



Discrimination of normal aging, MCI and AD with multimodal imaging measures on the medial temporal lobe

Jin Hyeong Jhoo^a, Dong Young Lee^{b,c,*}, Il Han Choo^b, Eun Hyun Seo^c, Jungsu S. Oh^b, Jae Sung Lee^d, Dong Soo Lee^d, Shin Gyeom Kim^e, Jong Chul Youn^f, Ki Woong Kim^g, Jong Inn Woo^{b,c,h}

^aDepartment of Neuropsychiatry, Kangwon National University Hospital, Chuncheon, Republic of Korea

^bDepartment of Neuropsychiatry, Seoul National University Hospital, Seoul, Republic of Korea

^cInterdisciplinary Program for Cognitive Science, Seoul National University, Seoul, Republic of Korea

^dDepartment of Nuclear Medicine, Seoul National University Hospital, Seoul, Republic of Korea

^eDepartment of Psychiatry, Soonchunhyang University Bucheon Hospital, Bucheon, Republic of Korea

^fDepartment of Neuropsychiatry, Kyunggi Provincial Hospital for the Elderly, Yongin, Republic of Korea

^gDepartment of Psychiatry, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

^hNeuroscience Research Institute of the Medical Research Center, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Article history:

Received 14 September 2009

Received in revised form 11 March 2010

Accepted 14 March 2010

Keywords:

Alzheimer's disease

Mild cognitive impairment

Hippocampal volume

Hippocampal glucose metabolism

Parahippocampal cingulum

Fractional anisotropy

ABSTRACT

This study aimed to compare the discrimination accuracy of hippocampal volume (HC-Vol), parahippocampal cingulum fractional anisotropy (PHC-FA), hippocampal glucose metabolism (HC-Glu), and any combination of the three measurements among normal control (NC), mild cognitive impairment (MCI), and Alzheimer's disease (AD). Three-dimensional MRI, diffusion tensor imaging, and FDG-PET were applied to age- and gender-matched 17 NC, 17 MCI, and 17 mild AD patients. Subjects also underwent a neuropsychological test battery including three verbal episodic memory tests. Logistic regression analyses were systematically conducted to select the best model for between-group discrimination. PHC-FA plus HC-Vol model, HC-Glu only model, and the model combining all three modalities were finally chosen for NC vs. MCI (discrimination accuracy: 79.4%), MCI vs. AD (73.5%), and NC vs. AD discrimination (94.1%), respectively. All the three imaging measures also showed significant correlation with all three episodic memory tests. These findings support that each imaging measure, respectively, and their combination have a stage-specific potential as a useful neuroimaging marker for detection and progression monitoring of early stage of AD.

© 2010 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

As the possibility of an effective disease modifying treatment against Alzheimer's disease (AD) increases (Cummings et al., 2007), accurate detection and progression monitoring of early stage of AD or its preclinical state, mild cognitive impairment (MCI) (Petersen, 2004), has recently become more important. In this context, various neuroimaging modalities have been increasingly recognized because of their potential to overcome the limitations of clinical and neuropsychological approaches (de Leon et al., 2007).

The neuropathological changes such as neurofibrillary tangle deposition, synaptic damage, and neuronal loss in AD have been known to begin in the medial temporal region encompassing the hippocampus (Braak and Braak, 1991; Ball, 1997). The hippocampus is also known to be a major neural structure related to episodic memory (Deweer et al., 2001), which is the earliest and most severely impaired

cognitive function in AD (Soininen and Scheltens, 1998). Therefore, many structural and functional neuroimaging studies for the early discrimination of AD process have focused on the hippocampus and adjacent medial temporal lobe (MTL) structures.

Numerous MRI studies consistently reported hippocampal atrophy in MCI (Pihlajamäki et al., 2009) or AD (De Santi et al., 2001; Kesslak et al., 1991; Jack et al., 1992; Chételat et al., 2008). Several recent studies have also shown hippocampal hypometabolism in mild AD or MCI (De Santi et al., 2001; Nestor et al., 2003; Mevel et al., 2007). In terms of the white matter tracts adjacent to the hippocampus or other medial temporal structures, several recent diffusion tensor imaging (DTI) studies demonstrated significantly decreased fractional anisotropy (FA), a quantitative measure sensitively reflecting the microstructural integrity of white matter fibers (Assaf and Pasternak, 2008), of the parahippocampal cingulum in MCI and AD (Zhang et al., 2007; Choo et al., 2008).

Therefore, each of these imaging measures on medial temporal structures, i.e., hippocampal volume, hippocampal glucose metabolism, and parahippocampal cingulum FA, has a potential for the early detection of preclinical or clinical AD. Several studies compare those modalities for this purpose (De Santi et al., 2001; Walhovd et al., 2009;

* Corresponding author. Department of Neuropsychiatry, Seoul National University Hospital, 28 Yongon-dong, Chongno-gu, Seoul 110-744, Republic of Korea. Tel.: +82 2 2072 2205; fax: +82 2 744 7241.

E-mail address: selfpsy@snu.ac.kr (D.Y. Lee).

Wang et al., 2009; Matsunari et al., 2007). However, most of them did not deal with all the three simultaneously for the same subjects. One recent study (Walhovd et al., 2009) investigated the combination effect of MR morphometry, PET metabolism, and DTI FA, and reported that higher diagnostic accuracy was achieved when three methods were combined, but it only focused on the differentiation MCI from normal control (NC) with no consideration of NC vs. AD or MCI vs. AD discrimination.

In this study, we compared the accuracy of hippocampal volume (HC-Vol), parahippocampal cingulum FA (PHC-FA), hippocampal glucose metabolism (HC-Glu), and any combination of the three measurements for NC vs. MCI, NC vs. MCI, and MCI vs. AD discrimination, respectively, in order to find out useful neuroimaging markers for early detection of preclinical and clinical AD. Additionally, we tried to investigate the relationship of the three imaging measures with episodic memory tests to validate these imaging measures as a tool for early AD progression monitoring.

2. Methods

2.1. Subjects

17 MCI and 17 AD patients were recruited from a cohort regularly followed at the Dementia & Age-Associated Cognitive Decline Clinic at Seoul National University Hospital. Individuals with MCI met recent criteria for amnesic MCI (9 single amnesic and 8 multi-domain amnesic MCI cases) (Windblad et al., 2004; Petersen, 2004). AD patients met both the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) criteria for dementia (American Psychiatric Association, 1994) and the National Institute of Neurological and Communication Disorders and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD (Mckhann et al., 1984). Age- and gender-matched 17 healthy normal controls were also selected from a pool of volunteers. All subjects were included after a standardized clinical assessment and neuropsychological testing, as described below.

The following exclusion criteria were applied to all subjects: any present serious medical, psychiatric, or neurological disorder that could affect mental function; evidence of focal brain lesions on MRI including lacunes; the presence of severe behavioral or communication problems that would make a clinical or MRI examination difficult; ambidextrousness or left-handedness; and the absence of a reliable informant.

The Institutional Review Board of Seoul National University Hospital, approved the study protocol and informed consent was obtained from all study subjects and their relatives.

2.2. Clinical and neuropsychological assessments

All subjects were examined by neuropsychiatrists with advanced training in neuropsychiatry and dementia research according to the protocol of the Korean Version of the CERAD Assessment Packet (Lee et al., 2002). The CERAD neuropsychological battery (Lee et al., 2002) including three verbal episodic memory tests, such as Word List Memory, Word List Recall, and Word List Recognition test, was applied by experienced clinical psychologists. A panel consisting of four neuropsychiatrists with expertise in dementia research made clinical decisions, including the assignment of Clinical Dementia Rating (CDR) index (Morris, 1993). All clinical assessments were carried out within 4 weeks of MRI and PET examination.

2.3. Image acquisition

MRI was performed using a 3.0-T GE whole body imaging system (GE VH/I; General Electric, Milwaukee, WI, USA). A dual spin-echo echo-planar imaging (EPI) sequence was used to acquire DTI images. MR

images with 25 non-collinear diffusion gradients and without diffusion gradient were acquired (TR = 10000 ms, TE = 77.1 ms, B-factor = 1000 s/mm², matrix = 128X128, slice thickness/gap = 3.5/0 mm, FOV = 240 mm, slice number = 38). A three-dimensional T1-weighted spoiled gradient recalled echo (SPGR) sequence was obtained for volumetric tracing and anatomical localization (TR = 22.0 ms, TE = 4.0 ms, slice thickness/gap = 1.4/0 mm, matrix = 256X192, FOV = 240 mm, Flip angle = 40°). Additionally, fluid-attenuated inversion recovery (FLAIR) and T2-weighted images were also obtained for qualitative clinical reading.

PET was performed using the ECAT EXACT 47 scanner (Siemens-CTI, Knoxville, TN, USA), which has an intrinsic resolution of 5.2 mm full width at half maximum (FWHM) and the images of 47 contiguous transverse planes with a 3.4 mm thickness for a longitudinal field of view of 16.2 cm. Before administering [¹⁸F] fluorodeoxyglucose (FDG), transmission scanning was performed, using three germanium-68 rod sources to correct the attenuation. Static emission scans began 30 min after the intravenous injection of 370 MBq (10 mCi) [¹⁸F] FDG and were continued for 30 min. All of the FDG-PET scans were performed in a dimly lit room with minimal auditory stimulation during both the injection and PET scanning. The subjects were in the supine position with their eyes closed during the scanning in order to minimize the confounding effects of any activity. The transaxial images were reconstructed using a filtered back-projection algorithm employing a Shepp-Logan filter with a cutoff frequency of 0.3 cycles/pixel as 128 × 128 × 47 matrices with a size of 2.1 × 2.1 × 3.4 mm.

2.4. Image processing

2.4.1. MRI volume data

The anatomical boundaries of hippocampus were traced manually on T1-weighted images using Analyze AVW 5.0 (Biomedical Imaging Resource, Mayo Foundation, Rochester, MN) by one of the authors (I.H.C.), who was blind to diagnosis, sex, or subject demographics. Sagittal SPGR sequence images were realigned to a standard orientation and reformatted using sinc interpolation to a 0.94-mm slice thickness in the axial plane. The standard alignment was based on the interhemispheric fissure, the lenses of both eyes, and the line connecting the anterior and posterior commissures in the sagittal plane. The method of tracing for hippocampus and intracranial volume (ICV) has been previously described in detail (Choo et al., 2008). In addition to hippocampus, for count normalization of PET metabolism data, the boundary of the pons was also traced on T1 weighted image as follows: On the mid-sagittal MRI section straight maximum anterior–posterior (A–P) line of pons was traced of which mid-point of A–P line was selected as center of circle. And then 7 circles including mid-sagittal section and bilateral 3 more slice sections with diameter of half length of maximum A–P line was traced for pons drawing.

2.4.2. DTI data

The details of DTI data postprocessing have been previously described in detail (Choo et al., 2008). In order to better compensate for the poor interslice resolution of DTI images, the authors interpolated the image volume along with slice direction to be spatially isotropic (Oh et al., 2007). One voxel size of the resulting DTI images was 0.94 mm × 0.94 mm × 0.94 mm. Fractional anisotropy and color-coded directionality diffusion maps were created off-line from the DTI images, using an IDL (Interactive Data Language, Research Systems Inc., Boulder, Colorado, USA) 6.0-based in-house program. In order to place the volumes of interests (VOI), FA and color-coded maps were overlaid. The color-coded directional maps (red: left-to-right direction, green: anterior-to-posterior direction, blue: superior-to-inferior direction) enable white matter fiber tracts to be easily visualized. A pair (left and right) of VOIs was placed on sagittal slices to select bilateral parahippocampal regions of the cingulum bundle (PHC) on the color-code maps of each subject (Fig. 1). To obtain

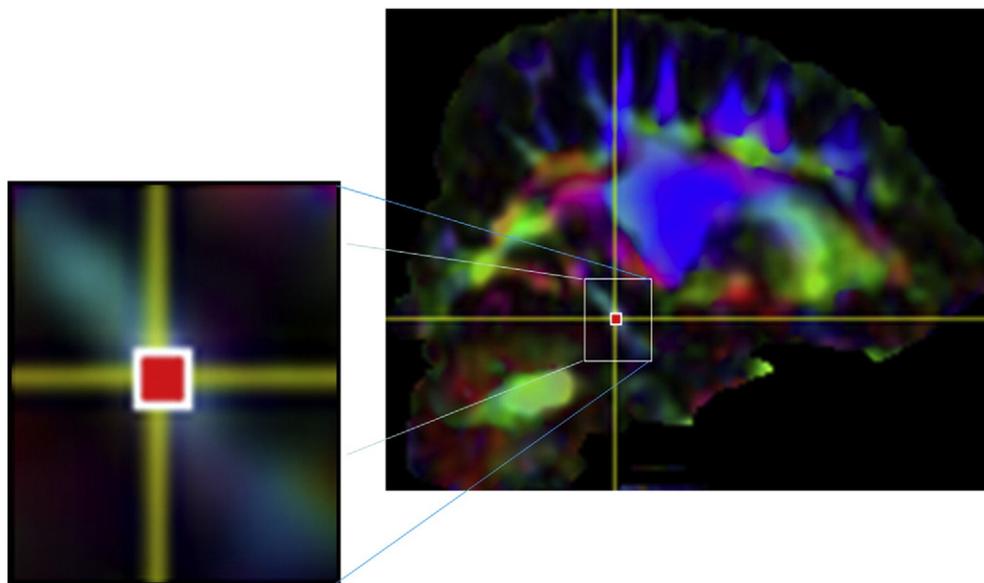


Fig. 1. Illustration of volume of interest (VOI) selection at the parahippocampal region of the cingulum on color-coded map for FA determination.

standardized conditions for analysis and to avoid data contamination by adjacent structures, all VOIs were fixed as $2.82 \times 2.82 \times 2.82 \text{ mm}^3$ (i.e., $3 \times 3 \times 3$ voxels). An experienced neuropsychiatrist (I.H.C.), blinded to subject information including diagnosis, placed the VOIs. FA values within each VOI were averaged. To determine the reliability of VOI measurements, the same rater, unaware of previous readings, repeated VOI drawings on 10 randomly selected subjects. Intrarater reliability, expressed as a mean intraclass correlation coefficient, was 0.98 for the DTI measurements.

2.4.3. FDG-PET data

Each subject PET image was coregistered with the corresponding 3D SPGR MRI image by 'coregister and reslice' option in Statistical Parametric Mapping (SPM) 99 software using 3D SPGR MRI image as a target image and PET image as an object image. Hippocampus and pons boundaries (Regions of interest: ROIs) traced on T1-weighted images using Analyze AVW 5.0 were applied to MRI coregistered PET images, and regional mean glucose metabolism were extracted by IDL-based in-house program. And then each hippocampal metabolism was count normalized by mean pons metabolism.

2.5. Statistics

The demographic and clinical data from the three groups were compared by one-way ANOVA. Raw scores of cognitive tests except MMSE were transformed to T-scores, based on normative data for the healthy elderly Korean population (Lee et al., 2004). For further analyses, each regional volume was normalized by ICV for individual subjects using the formula (absolute regional volume in mm^3/ICV in $\text{mm}^3) \times 100$. Since analyzed volumetric, FA and metabolism data did not show substantial deviation from normal distribution, parametric tests were used. Between-group differences in volumes, FA values, and metabolism levels were tested by one-way analyses of variance (ANOVAs) with Scheffe's post hoc tests. The level of statistical significance was set at two-tailed $P < 0.05$.

Logistic regression analyses were conducted to examine the ability of HC-Vol, PHC-FA, HC-Glu, and combinations of these measures to discriminate the three clinical groups. For these analyses, an average of left and right value was used for each image variable. We used the difference of -2 log likelihood ($-2LL$) to statistically compare the discrimination ability of various models with different numbers of independent variables. $-2LL$ is a quantity generated by the logistic

regression procedure and is directly proportional to the contribution of variables to the separation of groups. A smaller $-2LL$ means a better discriminative ability of the model. The probability distribution of $-2LL$ difference between simple (model 1) and more complex model (model 2) can be approximated by a chi-square distribution with $(df_2 - df_1)$ degrees of freedom, where df_1 and df_2 are the degrees of freedom of models 1 and 2 respectively. Therefore, $-2LL$ difference allows the direct statistical comparison of diagnostic models of different complexity (Hosmer, 1989). We included age as a covariate in every logistic regression model because it is known to influence on diagnostic performance of neuroimaging measures (Matsunari et al., 2007). The correlations between imaging and episodic memory measures were tested by partial correlation analyses controlling age.

3. Results

3.1. Subject characteristics

Subject demographic characteristics and the mean values of HC-Vol, PHC-FA, and HC-Glu are presented in Table 1.

3.2. Selection of diagnostic models for between-group discriminations

Before proceeding to model selection process using logistic regression analyses, we checked the intercorrelation matrix derived from the three imaging measures in order to identify any multicollinearity among them. HC-V showed significant, but weak correlation with HC-Glu (Pearson's $r = 0.295$, $P = 0.035$), while there was no significant correlation between PHC-FA and HC-Vol or between PHC-FA and HC-Glu ($r = 0.203$, $P = 0.154$; and $r = 0.268$, $P = 0.058$, respectively) across the three groups. Therefore, multicollinearity was not highly suspected between any two imaging modalities. The logistic regression analyses for model selection were conducted in three steps for every between-group differentiation (NC vs. MCI; NC vs. AD; and MCI vs. AD) (Tables 2, 3, and 4, respectively).

3.2.1. NC versus MCI

In the first step, as shown in Table 2, we tested the three one candidate models: Model V including only HC-Vol, model F including only PHC-FA, and model G including only HC-Glu as an independent variable. Among the three models, only model F was statistically

Table 1
Subject characteristics.

	NC	MCI	AD	P-value
No. (F/M)	17 (12/5)	17 (12/5)	17 (12/5)	–
Age, years	70.8 ± 5.4	70.8 ± 7.0	70.7 ± 5.8	0.999
Education, y	10.5 ± 3.3	7.7 ± 5.5	7.9 ± 5.5	0.185
CDR (0/0.5/1)	17/0/0	0/17/0	0/7/10	–
MMSE	28.4 ± 1.9	24.4 ± 3.7	16.9 ± 3.8	≤ 0.001*†‡
<i>Other CERAD</i>				
<i>Cognitive Tests</i>				
Verbal fluency	52.3 ± 10.4	45.1 ± 9.5	33.2 ± 7.1	≤ 0.001*†
Boston Naming	59.1 ± 7.1	46.9 ± 11.7	40.0 ± 16.8	≤ 0.001*†
Word List Memory	60.7 ± 12.6	36.6 ± 9.3	29.2 ± 12.9	≤ 0.001*†‡
Word List Recall	57.9 ± 9.7	34.5 ± 8.4	26.7 ± 8.3	≤ 0.001*†‡
Word List Recognition	52.4 ± 6.3	30.9 ± 12.2	15.8 ± 17.8	≤ 0.001*†‡
Constructional Praxis	52.4 ± 4.7	49.6 ± 13.0	41.5 ± 22.4	0.105
Constructional Recall	52.6 ± 9.7	41.1 ± 10.8	33.8 ± 9.8	≤ 0.001*†‡
HC-Vol	0.239 ± 0.032	0.214 ± 0.036	0.188 ± 0.031	< 0.001*
PHC-FA	0.359 ± 0.045	0.313 ± 0.042	0.298 ± 0.058	< 0.001*†
HC-Glu	0.998 ± 0.054	0.991 ± 0.080	0.917 ± 0.066	< 0.001*†

Data presented as means ± S.D. While MMSE scores are raw values, other CERAD cognitive test scores are age, education and gender-specific norm corrected *T*-scores. Group comparison by ANOVA. Scheffe's post hoc comparison of significant group differences: *NC vs. AD, †MCI vs. AD, ‡NC vs. MCI.

NC = normal controls; MCI = Mild Cognitive Impairment; AD = Alzheimer disease; CDR = clinical dementia rating; MMSE = Mini-mental state examination; HC-Vol = hippocampal volume (regional volume × 100/total intracranial volume); PHC-FA = parahippocampal cingulum fractional anisotropy; HC-Glu = hippocampal glucose metabolism (regional metabolism/pons metabolism).

significant and showed the lowest –2LL. Therefore, in the second step, we statistically tested the difference of –2LL between model F and the two candidate models which included not only PHC-FA, but also HC-Vol or HC-Glu (i.e., model FV and model FG). Model FV showed significantly lower –2LL than model F, but model FG did not. In the third step, the three candidate model (model FVG), which included all three variables, was compared to model FV. Model FVG was not statistically different from that of model FV. Because adding HC-Glu to the model FV did not make any significant improvement of discrimination ability, Model FV was selected as a final discrimination model between NC and MCI. Overall discrimination accuracy for the model was 79.4% (82.4% for NC and 76.5% for MCI).

Table 2
Results obtained from logistic regression analyses designed to select appropriate models for discrimination between NC and MCI.

Models	Variable	–2LL	χ^2	df	P-value	Diagnostic Accuracy (%)	Significance test* for –2LL difference
<i>One candidate model</i>							
Model V:	HC-Vol	42.68	4.45	2	0.108	70.6	
Model F:	PHC-FA	38.23	8.91	2	0.012	73.5	
Model G:	HC-Glu	47.05	0.08	2	0.959	50.0	
<i>Two candidate model</i>							
Model FV:	PHC-FA+ HC-Vol	30.44	16.70	3	0.001	79.4	Model FV vs. F: P = 0.005
Model FG:	PHC-FA+ HC-Glu	38.21	8.92	3	0.030	70.6	Model FG vs. F: P = 1.000
<i>Three candidate model</i>							
Model FVG:	PHC-FA+ HC-Vol+ HC-Glu	30.20	16.93	4	0.002	79.4	Model FVG vs. FV: P = 0.990

All models contain age as a covariate. Model indicated by bold characters is the finally selected one. –2LL = –2 log likelihood; NC = normal controls; MCI = Mild Cognitive Impairment; AD = Alzheimer disease; HC-Vol = hippocampal volume; PHC-FA = parahippocampal cingulum fractional anisotropy; HC-Glu = hippocampal glucose metabolism.

* *df* = 1 for every test.

Table 3
Results obtained from logistic regression analyses designed to select appropriate models for discrimination between NC and AD.

Models	Variable	–2LL	χ^2	df	P-value	Diagnostic accuracy (%)	Significance test* for –2LL difference
<i>One candidate model</i>							
Model V:	HC-Vol	30.43	16.70	2	< 0.001	82.4	
Model F:	PHC-FA	36.38	10.75	2	0.005	73.5	
Model G:	HC-Glu	33.98	13.16	2	0.001	82.4	
<i>Two candidate model</i>							
Model V:	HC-Vol+ VG:	19.46	27.67	3	< 0.001	88.2	Model VG vs. V: P = 0.009
Model V:	HC-Vol+ VF:	24.05	23.08	3	< 0.001	79.4	Model VF vs. V: P = 0.012
<i>Three candidate model</i>							
Model V:	HC-Vol+ VG:	11.22	35.91	4	< 0.001	94.1	Model VGF vs. VG: P = 0.021
	HC-Glu+ PHC-FA						

All models contain age as a covariate. Model indicated by bold characters is the finally selected one. –2LL = –2 log likelihood; NC = normal controls; MCI = mild cognitive impairment; AD = Alzheimer disease; HC-Vol = hippocampal volume; PHC-FA = parahippocampal cingulum fractional anisotropy; HC-Glu = hippocampal glucose metabolism.

* *df* = 1 for every test.

3.2.2. NC versus AD

The first step tests showed that all one candidate models were highly significant as shown in Table 3. Because model V had the lowest –2LL among them, in the second step, we tested the difference of –2LL between model V and model VG, or between model V and model VF. Both models VG and VF showed significantly lower –2LL than model V. Because model VG had relatively lower –2LL value, than model VF, in the third step, the three candidate model (Model VGF) was compared to model VG. Model VGF was significantly lower –2LL than of model VG, and model VGF was selected as a final discrimination model between NC and MCI. Overall accuracy for the model was 94.1% (94.1% for NC and 94.1% for AD).

3.2.3. MCI versus AD

The first step tests showed that only model G was significant, but model F and V were not, as shown in Table 4. In the second step, we

Table 4
Results obtained from logistic regression analyses designed to select appropriate models for discrimination between MCI and AD.

Models	Variable	–2LL	χ^2	df	P-value	Diagnostic accuracy (%)	Significance test* for –2LL difference
<i>One candidate model</i>							
Model V:	HC-Vol	42.14	4.99	2	0.082	64.7	
Model F:	PHC-FA	46.30	0.83	2	0.657	64.7	
Model G:	HC-Glu	39.12	8.02	2	0.018	73.5	
<i>Two candidate model</i>							
Model GV:	HC-Glu+ HC-Vol	35.73	11.41	3	0.010	73.5	Model GV vs. G: P = 0.069
Model GF:	HC-Glu+ PHC-FA	39.11	8.02	3	0.046	73.5	Model GF vs. G: P = 0.920
<i>Three candidate model</i>							
Model GVF:	HC-Vol+ HC-Glu+ PHC-FA	35.73	11.41	4	0.022	73.5	Model GVF vs. GV: P = 1.000

All models contain age as a covariate. Model indicated by bold characters is the finally selected one. –2LL = –2 log likelihood; NC = normal controls; MCI = Mild Cognitive Impairment; AD = Alzheimer disease; HC-Vol = hippocampal volume; PHC-FA = parahippocampal cingulum fractional anisotropy; HC-Glu = hippocampal glucose metabolism.

* *df* = 1 for every test.

compared –2LL between model G and model GV, or between model G and model GF. There was no significant difference of –2LL between model G and any of the two candidate models. Because adding HC-V and PHC-FA to the model G did not make any significant improvement of discrimination ability as well, model G was selected as a final discrimination model between MCI and AD. Overall accuracy for the model was 73.5% (70.6% for MCI and 76.5% for AD).

3.3. Relationship of imaging measures with episodic memory tests

All the three measures showed significant correlation with Word List Memory (partial correlation coefficient (r_p) = 0.526, $P \leq 0.001$ for HC-Vol; $r_p = 0.392$, $P = 0.005$ for PHC-FA; and $r_p = 0.369$, $P = 0.008$ for HC-Glu, respectively), Word List Recall ($r_p = 0.554$, $P \leq 0.001$ for HC-Vol; $r_p = 0.421$, $P = 0.002$ for PHC-FA; and $r_p = 0.346$, $P = 0.014$ for HC-Glu, respectively), and Word List Recognition scores ($r_p = 0.538$, $P \leq 0.001$ for HC-Vol; $r_p = 0.501$, $P \leq 0.001$ for PHC-FA; and $r_p = 0.341$, $P = 0.015$ for HC-Glu, respectively).

4. Discussion

4.1. NC versus MCI discrimination

When only one candidate models were considered, PHC-FA was the only significant discriminator between NC and MCI, whereas HC-Vol or HC-Glu was not. Such a superior diagnostic ability of PHC-FA for MCI, corresponding to the preclinical state of AD, is probably explained by its direct connection with the entorhinal cortex, in which the earliest neuropathological changes of AD occur (Braak and Braak, 1991). Although the cingulum is known to start from the medial temporal region (Wakana et al., 2004; Catani et al., 2002), primate studies indicated that the entorhinal cortex among the medial temporal structures is specifically connected with the cingulum (Mufson and Pandya, 1984; Amaral et al., 1984). Therefore, early neuronal loss in the entorhinal cortex is likely to contribute to PHC-FA decrease through anterograde Wallerian degeneration of the axonal fibers. A previous study from our group also showed that the FA reduction of the cingulum was highly correlated with entorhinal cortex volume (Choo et al., 2008).

Although not statistically significant ($P = 0.108$), the HC-Vol alone model show discrimination accuracy comparable to that of the PHC-FA alone model (70.6% for HC-Vol vs. 73.5% for PHC-FA). Previous reports also supported significant hippocampal atrophy in MCI (Pihlajamäki et al., 2009). Given relatively small sample size of this study, therefore, this result on the ability of HC-Vol as a single discriminator for NC vs. MCI should be cautiously interpreted.

HC-Vol also improved the diagnostic accuracy of a discrimination model when combined with PHC-FA. Moreover, for NC vs. AD discrimination, HC-Vol had better diagnostic ability than PHC-FA. These finding may be explained by published neuropathological evidences indicating that the early pathological changes of AD progress from the entorhinal cortex to hippocampus and other limbic structures (Braak and Braak, 1991). Previous volumetric studies also indicated that entorhinal cortex atrophy dominates hippocampal atrophy in MCI, whereas in mild AD hippocampal atrophy is more pronounced (Penny et al., 2004).

In contrast to HC-Vol, HC-Glu did not contribute to NC vs. MCI discrimination even when combined with PHC-FA. There are some controversies about hippocampal hypometabolism at the early stage of AD process, and a number of studies failed to demonstrate hippocampal hypometabolism (Minoshima, et al., 1997; Desgranges et al., 1998; Ibáñez et al., 1998; Herholz et al., 2002). Although several studies using stringent methodology detected genuine hippocampal hypometabolism in AD (De Santi et al., 2001; Nestor et al., 2003; Mevel et al., 2007), and even in MCI (De Santi et al., 2001; Nestor et al., 2003), a recent comparison study strongly indicated that hypometabolism is less severe than atrophy in the hippocampus of very mild AD (Chételat et al., 2008) like our finding for MCI. This less prominent hippocampal hypometabolism is probably associated with functional compensation for local neuronal damage or entorhinal disconnection occurring during the early process of AD (Chételat et al., 2008; Caroli et al., 2008; Dickerson et al., 2005; DeKosky et al., 2002).

As far as we know, this is the first study demonstrating that the PHC-FA has better diagnostic ability for MCI, than HC-Glu or even HC-Vol. In contrast to our finding, one recent MCI study comparing three imaging modalities including MRI morphometry, glucose metabolism, and white matter FA indicated that any MTL region white matter FA alone cannot discriminates MCI from NC (Walhovd et al., 2009). This discrepancy on medial temporal white matter FA seems to be related with the locations of white matter ROI. We focused homogeneously oriented specific white matter tract, whereas Walhovd et al. (2009) selected white matter just underneath cortex with a 5 mm distance limit excluding deep cortical area. Regional white matter FA values are highly variable even in normal healthy brain, depending on the baseline structure of the region, i.e., the innate homogeneity of fiber orientation within the region (Wiegell et al., 2000). The subcortical white matter area usually contains heterogeneously oriented fibers, such as U-fibers and fanning portion of specific white matter tracts. Therefore, FA values in such area are very low and even heterogeneous within a single ROI (Wiegell et al., 2000), and this can make it difficult to detect real pathological alterations.

4.2. NC versus AD discrimination

In contrast to NC vs. MCI discrimination, all the three measurements significantly contributed to NC vs. AD discrimination. Between-group comparison for mean values also demonstrated a similar finding (Table 2). Our results for individual imaging methods are generally in line with previous reports indicating hippocampal atrophy (De Santi et al., 2001; Kesslak et al., 1991; Jack et al., 1992; Chételat et al., 2008), hypometabolism (De Santi et al., 2001; Nestor et al., 2003; Mevel et al., 2007), and parahippocampal cingulum FA alteration (Zhang et al., 2007; Choo et al., 2008) in clinical AD. However, in terms of relative diagnostic accuracy for AD, our study showed that HC-Glu was equal to HC-Vol (82.4% vs. 82.4%), whereas a previous study using similar ROI approach suggested that HC-Glu is better than HC-Vol (91.0% vs. 83.0%) (De Santi et al., 2001). This discrepancy may be related with different dementia severity of subjects. We included only very mild or mild AD cases, whereas mild to moderately severe AD cases were included in De Santi et al.'s (2001) study.

4.2. NC versus AD discrimination

Compared to any one or two measurement models, the model including all the three modalities showed better diagnostic accuracy. These findings suggest that HC-Vol, HC-Glu, and PHC-FA can additively contribute to the early clinical diagnosis of AD.

4.3. MCI versus AD discrimination

Only HC-Glu showed significant discrimination ability, whereas HC-Vol and PHC-FA did not, for MCI vs. AD. There was no additive effect of HC-Vol or PHC-FA as well. This is clearly contrasted to the findings for NC vs. MCI discrimination. The results indicating significant difference of PHC-FA between NC and MCI, but not between MCI and AD is in line with previous reports (Zhang et al., 2007; Choo et al., 2008). One previous study also supported that HC-Glu is a significant classifier for MCI vs. AD, but HC-Vol is not (Zhang et al., 2007). These results for MCI vs. AD, considered together with those from NC vs. MCI discrimination, suggest that while structural alteration of the MTL is prominent even at the preclinical stage of AD, functional change of the structure gradually become conspicuous at early clinical stage. Therefore, in terms of disease detection using medial temporal imaging, structural measures are likely to be more

appropriate for preclinical stage, whereas functional measures are for early clinical stage.

This unexpected delayed functional decline may be associated with the failure of early compensatory mechanism within the hippocampus (Chételat et al., 2008). Initial cholinergic up regulation is known to be one of the major possible compensatory mechanisms in the hippocampus (DeKosky et al., 2002) and precede amyloid plaque formation (Bell and Claudio, 2006), followed by a subsequent decline as the disease evolves (DeKosky et al., 2002; Bell and Claudio, 2006). Amyloid deposition is known to be related to cholinergic suppression (Auld et al., 2002), and contrast to neurofibrillary tangle accumulation, its density in the hippocampus is very low in the initial stage of AD, and gradually increases with disease progression (Braak and Braak, 1991).

4.4. Validation of early AD progression monitoring ability

All the three imaging measures showed significant correlation with all three episodic memory tests. These results in general consistent with previous reports for each imaging measure (Deweert et al., 2001; Choo et al., 2008; Chételat et al., 2003; Müller et al., 2005) and support the potential of the three imaging measures as a surrogate marker for disease progression monitoring or treatment efficacy monitoring in early AD trials. The degree of correlation, as expressed by r_p , with the memory measures, however, was relatively weak for HC-Glu compared to other two imaging measures, indicating relative disadvantage of HC-Glu as a progression measure. This weak correlation is probably related with a mismatch between functional compensation of the hippocampus (preserved metabolism) and decline of episodic memory in MCI stage (Petersen, 2004; Soininen and Scheltens, 1998). However, referring to the result for MCI vs. AD discrimination, this weakness of HC-Glu appears not matter after MCI stage.

4.5. Limitations

Some possible limitations should be discussed. First, cautious generalization of the results are needed because of relatively small sample size and potential selection bias related to subject recruitment from tertiary hospital cohort. Second, our study is not based on longitudinal follow-up, but only based on a cross-sectional observation. Therefore, the validity of the three imaging measures as a surrogate marker for progression of early AD should be cautiously interpreted and confirmed by further longitudinal studies. Third, we did not apply partial volume effect (PVE) correction to PET data and hippocampal hypometabolism in AD might be overestimated. Given that PVE due to brain atrophy does not alter the general pattern of differences of metabolism between AD, especially mild cases, and CN (Ibáñez et al., 1998), however, it is not likely to prominently affect our results.

5. Conclusions

Our results suggest that, in terms of between-group discrimination by using multimodal imaging measures of MTL structures, stage-specific potential of each imaging measure or their combination as a useful neuroimaging marker for early detection or progression monitoring of MCI and clinical AD. Parahippocampal cingulum FA decline, especially considered with hippocampal atrophy, is probably helpful for MCI discrimination from normal aging, whereas hippocampal hypometabolism is more useful for the early discrimination of clinical AD from MCI. Moreover, all the three imaging measures, HC-Vol, HC-Glu, and PHC-FA, are likely to contribute additively to the accurate detection of mild AD. In spite of the innate limitation of a cross-sectional approach, our results also support a potential of all the

three imaging measures as a valid surrogate marker for progression monitoring for early stage of AD.

Acknowledgements

This study was supported by a grant of the Korea Healthcare Technology R&D Project, Ministry of Health, Welfare & Family Affairs, Republic of Korea (Grant No.: A070001) and by a grant from Kangwon National University.

References

- Amaral, D.G., Insausti, R., Cowan, W.M., 1984. The commissural connections of the monkey hippocampal formation. *The Journal of Comparative Neurology* 224, 307–336.
- American Psychiatric Association, 1994. *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. American Psychiatric Association, Washington DC.
- Assaf, Y., Pasternak, O., 2008. Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *Journal of Molecular Neuroscience* 34, 51–61.
- Auld, D.S., Kornecook, T.J., Bastianetto, S., Quirion, R., 2002. Alzheimer's disease and the basal forebrain cholinergic system: relations to beta-amyloid peptides, cognition, and treatment strategies. *Progress in Neurobiology* 68, 209–245.
- Ball, M.J., 1997. Neuronal loss, neurofibrillary tangles and granulovacuolar degeneration in the hippocampus with ageing and dementia. *Acta Neuropathologica* 37, 111–118.
- Bell, K.F., Claudio, C., 2006. Altered synaptic function in Alzheimer's disease. *European Journal of Pharmacology* 545, 11–21.
- Braak, H., Braak, E., 1991. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica* 82, 239–259.
- Caroli, A., Geroldi, C., Nobili, F., Barnden, L.R., Guerra, U.P., Bonetti, M., Frisoni, G.B., 2010. Functional compensation in incipient Alzheimer's disease. *Neurobiology of Aging* 31, 387–397.
- Catani, M., Howard, R.J., Pajevic, S., Jones, D.K., 2002. Virtual in vivo interactive dissection of white matter fasciculi in the human brain. *NeuroImage* 17, 77–94.
- Chételat, G., Desgranges, B., de la Sayette, V., Viader, F., Berkouk, K., Landeau, B., Lalevee, C., Le Doze, F., Dupuy, B., Hannequin, D., Baron, J.C., Eustache, F., 2003. Dissociating atrophy and hypometabolism impact on episodic memory in mild cognitive impairment. *Brain* 126, 1955–1967.
- Chételat, G., Desgranges, B., Landeau, B., Mezenge, F., Poline, J.B., de la Sayette, V., Viader, F., Eustache, F., Baron, J.C., 2008. Direct voxel-based comparison between grey matter hypometabolism and atrophy in Alzheimer's disease. *Brain* 131, 60–71.
- Choo, I.H., Lee, D.Y., Oh, J.S., Lee, J.S., Lee, D.S., Song, I.C., Youn, J.C., Kim, S.G., Kim, K.W., Jhoo, J.H., Woo, J.L., 2010. Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging* 32, 772–779.
- Cummings, J.L., Doody, R., Clark, C., 2007. Disease-modifying therapies for Alzheimer's disease: challenges to early intervention. *Neurology* 69, 1622–1634.
- de Leon, M.J., Mosconi, L., Blennow, K., De Santi, S., Zinkowski, R., Mehta, P.D., Pratico, D., Tsui, W., Saint Louis, L.A., Sobanska, L., Brys, M., Li, Y., Rich, K., Rinne, J., Rusinek, H., 2007. Imaging and CSF studies in the preclinical diagnosis of Alzheimer's disease. *Annals of the New York Academy of Sciences* 1097, 114–145.
- De Santi, S., de Leon, M.J., Rusinek, H., Convit, A., Tarshish, C.Y., Roche, A., Tsui, W.H., Kandil, E., Boppana, M., Daisley, K., Wang, G.J., Schlyer, D., Fowler, J., 2001. Hippocampal formation glucose metabolism and volume losses in MCI and AD. *Neurobiology of Aging* 22, 529–539.
- DeKosky, S.T., Ikonomic, M.D., Styren, S.D., Beckett, L., Wisniewski, S., Bennett, D.A., Cochran, E.J., Kordower, J.H., Mufson, E.J., 2002. Upregulation of choline acetyltransferase activity in hippocampus and frontal cortex of elderly subjects with mild cognitive impairment. *Annals of Neurology* 51, 145–155.
- Desgranges, B., Baron, J.C., de la Sayette, V., Petit-Taboué, M.C., Benali, K., Landeau, B., Lechevalier, B., Eustache, F., 1998. The neural substrates of memory systems impairment in Alzheimer's disease. A PET study of resting brain glucose utilization. *Brain* 121, 611–631.
- Deweert, B., Pillon, B., Pochon, J.B., Dubois, B., 2001. Is the HM story only a "remote memory"? Some facts about hippocampus and memory in humans. *Behavioural Brain Research* 127, 209–224.
- Dickerson, B.C., Salat, D.H., Greve, D.N., Chua, E.F., Rand-Giovannetti, E., Rentz, D.M., Bertram, L., Mullin, K., Tanzi, R.E., Blacker, D., Albert, M.S., Sperling, R.A., 2005. Increased hippocampal activation in mild cognitive impairment compared to normal aging and AD. *Neurology* 65, 404–411.
- Herholz, K., Schopphoff, H., Schmidt, M., Mielke, R., Eschner, W., Scheidhauer, K., Schicha, H., Heiss, W.D., Ebmeier, K., 2002. Direct comparison of spatially normalized PET and SPECT scans in Alzheimer's disease. *Journal of Nuclear Medicine* 43, 21–26.
- Hosmer, D., 1989. *Applied Logistic Regression*. John Wiley & sons, Toronto.
- Ibáñez, V., Pietrini, P., Alexander, G.E., Furey, M.L., Teichberg, D., Rajapakse, J.C., Rapoport, S.I., Schapiro, M.B., Horwitz, B., 1998. Regional glucose metabolic abnormalities are not the result of atrophy in Alzheimer's disease. *Neurology* 50, 1585–1593.
- Jack Jr., C.R., Petersen, R.C., O'Brien, P.C., Tangalos, E.G., 1992. MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease. *Neurology* 42, 183–188.
- Kesslak, J.P., Nalcioglu, O., Cotman, C.W., 1991. Quantification of magnetic resonance scans for hippocampal and parahippocampal atrophy in Alzheimer's disease. *Neurology* 41, 51–54.

- Lee, D.Y., Lee, K.U., Lee, J.H., Kim, K.W., Jhoo, J.H., Kim, S.Y., Youn, J.C., Woo, S.I., Ha, J., Woo, J.I., 2004. A normative study of the CERAD neuropsychological assessment battery in the Korean elderly. *Journal of the International Neuropsychological Society* 10, 72–81.
- Lee, J.H., Lee, K.U., Lee, D.Y., Kim, K.W., Jhoo, J.H., Kim, J.H., Lee, K.H., Kim, S.Y., Han, S.H., Woo, J.I., 2002. Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): Clinical and neuropsychological assessment battery. *Journal Gerontol Psychol Science* 57, 47–53.
- Matsunari, I., Samuraki, M., Chen, W.P., Yanase, D., Takeda, N., Ono, K., Yoshita, M., Matsuda, H., Yamada, M., Kinuya, S., 2007. Comparison of 18F-FDG PET and optimized voxel-based morphometry for detection of Alzheimer's disease: aging effect on diagnostic performance. *Journal of Nuclear Medicine* 48, 1961–1970.
- Mckhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E.M., 1984. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task force on Alzheimer's disease. *Neurology* 34, 939–944.
- Mevel, K., Desgranges, B., Baron, J.C., Landeau, B., De la Sayette, V., Viader, F., Eustache, F., Chételat, G., 2007. Detecting hippocampal hypometabolism in mild cognitive impairment using automatic voxel-based approaches. *Neuroimage* 37, 18–25.
- Minoshima, S., Giordani, B., Berent, S., Frey, K.A., Foster, N.L., Kuhl, D.E., 1997. Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Annals of Neurology* 42, 85–94.
- Morris, J.C., 1993. The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414.
- Mufson, E.J., Pandya, D.N., 1984. Some observations on the course and composition of the cingulum bundle in the rhesus monkey. *The Journal of Comparative Neurology* 225, 31–43.
- Müller, M.J., Greverus, D., Dellani, P.R., Weibrich, C., Wille, P.R., Scheurich, A., Stoeter, P., Fellgiebel, A., 2005. Functional implications of hippocampal volume and diffusivity in mild cognitive impairment. *NeuroImage* 28, 1033–1042.
- Nestor, P.J., Fryer, T.D., Smielewski, P., Hodges, J.R., 2003. Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Annals of Neurology* 54, 343–351.
- Oh, J.S., Song, I.C., Lee, J.S., Kang, H., Park, K.S., Kang, E., Lee, D.S., 2007. Tractography-guided statistics (TGIS) in diffusion tensor imaging for the detection of gender difference of fiber integrity in the midsagittal and parasagittal corpora callosa. *NeuroImage* 36, 606–616.
- Pennanen, C., Kivipelto, M., Tuomainen, S., Hartikainen, P., Hänninen, T., Laakso, M.P., Hallikainen, M., Vanhanen, M., Nissinen, A., Helkala, E.L., Vainio, P., Vanninen, R., Partanen, K., Soininen, H., 2004. Hippocampus and entorhinal cortex in mild cognitive impairment and early AD. *Neurobiology of Aging* 25, 303–310.
- Petersen, R.C., 2004. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine* 256, 183–194.
- Pihlajamäki, M., Jauhiainen, A.M., Soininen, H., 2009. Structural and functional MRI in mild cognitive impairment. *Current Alzheimer Research* 6, 179–185.
- Soininen, H.S., Scheltens, P., 1998. Early diagnostic indices for the prevention of Alzheimer's disease. *Annali Medici* 30, 553–559.
- Wakana, S., Jiang, H., Nagae-Poetscher, L.M., van Zijl, P.C.M., Mori, S., 2004. Fiber tract-based atlas of human white matter anatomy. *Radiology* 230, 77–87.
- Walhovd, K.B., Fjell, A.M., Amlie, I., Grambaite, R., Stenset, V., Bjørnerud, A., Reinvang, I., Gjerstad, L., Cappelen, T., Due-Tønnessen, P., Fladby, T., 2009. Multimodal imaging in mild cognitive impairment: Metabolism, morphometry and diffusion of the temporal-parietal memory network. *Neuroimage* 45, 215–223.
- Wang, L., Goldstein, F.C., Veledar, E., Levey, A.I., Lah, J.J., Meltzer, C.C., Holder, C.A., Mao, H., 2009. Alterations in cortical thickness and white matter integrity in mild cognitive impairment measured by whole-brain cortical thickness mapping and diffusion tensor imaging. *American Journal of Neuroradiology* 30, 893–899.
- Wiegell, M.R., Larsson, H.B.W., Wedeen, V.J., 2000. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology* 217, 897–903.
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.O., et al., 2004 Sep. Mild cognitive impairment—beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *Journal of Internal Medicine* 256 (3), 240–246.
- Zhang, Y., Schuff, N., Jahng, G.H., Bayne, W., Mori, S., Schad, L., Mueller, S., Du, A.T., Kramer, J.H., Yaffe, K., Chui, H., Jagust, W.J., Miller, B.L., Weiner, M.W., 2007. Diffusion tensor imaging of cingulum fibers in mild cognitive impairment and Alzheimer disease. *Neurology* 68, 13–19.