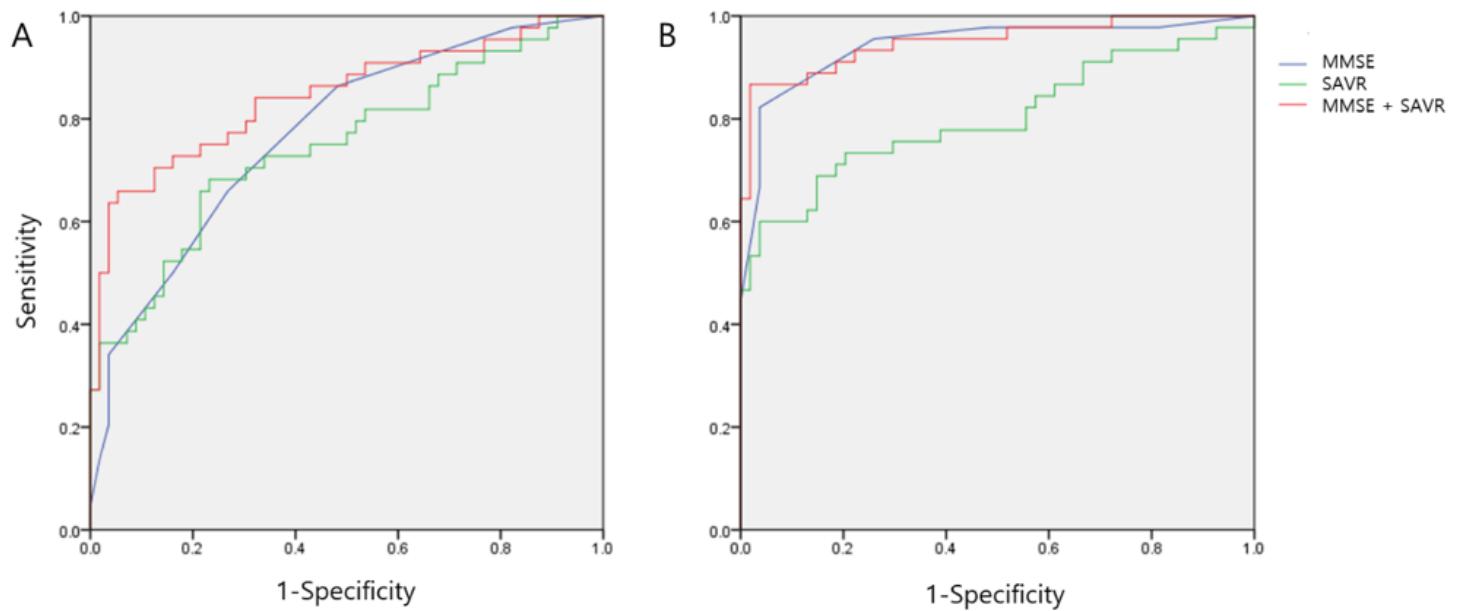


Figure 2

ROC curves of MMSE, medial temporal SAVR, and the combined prediction models for discrimination of MCI from NC (A) and of ADD from NC (B). MMSE: Mini Mental Status Examination, SAVR: Standardized Atrophy Volume Ratio, ADD: Alzheimer's Disease Dementia, MCI: Mild Cognitive Impairment, NC: Normal Control

Supplementary Files

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

- [Checklist.pdf](#)