Neuropsychologia xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease

Yu-Ling Chang[a,*], Mark W. Bondi[a,d], Christine Fennema-Notestine[a,b], Linda K. McEvoy[b], Donald J. Hagler Jr.[b], Mark W. Jacobson[a,d], Anders M. Dale[b,c], the Alzheimer's Disease Neuroimaging Initiative[1]

[a] Department of Psychiatry, University of California, San Diego, 8950 Villa La Jolla Drive Suite C101, La Jolla, CA 92037, United States
[b] Department of Radiology, University of California, San Diego, CA, United States
[c] Department of Neurosciences, University of California, San Diego, La Jolla, CA, United States
[d] Veterans Affairs San Diego Healthcare System, San Diego, CA, United States

ARTICLE INFO

Article history:
Received 4 September 2009
Received in revised form 9 December 2009
Accepted 16 December 2009
Available online xxx

Keywords:
Amnestic MCI
Early detection
Episodic memory
Longitudinal outcome
MR morphometry

ABSTRACT

Understanding the underlying qualitative features of memory deficits in mild cognitive impairment (MCI) can provide critical information for early detection of Alzheimer's disease (AD). This study sought to investigate the utility of both learning and retention measures in (a) the diagnosis of MCI, (b) predicting progression to AD, and (c) examining their underlying brain morphometric correlates. A total of 607 participants were assigned to three MCI groups (high learning–low retention; low learning–high retention; low learning–low retention) and one control group (high learning–high retention) based on scores above or below a 1.5 SD cutoff on learning and retention indices of the Rey Auditory Verbal Learning Test. Our results demonstrated that MCI individuals with predominantly a learning deficit showed a widespread pattern of gray matter loss at baseline, whereas individuals with a retention deficit showed more focal gray matter loss. Moreover, either learning or retention measures provided good predictive value for longitudinal clinical outcome over two years, although impaired learning had modestly better predictive power than impaired retention. As expected, impairments in both measures provided the best predictive power. Thus, the conventional practice of relying solely on the use of delayed recall or retention measures in studies of amnestic MCI misses an important subset of older adults at risk of developing AD. Overall, our results highlight the importance of including learning measures in addition to retention measures when making a diagnosis of MCI and for predicting clinical outcome.

1. Introduction

Memory deficits are one of the hallmark features of Alzheimer's disease (AD) and are regarded as essential for the diagnosis (see Salmon & Bondi, 2009, for discussion). Information processing models provide evidence of three distinct processes involved in memory: encoding, retention, and retrieval of information (Lucas, 2005). Encoding or learning is the process by which information is acquired and transformed into a stored mental representation. Retention refers to the process by which the encoded information is maintained over time in the absence of active rehearsal. Studies of learning and memory in AD have found deficits in both learning and retention (i.e., accelerated forgetting) of episodic material, although there is not consensus that both are impaired in the earliest stages of AD. Some researchers argue that prodromal AD is characterized predominantly by an acquisition deficit (Greene, Baddeley, & Hodges, 1996; Grober & Kawas, 1997; Weingartner et al., 1981), whereas others put greater emphasis on a deficit in retention (Hart, Kwentus, Harkins, & Taylor, 1988; Moss, Albert, Butters, & Payne, 1986).

Lesion and functional imaging studies have suggested that learning and retention processes often are correlated but also show some independence and reflect different underlying neural processes (Moulin, James, Freeman, & Jones, 2004). For instance, delayed recall and/or retention tasks are primarily based on long-term memory (LTM) with critical involvement of the medial temporal lobe (MTL), including hippocampus and entorhinal cortex (Leube, Erb, Grodd, Bartels, & Kircher, 2001; Moscovitch et al., 2005; Parsons, Haut, Lemieux, Moran, & Leach, 2006; Powell et al.,

* Corresponding author. Tel.: +1 858 246 0942; fax: +1 858 534 1078.
E-mail address: yuling@ucsd.edu (Y.-L. Chang).

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu\ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. Complete listing of ADNI investigators available at http://www.loni.ucla.edu/ADNI/Data/ADNI_Authorship_List.pdf.

0028-3932/$ – see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.neuropsychologia.2009.12.024

2005; Squire, Stark, & Clark, 2004; Strange, Otten, Josephs, Rugg, & Dolan, 2002; Weintrob, Saling, Berkovic, & Reutens, 2007). On the other hand, learning tasks, often assessed by performance on immediate recall of story material or word lists, do not rely on LTM but are considered dependent on working memory processes (i.e., the phonological loop and episodic buffer; Baddeley, 2000). Learning often involves widely distributed neural substrates, including medial temporal, frontal, and parietal cortices (Axmacher, Schmitz, Weinreich, Elger, & Fell, 2008; Cabeza & Nyberg, 2000; Fujii et al., 2002; Hannula & Ranganath, 2008; Leube et al., 2008; Mayes & Montaldi, 1999).

Mild cognitive impairment (MCI) is well established as a risk state for the development of AD (see Petersen et al., 2001, for discussion) and, since its inception (Petersen et al., 1999), the definition has required a deficit in objective memory, which has been overwhelmingly interpreted as a retention deficit. Indeed, a large number of studies have shown that a decrement in episodic memory, particularly on measures of delayed recall, is a strong predictor of future AD (Albert, Moss, Tanzi, & Jones, 2001; Arnaiz & Almkvist, 2003; Backman, Jones, Berger, Laukka, & Small, 2005; Bondi et al., 1994; Bondi, Salmon, Galasko, Thomas, & Thal, 1999; Grober et al., 2008; Grober & Kawas, 1997; Twamley, Ropacki, & Bondi, 2006). Not surprisingly, studies of MCI have relied almost exclusively on delayed recall or retention measures in diagnosis (Arnaiz & Almkvist, 2003), and more recent conceptualizations of MCI continue to rely on the retention deficit in classifying whether an individual has an 'amnestic' or 'non-amnestic' form of the disorder (Petersen & Morris, 2005). Although retention measures have undoubtedly proven to be useful in MCI diagnosis and prodromal AD detection (Arnaiz & Almkvist, 2003), it is still an open question whether learning measures are useful as well. Furthermore, the corresponding brain morphometric changes in older adults with impaired learning, retention, or both abilities in MCI are poorly characterized.

Thus, in the present study of publicly available data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), we classified participants as MCI versus normally aging based on their learning and retention performances on a commonly used verbal memory test (Rey Auditory Verbal Learning Test; Rey, 1941) and then examined their brain morphometry. The RAVLT has enjoyed widespread use in clinical neuropsychological assessment of older adults as a sensitive measure of word list learning and memory, and the Mayo Older Americans Normative Studies (MOANS; Ivnik et al., 1992) provide some of the best normative reference standards for the demographic adjustment of age and education effects on RAVLT test performance. Furthermore, the RAVLT and its MOANS complement of normative data provide two summary indices of learning and retention (see below for details) for use in the present study. With these measures, we predicted that the brain morphometry of MCI individuals with predominantly retention deficits could be differentiated from MCI individuals with predominantly learning deficits. Specifically, we predicted that MCI individuals with retention deficits (either with or without learning impairment) would demonstrate circumscribed atrophy in mesial temporal regions (i.e., smaller hippocampal volumes; reduced cortical thickness in entorhinal and/or parahippocampal areas) relative to individuals with intact retention ability (Leube et al., 2001; Moscovitch et al., 2005; Parsons et al., 2006; Powell et al., 2005; Squire et al., 2004; Strange, Otten, Josephs, Rugg, & Dolan, 2002; Weintrob et al., 2007). In contrast, MCI individuals with learning deficits (either with or without retention impairment) would show a more widespread pattern of cortical thinning involving frontal and parietal regions, in addition to the mesial temporal regions, than individuals with intact learning ability (Axmacher et al., 2008; Fujii et al., 2002; Hannula & Ranganath, 2008; Leube et al., 2008; Mayes & Montaldi, 1999).

We further examined the two-year clinical outcome of these participants with the goal of identifying predictors of progression to dementia related to initial learning and retention performance. We predicted that MCI individuals with either learning or retention deficits would have a higher risk of developing AD compared to individuals without learning and retention deficits. A meta-analytic study of the cognitive impairments in prodromal AD by Backman et al. (2005) supports the sensitivity of both learning and retention measures for predicting AD progression, although in their meta-analysis delayed recall ($d = 1.23$) surpassed immediate recall ($d = 0.96$). However, based on evidence that MCI individuals with more widespread gray matter loss at baseline have been shown to progress more rapidly to AD relative to those with focal gray matter loss (McEvoy et al., 2009; Whitwell et al., 2008), we predicted that MCI individuals with impaired learning ability would show a more widespread pattern of cortical atrophy and be more prone to develop AD in the longitudinal follow-up than individuals with retention deficits only.

2. Methods

The raw data used in the current study were obtained from the ADNI database (www.loni.ucla.edu\ADNI). The ADNI was launched in 2003 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies and non-profit organizations, as a $60 million, five-year public–private partnership. ADNI's goal is to test whether serial magnetic resonance imaging (MRI), positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. The Principal Investigator of this initiative is Michael W. Weiner, M.D., VA Medical Center and University of California – San Francisco. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations. Participants have been recruited from over 50 sites across the U.S. and Canada (see www.adni-info.org). This study was approved by an ethical standards committee on human experimentation at each institution. Written informed consent was obtained from all participants or authorized representatives participating in the study. The study is conducted in compliance with Health Insurance Portability and Accountability Act regulations.

2.1. Participants

ADNI general eligibility criteria are described at http://www.adni-info.org/index.php?option=com_content&task=view&id=9& Itemid=43. Briefly, participants were 55–90 years old, non-depressed, with a modified Hachinski score of 4 or less, and had a study partner able to provide an independent evaluation of functioning. Healthy control participants had a Clinical Dementia Rating (CDR; Hughes, Berg, Danziger, Coben, & Martin, 1982) score of 0. Participants classified as MCI within ADNI had a subjective memory complaint, objective memory loss measured by education-adjusted scores on modified Wechsler Memory Scale Logical Memory II (LM II), a CDR score of 0.5, preserved activities of daily living, and an absence of dementia (Petersen et al., 2001). Though not used in the present study, AD subjects in ADNI had MMSE scores between 20 and 26, global CDR of 0.5 or 1.0, and met NINCDS/ADRDA criteria for probable AD (McKhann et al., 1984). The present study used data collected prior to March 2009, and only individuals who were classified by the ADNI criteria as healthy control (HC) or MCI at baseline were included ($n = 616$). Nine (5 HC and 4 MCI) of the total 616 participants were excluded from the study due to missing data on the verbal memory measures. Due to exclusion of MR images that did not pass local quality control, baseline MR morphometric data at baseline were available for 551 of the 607 individuals. Two-year follow-up clinical outcome data (i.e., progression to AD) were available for 423 participants.

Classification of individuals as HC or MCI in ADNI was based on education-adjusted scores on the modified LM II. However, the modified LM II score does not provide the same level of information about each individual's learning and retention abilities as does the Rey Auditory Verbal Learning Test (RAVLT). Moreover, education-adjusted scores can potentially result in misclassification for some borderline cases if age is not taken into account. Since the purpose of the current study was to examine the relative utility of learning versus retention measures in predicting progression to AD and the underlying brain substrate correlates, we re-classified all participants ($n = 607$) into one of four subgroups based on their scores on the well-established MOANS learning and retention indices (Ivnik et al., 1992) of the RAVLT (see below for details) irrespective of their classification as HC or MCI within ADNI. Here the three learning–retention impaired groups as defined below were identified as MCI groups; the group with intact learning and retention abilities was identified as the control group.

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

2.2. Learning/retention group assignment

All participants were divided into four groups based on their performance on the RAVLT, which was administered as part of a larger battery of neuropsychological tests. The RAVLT, a 15-item list-learning task, was presented verbally over five trials and participants were asked to recall as many words as possible after each trail. The first trial represents immediate word span (Trial 1). After Trial 5, a new list of words was presented (considered an interference list) and free recall of the new list (List B) was elicited. Immediately afterwards the participant was asked to recall items from the first list (short delayed recall). Long delay free recall and a recognition trial were given following a 20-min delay period.

These RAVLT component scores (e.g., Trial 1 score, short delayed recall) were converted to MOANS age-corrected scaled scores (AcSS, mean = 10, standard deviation = 3) (Ivnik et al., 1992). Summary indices for the RAVLT were then obtained and calculated in a fashion consistent with the variables provided for the MOANS norms (Harris, Ivnik, & Smith, 2002; Ivnik et al., 1992). The summary indices are as follows: (1) learning over trials (LOT) = the sum of words remembered across Trials 1–5, corrected for immediate word span (Trial 1); (2) short-term percent retention (STPR) = short delayed recall expressed as a proportion of Trial 5 recall; (3) long-term percent retention (LTPR) = long delayed recall score expressed as a proportion of Trial 5 recall. In addition to the above summary scores, two RAVLT indices were derived by grouping and summing MOANS AcSS that reflect learning efficiency and percent retention (Ivnik et al., 1992). These indices were expressed as standard scores (mean = 100, standard deviation = 15) and were derived in the following manner:

1. Learning Efficiency Index (LEI): This index is derived from the summation of AcSS scores for Trial 1 (which reflects immediate word span) and LOT (which reflects the ability to improve beyond immediate word span).
2. Percent Retention Index (PRI): This summary index reflects the amount of data remembered following the short and long delay, relative to the amount of data that was originally learned. It is obtained by summing AcSS scores for STPR and LTPR.

Each participant was assigned to one of the following four groups on the basis of scores above or below 1.5 SDs on the LEI and PRI: (1) high learning–high retention (HL–HR), (2) high learning–low retention (HL–LR), (3) low learning-high retention (LL–HR), or (4) low learning–low retention (LL–LR) group.

2.3. MR scanning and brain morphometry

Image acquisition and analysis methods were developed within the NIH/NCRR sponsored Morphometry Biomedical Informatics Research Network (mBIRN) and the ADNI (Fennema-Notestine et al., 2006; Han et al., 2006; Jack et al., 2008; Jovicich et al., 2006). Data were collected across a variety of 1.5 T scanners. Protocols are described in detail at http://www.loni.ucla.edu/ADNI/Research/Cores/index.shtml. Two T1-weighted volumes were acquired for each participant. These raw DICOM MRI scans were downloaded from the public ADNI site (http://www.loni.ucla.edu/ADNI/Data/index.shtml). Locally, images were reviewed for quality, automatically corrected for spatial distortion due to gradient nonlinearity (Jovicich et al., 2006) and B1 field inhomogeneity (Sled, Zijdenbos, & Evans, 1998), registered, and averaged to improve signal-to-noise. Volumetric segmentation (Fischl et al., 2002, 2004) and cortical surface reconstruction (Dale, Fischl, & Sereno, 1999; Dale & Sereno, 1993; Fischl, Sereno, & Dale, 1999; Fischl et al., 2004) methods based on FreeSurfer software, optimized for use on large, multi-site datasets, were used. To measure thickness, the cortical surface was reconstructed (Dale et al., 1999; Dale & Sereno, 1993) and parcellated into distinct regions of interest (ROIs) (Desikan et al., 2006; Fischl et al., 2004). Details of the application of these methods to the ADNI data have been described in full elsewhere (Fennema-Notestine et al., in press). To limit the number of multiple comparisons, only regions assumed to be involved in early AD pathology (Fennema-Notestine et al., 2006; Han et al., 2006; Jack et al., 2008; Jovicich et al., 2006) were included in the present analyses, including bilateral hippocampal formation (volumetric measures; not pictured) which included dentate gyrus, CA fields, subiculum/parasubiculum and the fimbria (Makris et al., 1999), frontal, other temporal, parietal lobe areas, and cingulate regions bilaterally (thickness measures) (see ROIs listed in Table 2). Defined frontal ROIs included the frontal pole, caudal and rostral portions of the middle frontal cortex, lateral and medial regions of the orbitofrontal cortex, superior frontal cortex, the par orbitalis, and the frontal operculum (comprised of the pars opercularis and pars triangularis). Temporal ROIs, in addition to the hippocampus, included entorhinal cortex, fusiform gyrus, the superior, middle and inferior temporal gyri, and the temporal pole. Finally, parietal ROIs included the supramarginal gyrus, superior and inferior parietal cortex, and the precuneus. To further decrease numbers of comparisons, the caudal and rostral anterior cingulate regions were combined as anterior cingulate cortex (ACC); the isthmus and posterior cingulate regions were combined as posterior cingulate cortex (PCC); and as mentioned the pars opercularis and pars triangularis were combined as the frontal operculum. Baseline volumetric data were corrected for individual differences in head size by regressing the estimated total intracranial volume (eTIV) as in Buckner et al. (2004).

2.4. Statistical analysis

Group comparisons were performed with analyses of variance (ANOVAs) or Chi-square tests for demographic variables. To assess group differences in morphometric variables at baseline, a repeated measure analysis of variance (ANOVA) was performed with learning–retention group as the between-subject factor, and hemisphere (left versus right) and ROIs (including all ROIs in a single model) as within-subjects variables. Prior to analyses, effects of age and gender were regressed from all thickness and volumetric measures, and standardized residual values were used for analyses; bilateral hippocampal volumes also were corrected for differences in head size by regressing the eTIV volume (Buckner et al., 2004). When significant group effects were observed for a given ROI, univariate analyses were performed and the α level was set to $p < 0.002$ (Bonferroni correction). Effect sizes were calculated for pairwise comparisons on morphometric variables using Cohen's d (Cohen, 1977), computed by dividing the mean difference between groups by the pooled standard deviation.

To examine the unique relationship between learning or retention and morphometry, partial correlations were performed after controlling for the effects of gender and education. We did not enter age as a control variable because both LEI and PRI were calculated via age-corrected norms and morphometric variables were already corrected for age effects prior to analysis. In addition, since retention scores are often highly correlated with learning scores (in the present study, $r = 0.64$, $p < 0.001$), we conducted a separate correlation analysis between retention scores and morphometric variables after controlling for the effects of learning, gender and education to obtain a better estimate of the relationship between retention and morphometry. Separate partial correlations, controlling for apolipoprotein E (APOE) genotype in addition to gender and education variables, were performed due to a larger proportion of $\varepsilon 4$ carriers in the LR-LL group than in the other three groups. The α level for the partial correlation analyses was set to 0.001 based on Bonferroni corrections.

To evaluate the predictive value of several covariates on clinical outcome, defined by AD conversion during the two-year follow-up, we conducted binary logistic regression. Specifically, the predictors included in the model were age, gender, education, MMSE scores at baseline, presence of at least one APOE $\varepsilon 4$ allele, and learning–retention group membership. Significant predictors were selected using the stepwise selection (LR) method with $\alpha \leq 0.05$. Odds ratios and 95% confidence intervals were calculated to quantify the effect of significant predictors. All analyses were conducted in SPSS (Version 17.0).

3. Results

3.1. Demographic and clinical characteristics

The demographic and clinical characteristics for the four subgroups are presented in Table 1. The four groups did not differ on level of education ($F_{(3,603)} = 2.60$, $p > 0.05$), although they showed a significant difference in age ($F_{(3,603)} = 9.49$, $p < 0.001$). Individuals in the LL–LR groups were significantly younger than the HL–LR and HL–HR groups (both p-values < 0.005). The groups also differed in gender distribution ($\chi^2_{(3,\,N=607)} = 16.70$, $p < 0.005$): the LL–HR group contained more men than the other three groups ($p < 0.05$), and the LL–LR group contained more men than the HL–HR group ($p < 0.005$). Groups significantly differed on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) scores ($F_{(3,603)} = 50.08$, $p < 0.001$) with the HL–HR group showing higher scores relative to the other three groups (all p-values < 0.001), and the HL–LR group showing significantly higher scores than the LL–LR and the LL–HR groups (both p-values < 0.005). Moreover, a significant difference in frequency of APOE $\varepsilon 4$ carriers among the four groups was found ($\chi^2_{(3,\,N=594)} = 52.83$, $p < 0.001$). The LL–LR group demonstrated the highest frequency and the HL–HR group showed the lowest frequency of APOE $\varepsilon 4$ carriers among groups (LL–LR > LL–HR = HL–LR > HL–HR, all p-values < 0.05). In summary, the LL–LR group tended to be younger and have a higher frequency of APOE $\varepsilon 4$ carriers than other groups, and the LL–LR and LL–HR groups showed lower MMSE scores than the other two groups.

3.2. Regional differences in morphometry by group

Results of the ANOVA for ROIs (including all ROIs in a single model) revealed a main effect of group ($F_{(3,547)} = 23.26$, $p < 0.001$) and a region by group interaction ($F_{(63,1587)} = 2.34$, $p < 0.001$). There

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

Table 1
Demographic and global cognitive characteristics of the four groups based on the learning–retention classification scheme.

	LL–LR, $n = 185$ (mean, S.D.)	LL–HR, $n = 53$ (mean, S.D.)	HL–LR, $n = 124$ (mean, S.D.)	HL–HR, $n = 245$ (mean, S.D.)
Age	73.75 (7.59)**	75.82 (7.08)	76.66 (5.52)	76.97 (5.95)
Education	15.44 (2.96)	15.25 (3.59)	15.89 (2.85)	16.13 (2.82)
Gender (% men)	67%‡	76%#	58%	51%
MMSE	26.85 (1.75)	26.68 (1.76)	27.90 (1.76)†	28.64 (1.47)††
% APOE ε4+	64%*	45%	39%	28%§
Baseline ADNI diagnosis				
Healthy controls	8	8	39	167
MCI	177	45	85	78

‡ The LL–LR group was significantly different from the HL–HR group ($p < 0.05$).
† The HL–LR group was significantly different from the LL–LR and the LL–HR groups ($p < 0.005$).
†† The HL–HR group was significantly different from the other three groups.
* The LL–LR group was significantly different from the other three groups ($p < 0.05$).
** The LL–LR group was significantly different from the HL–LR and HL–HR groups ($p < 0.005$).
# The LL–HR group was significantly different from the other three groups ($p < 0.05$).
§ The HL–HR group was significantly different from the LL–HR and the HL–LR groups ($p < 0.001$).

was no main effect of hemisphere ($F_{(1,547)} = 0.22$, $p = 0.88$), group by hemisphere interaction ($F_{(3,547)} = 0.33$, $p = 0.80$), or group by region by hemisphere interaction ($F_{(63,1587)} = .87$, $p = 0.76$). Therefore, follow-up analyses were collapsed across right and left hemisphere values by averaging the volumes or cortical thickness in both hemispheres. Univariate ANOVAs revealed group differences across all ROIs except for ACC ($F_{(3,547)} = 2.85$, $p = 0.04$). MR morphometric measures by group for selected ROIs are presented in Fig. 1. Continuous surface maps of cortical thickness between groups are shown in Fig. 2. Post hoc comparisons were performed within each ROI among groups and the results, ordered by effect size magnitude, were as follows:

LL–LR versus HL–HR: Relative to the HL–HR group, the LL–LR group showed greater volume reduction or cortical thinning in entorhinal cortex (Cohen's $d = 1.01$), hippocampus ($d = 1.00$), middle temporal ($d = 0.83$), inferior temporal ($d = 0.81$), fusiform cortex ($d = 0.78$), temporal pole ($d = 0.67$), inferior parietal ($d = 0.65$), pre-

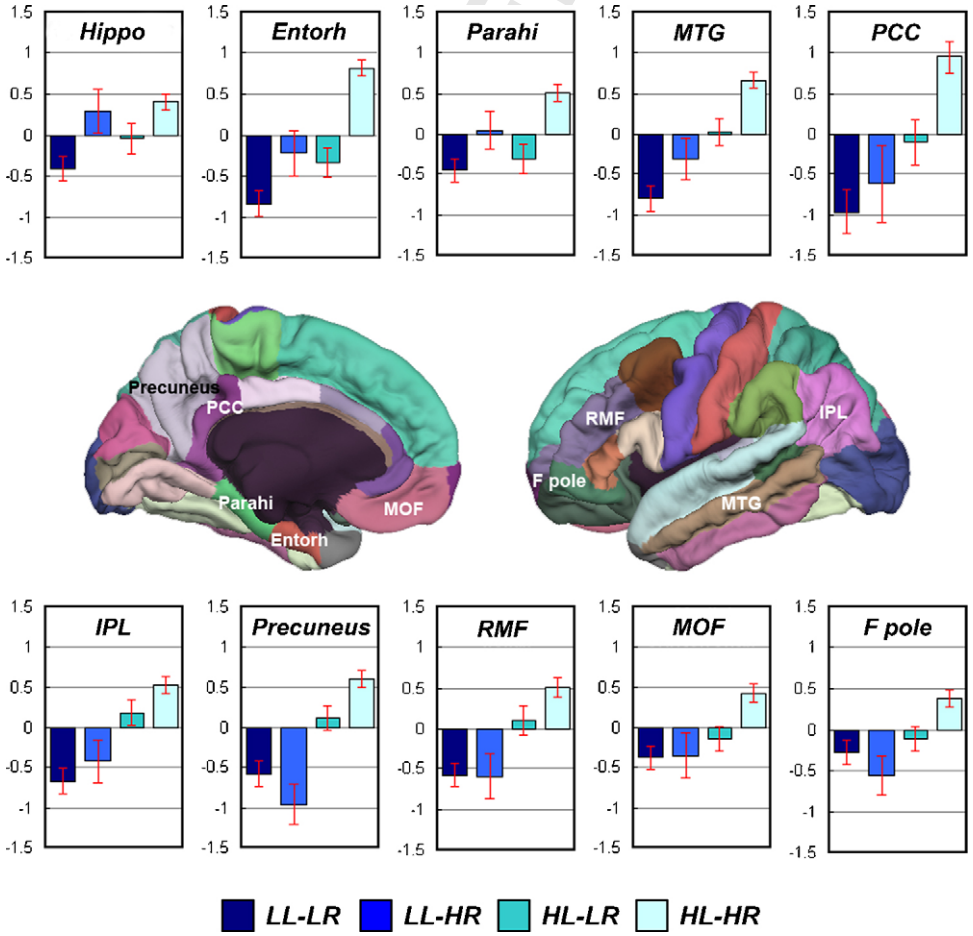
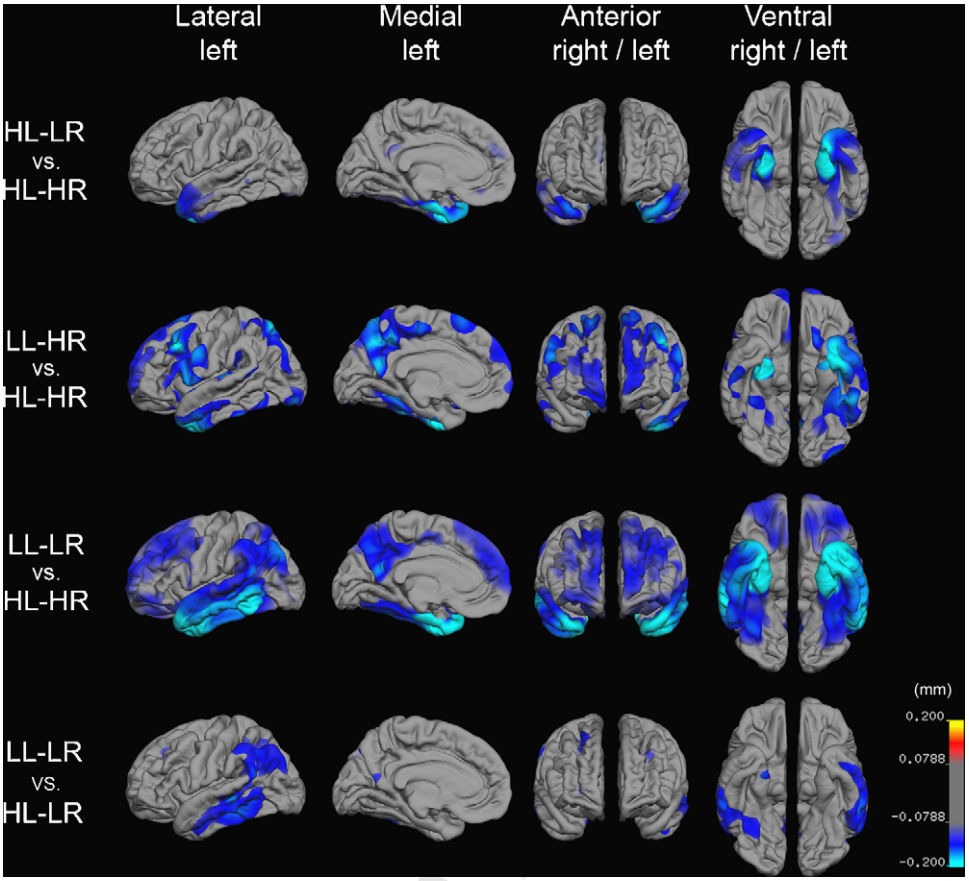


Fig. 1. Bar chart showing MR values for hippocampal volume and thickness for selected regions for the four groups. All values are standardized residuals (*z*-score) after the effects of age and gender have been regressed out. The hippocampal volumes are also controlled for the effect of eTIV. Error bars: standard error of the mean. Hippo: hippocampus; Entorh: entorhinal cortex; Parahi: parahippocampal gyrus; MTG: middle temporal gyrus; PCC: posterior cingulate cortices; IPL: inferior parietal lobule; RMF: rostral middle frontal; MOF: medial orbitofrontal gyrus; F pole: frontal pole.

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024



Q4 Fig. 2. Reconstructed cortical surface maps representing the average mean difference in thickness (mm, $p < 0.002$) for the three groups with learning and/or retention impairment relative to memory intact group (top three rows), and the LL–LR group relative to the HL–LR group (bottom row), after controlling for the effects of age and gender. Blue and cyan indicate thinning whereas red and yellow indicate thickening. Relative to the HL–HR group, the two groups with impaired learning ability showed a more widespread pattern of cortical thinning, involving temporal, frontal regions, and PCC. In contrast, the low retention group (the HL–LR group) demonstrated significantly thinner gray matter in medial temporal areas and PCC relative to the HL–HR group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

cuneus ($d = 0.63$), superior temporal ($d = 0.61$), rostral middle frontal ($d = 0.61$), PCC region ($d = 0.60$), superior frontal ($d = 0.55$), parahippocampus ($d = 0.54$), supramarginal gyrus ($d = 0.54$), caudal middle frontal ($d = 0.51$), pars orbitalis ($d = 0.50$), lateral orbitofrontal ($d = 0.47$), medial orbitofrontal ($d = 0.46$), superior parietal ($d = 0.45$), frontal operculum ($d = 0.42$), and frontal pole ($d = 0.38$).

LL–HR versus HL–HR: Relative to the HL–HR group, the LL–HR group showed greater volume reduction or cortical thinning in precuneus (Cohen's $d = 0.94$), caudal middle frontal ($d = 0.74$), hippocampus ($d = 0.70$), fusiform cortex ($d = 0.68$), superior parietal ($d = 0.67$), entorhinal cortex ($d = 0.64$), middle temporal ($d = 0.61$), inferior temporal ($d = 0.60$), rostral middle frontal ($d = 0.60$), frontal pole ($d = 0.58$), frontal operculum ($d = 0.57$), temporal pole ($d = 0.55$), superior frontal ($d = 0.55$), inferior parietal ($d = 0.55$), superior temporal ($d = 0.53$), PCC region ($d = 0.52$), supramarginal gyrus ($d = 0.51$), and medial orbitofrontal cortex ($d = 0.43$).

HL–LR versus HL–HR: Relative to the HL–HR group, the HL–LR group showed greater volume reduction or cortical thinning in hippocampus (Cohen's $d = 0.77$), entorhinal cortex ($d = 0.69$), fusiform cortex ($d = 0.54$), temporal pole ($d = 0.50$), parahippocampus ($d = 0.46$), middle temporal ($d = 0.39$), inferior temporal ($d = 0.38$), superior temporal ($d = 0.37$), PCC region ($d = 0.35$), and medial orbitofrontal cortex ($d = 0.34$).

LL–LR versus HL–LR: The LL–LR group showed greater cortical thinning within inferior parietal (Cohen's $d = 0.45$), middle temporal ($d = 0.43$), supramarginal ($d = 0.38$), precuneus ($d = 0.37$), and rostral middle frontal ($d = 0.36$) regions compared to the HL–LR group.

Overall, the two impaired learning groups (the LL–LR and the LL–HR groups) relative to their counterparts (the HL–LR and the HL–HR groups, respectively) showed greater volume reduction or cortical thinning in areas beyond temporal lobe including prefrontal and parietal regions.

3.3. Relationship between learning and memory measures and morphometry

To examine the relationship between learning and memory and morphometric measures, we first performed partial correlations controlling for the effects of gender and education on all four groups. Results are presented in Table 2 (left columns). Learning and retention scores were both significantly correlated with all lateral frontal, medial frontal, lateral temporal, medial temporal, anterior temporal, parietal, ACC, and PCC ROIs included in the present study. We further examined the relationship between retention ability and morphometric measures by adding the learning scores as a controlled variable in the partial correlation analysis (PRI 2, Table 2). The results showed that the PRI 2 retention scores were significantly correlated with anterior, medial, and ventral temporal lobe which included hippocampal volumes, and cortical thickness of entorhinal, parahippocampal, temporal pole, and fusiform regions.

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

Table 2
Partial correlation coefficients between learning and retention function and volumetric measure of hippocampus and cortical thickness measures of frontal, parietal, other temporal lobe regions as well as cingulate cortices with all participants. The left column indicates results controlling for the effects of gender and education; the right column indicates results controlling for the effects of gender, education, and APOE genotype (the age effect has been controlled for both cognitive and morphometric variables before entering the analyses).

	LEI (learning)		PRI 1 (retention)		PRI 2 (retention controlling for learning)	
Cingulate cortex						
Anterior	0.13*	0.11	0.11	0.09	0.03	0.02
Posterior	0.24*	0.22*	0.21*	0.17*	0.06	0.05
Frontal						
F pole	0.17*	0.14*	0.15*	0.12*	0.05	0.04
Caudal middle F	0.20*	0.16*	0.14*	0.11	0.02	0.01
Rostral middle F	0.27*	0.23*	0.20*	0.16*	0.04	0.03
Lateral orbitoF	0.19*	0.16*	0.16*	0.13	0.06	0.05
Medial orbitoF	0.20*	0.17*	0.19*	0.17*	0.09	0.08
Superior F	0.23*	0.19*	0.19*	0.15*	0.05	0.04
Pars orbitalis	0.23*	0.20*	0.20*	0.18*	0.07	0.07
Operculum	0.21*	0.18*	0.17*	0.14*	0.05	0.04
Temporal						
Hippocampus	0.39*	0.36*	0.41*	0.37*	0.22*	0.21*
Parahippocampus	0.24*	0.22*	0.25*	0.23*	0.13*	0.12
Entorhinal	0.35*	0.33*	0.35*	0.33*	0.19*	0.17*
Fusiform	0.29*	0.27*	0.27*	0.25*	0.11*	0.11
T pole	0.27*	0.26*	0.25*	0.25*	0.12*	0.12
Superior T	0.26*	0.23*	0.23*	0.20*	0.07	0.07
Middle T	0.32*	0.29*	0.27*	0.24*	0.09	0.08
Inferior T	0.32*	0.30*	0.26*	0.24*	0.08	0.07
Parietal						
Supramarginal	0.23*	0.21*	0.17*	0.15*	0.03	0.03
Superior P	0.20*	0.17*	0.13*	0.12	0.01	0.01
Inferior P	0.27*	0.24*	0.19*	0.17*	0.03	0.02
Precuneus	0.28*	0.25*	0.18*	0.16*	0.01	0.01

PRI 1 indicates results without controlling for LEI; PRI 2 indicates results with controlling LEI in addition to other covariates. F: frontal; T: temporal.
* $p < 0.001$.

Because the LR–LL group had more ε4 carriers than the other three groups, a second set of partial correlation analyses were performed after controlling for the effects of APOE genotype, gender and education. A similar pattern of findings was observed, although in some cases the correlation coefficients were modestly reduced when APOE genotype was included. Specifically, the learning scores remained significantly correlated with widespread brain regions across frontal, temporal, and parietal areas, although the ACC was no longer significantly correlated with learning scores once APOE genotype was included. In addition, retention scores remained significantly correlated with morphometric measures in medial temporal lobe regions (i.e., hippocampus and entorhinal regions), although relationships with other temporal lobe ROIs, including parahippocampus, fusiform, and temporal pole, were no longer statistically significant after controlling for APOE genotype (Table 2, right columns).

3.4. Conversion rates to probable AD and prediction of AD conversion over two-year follow-up

Over the two-year follow-up period, the four groups significantly differed in the AD conversion rate ($\chi^2_{(3, N=423)} = 117.94$, $p < 0.001$). The LL–LR group (61.1%) showed a significantly higher AD conversion rate compared to the HL–LR (30.4%, $p < 0.001$) and the HL–HR (4.0%, $p < 0.001$) groups but a more comparable rate to the LL–HR group (43.3%, $p = 0.06$). There was no significant difference in the AD conversion rate between the LL–HR and the HL–LR group ($p = 0.27$), but both groups showed a significantly higher AD conversion rate relative to the HL–HR group (both p-values < 0.001) (see left side of Fig. 3).

We then combined individuals with impaired learning ability regardless of the level of their retention ability (i.e., the LL–HR and the LL–LR groups) into a "combined low learning group",

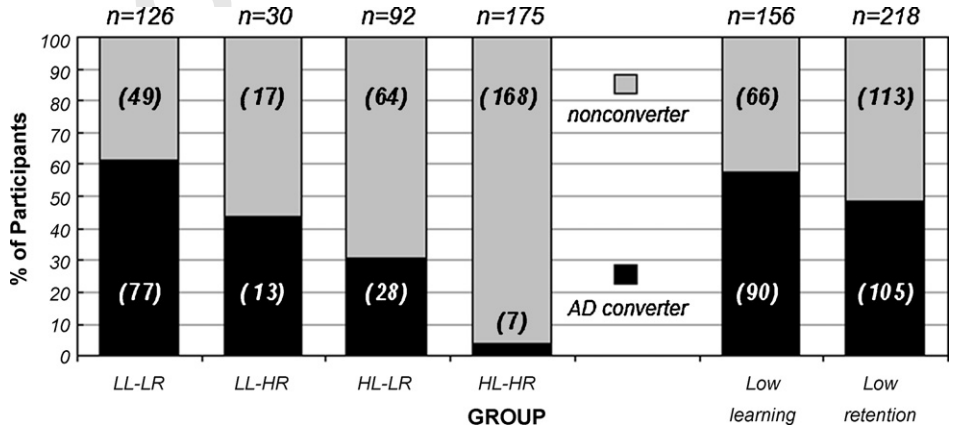


Fig. 3. Bar chart of AD conversion rate over two years for the four groups (left) as well as for the combined low learning/retention groups (right).

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

Table 3
Logistic regression with age, gender, education, MMSE scores, presence of at least one APOE ε4 allele, and the learning–retention group membership as predictors, and AD conversion over two years as the outcome measure.

	Odds ratio	95% CI	Wald χ^2	*p*-value
MMSE scores	0.69	0.59–0.81	21.26	<0.001
APOE ε4+ allele	2.04	1.19–3.50	6.64	<0.05
Learning–retention group*				
LL–LR	17.84	7.38–43.10	40.99	<0.001
LL–HR	9.01	2.98–27.21	15.18	<0.001
HL–LR	8.48	3.45–20.86	21.65	<0.001

* The HL–HR as the reference group.

and similarly combined individuals with impaired retention ability regardless of their learning ability (i.e., the HL–LR and the LL–LR groups) into a "combined low retention group." A Chi-square analysis (see right side of Fig. 3) revealed that the combined low learning group had significantly more participants who developed AD over a two-year follow-up period (57.7%) than the combined low retention group (48.2%) ($p < 0.001$).

We next sought to determine whether the classification of memory deficits would be useful in the assessment of the likelihood of an AD diagnosis at follow-up. Table 3 shows the results of logistic regression models predicting these outcomes. The multivariate logistic regression results showed that age, gender, and level of education were not significant predictors of progression to AD (*p*-values > 0.05), whereas MMSE, APOE, and group membership in the learning–retention scheme were significantly associated with the likelihood of AD conversion over two-year follow-up. Specifically, the likelihood of progression to AD was predicted by presence of at least one APOE ε4 allele, low MMSE scores, and by membership in the LL–LR, LL–HR, or the HL–LR groups. Individuals with the presence of at least one APOE ε4 allele showed two-fold increase in likelihood of AD progression relative to those did not have an APOE ε4 allele; individuals in the LL–LR group demonstrated approximately an 18-fold increase in risk of converting to AD compared to those with intact learning and retention abilities. Individuals in the LL–HR or the HL–LR groups showed an eight- to nine-fold increase in likelihood of AD progression over a two-year period.

As post hoc analyses, we conducted separate ANOVAs or Chi-square tests to examine the demographic and global cognitive variables between individuals with versus without follow-up data for each of the four groups to look for potential selective attrition effects. The result showed that across the four groups, individuals with or without follow-up data did not differ in age, education level, gender distribution, APOE ε4 status, or MMSE scores (all *p*-values > 0.05), suggesting that it is unlikely that selective attrition occurred (Table 4).

4. Discussion

To date, studies of MCI have relied almost exclusively on delayed recall or retention measures in rendering the diagnosis (Arnaiz & Almkvist, 2003), and there has been a relative dearth of research parsing the underlying components of the memory problems that characterize individuals with MCI. Thus, we investigated qualitative differences in learning versus retention, and their relation to morphometric measures and disease progression, in MCI. We presented behavioral evidence showing that MCI individuals can be characterized by impairments either in learning, retention, or both, based on a commonly used verbal memory task—RAVLT. Furthermore, individuals with different deficit profiles in learning and retention presented distinct patterns of brain morphometry. We then examined the AD progression rate among the MCI groups over a two-year follow-up and found that either impaired learning or impaired retention increased risk of future development of AD, but that individuals with both impaired learning and retention abilities showed the highest risk of AD conversion.

With respect to brain morphometry, we predicted that individuals with impaired learning ability would show a more widespread pattern of brain atrophy involving frontal, temporal, and parietal lobe regions, while individuals with impaired retention ability would demonstrate more circumscribed atrophy primarily involving medial temporal regions. Results based on group comparisons and partial correlation analyses supported our predictions and were consistent with previous neuroimaging and lesion studies (Moscovitch et al., 2005; Parsons et al., 2006; Powell et al., 2005; Rosen et al., 2005a; Shankle et al., 2005; Squire et al., 2004; Weintrob et al., 2007), suggesting that learning measures involve a broader neural network whereas retention measures show a more focal gray matter involvement.

Interestingly, relative to the healthy control (or HL–HR) group, the HL–LR group showed cortical thinning in medial orbitofrontal areas in addition to temporal lobe areas. The work of Stuss and coworkers has shown that medial orbitofrontal areas are associated with the ability to inhibit irrelevant information (Happaney, Zelazo, & Stuss, 2004; Stuss & Alexander, 2007). It is possible that good retention ability requires not only the integrity of medial temporal structures but also the ability to inhibit irrelevant information (e.g., words from the interference trial in the RAVLT) mediated by medial orbitofrontal regions (Stuss & Alexander, 2007).

Cortical thinning in PCC was found in all MCI groups relative to the HL–HR group. This was not unexpected given that the PCC is considered part of the limbic system and has reciprocal connections with the medial temporal lobe, including entorhinal cortex and hippocampal formation (Kobayashi & Amaral, 2003, 2007). Hypometabolism and volumetric reduction in PCC has been identified as a feature of early AD (Choo et al., in press; Chua, Wen, Slavin, & Sachdev, 2008; Pengas, Hodges, Watson, & Nestor, in press),

Table 4
Demographic and global cognitive characteristics of individuals with or without clinical follow-up outcome data (i.e., AD conversion) for the four groups.

Follow-up data	LL–LR		LL–HR		HL–LR		HL–HR	
	Yes, *n* = 126	No, *n* = 59	Yes, *n* = 30	No, *n* = 23	Yes, *n* = 92	No, *n* = 32	Yes, *n* = 175	No, *n* = 70
Age	73.88 (7.08)	73.47 (8.63)	76.13 (7.31)	75.42 (6.91)	76.50 (5.68)	77.14 (5.08)	77.01 (5.50)	76.88 (6.99)
Education	15.69 (2.87)	14.92 (3.11)	15.87 (3.27)	14.43 (3.89)	15.83 (2.80)	16.06 (3.05)	16.13 (2.74)	16.13 (3.02)
Gender (% men)	67%	66%	83%	65%	59%	56%	53%	47%
MMSE	26.83 (1.73)	26.92 (1.80)	26.73 (1.80)	26.61 (1.75)	28.02 (1.72)	27.53 (1.85)	28.31 (1.56)	28.78 (1.41)
% APOE ε4+	69%	53%	47%	43%	37%	47%	30%	24%

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

and several recent studies have reported PCC hypometabolism or/and volume reduction in individuals with MCI (Choo et al., in press; Chua et al., 2008; Fennema-Notestine et al., in press; Pengas, Hodges, Watson, & Nestor, in press). Overall, our findings were consistent with prior studies that suggest that PCC abnormality can be detected in a prodromal stage of AD. In addition, we found significant cortical thinning in the lateral temporal lobe regions in all MCI groups relative to the HL–HR group. Lateral temporal areas, particularly middle and inferior temporal gyri, have been implicated in the progression of AD (McEvoy et al., 2009; Whitwell et al., 2007; Whitwell et al., 2008). Although atrophy of the superior temporal gyrus has typically been observed only after a diagnosis of probable AD (Scahill, Schott, Stevens, Rossor, & Fox, 2002; Whitwell et al., 2007), our results, in accord with some recent studies (Chang et al., in press; Fan, Batmanghelich, Clark, & Davatzikos, 2008; McEvoy et al., 2009), showed significant atrophy in this area in the MCI groups, suggesting that atrophy of the lateral temporal gyrus can occur prior to a diagnosis of probable AD and may be associated with a higher risk of imminent clinical decline.

This possibility is further supported by studies demonstrating that measures of semantic knowledge show significant declines during prodromal AD (Cuetos, Arango-Lasprilla, Uribe, Valencia, & Lopera, 2007; Mickes et al., 2007; Powell et al., 2006), and that these cognitive operations may be relatively independent of episodic memory deficits (Koenig, Smith, Moore, Glosser, & Grossman, 2007). For example, Mickes et al. (2007) have shown in a detailed neuropsychological study of prodromal AD that both semantic memory and episodic memory functions declined rapidly in a three-year period progressing to AD, whereas executive function deficits were not particularly prominent. Mickes et al. concluded that cognitive abilities thought to be subserved by the medial and lateral temporal lobes (episodic and semantic memory, respectively) may be more prominently impaired than cognitive functions subserved by the frontal lobes (executive functions). These findings map nicely onto the known neuropathologic encroachments of AD early on in the disease process (Braak & Braak, 1991) and are also consistent with recent reports of decreased semantic access in non-demented APOE ε4 older adults (Rosen et al., 2005b) and the ability of language tasks to predict pathologic AD six years later (Powell et al., 2006).

Although distinguishable morphometric patterns were found between the poor learning or retention groups and the HL–HR group, the correlation coefficients observed between learning, retention and morphometric measures were generally low (r's = 0.11–0.41), suggesting that much of the variance in learning and retention scores is not explained by brain morphometry. It is likely that there are some factors of interindividual differences that may have also contributed to the differential morphometric patterns observed among groups. For example, the APOE ε4 allele has been documented as a genetic risk factor for late-onset AD (Bennett et al., 2003; Bondi et al., 1994; Bondi et al., 1999; Modrego, 2006). Some studies suggest that the APOE ε4 genotype, particularly for individuals who progress to AD over time, is associated with more widespread brain atrophy involving areas of medial temporal, frontal, and parietal regions (Hamalainen et al., 2008). Consistent with this view, we found that the LL–LR group had the highest frequency of APOE ε4 carriers among the three MCI groups and showed the most widespread pattern of gray matter atrophy relative to the other groups.

Another goal of the current study was to determine the relative utility of learning and retention measures in predicting AD progression among the four groups. Not surprisingly, individuals with both learning and retention impairments at baseline had the highest risk for progression to AD over two years. Learning impairment with intact retention, and retention impairment with intact learning were also each associated with an increased risk for developing AD, although our ability to directly compare the conversion rates of these two important subgroups (i.e., LL–HR versus HL–LR) was likely underpowered due to their relatively small sample sizes. However, individuals with learning deficits (regardless of the level of their retention abilities) at baseline showed a significantly higher likelihood of developing AD over two years compared to those with a retention deficit (regardless of the level of their learning abilities). These results are consistent with prior studies that have also reported differential sensitivity of learning and retention measures. For example, Grober and Kawas (1997), perhaps the first to show the utility of learning measures in prodromal AD, found that individuals in the prodromal stage of dementia recalled significantly fewer words during the learning trials of the free and cued selective reminding procedure than did matched control participants, whereas their retention of material over the 30-min delay period was identical to that of control participants, suggesting that learning variables may be a more sensitive measure for predicting AD conversion than retention. Also, Rabin et al. (2009) investigated the discriminative ability of several widely used clinical memory tests to classify individuals as MCI or healthy older adults. They found that the total learning score on a list-learning task appeared to be the most sensitive diagnostic index for distinguishing MCI from healthy aging. Together with the current results showing that learning impairment is associated with a higher rate of progression to AD than retention deficits in the absence of learning impairments, these findings suggest that learning measures can be as useful as retention measures in predicting progression from MCI to AD, and suggest that the use of only delayed recall or retention measures in studies of amnestic MCI potentially misses an important subset of older adults at risk of developing AD.

Our cross-sectional results showed that individuals with impaired learning or retention could be distinguished from elderly individuals without memory impairment not only from this neuropsychological perspective but also in terms of brain morphometry. Buckner (2004) suggests that AD pathology, even in the early stages, involves both hippocampal and frontal regions, though via different mechanisms. Some studies have also demonstrated that MCI individuals with more widespread gray matter loss at baseline progress more rapidly to AD relative to those with focal gray matter loss (McEvoy et al., 2009; Whitwell et al., 2008). Consistent with these studies, relative to the impaired retention group, the impaired learning group showed a more widespread pattern of gray matter loss at baseline involving frontal, temporal, and other cortical regions; and these individuals showed a higher progression rate to AD during the follow-up period. Overall, our finding suggests that learning ability, given its involvement in multiple cortical regions and likely reliance on other neuropsychological mechanisms such as attention and concentration, can be a sensitive indicator of imminent clinical decline in the prodromal period.

Some of the initial factor analytic studies of the neuropsychological measures comprising the MOANS core battery, which includes the RAVLT, also support this notion. Specifically, Smith, Ivnik, Malec, and Kokmen (1992), Smith, Ivnik, Malec, and Tangalos (1993) demonstrated that the RAVLT learning index loaded on a Learning factor along with a number of other learning and working memory measures (WMS-R Logical Memory I, Visual Reproduction I, Visual Associates, Paired Associates), whereas the RAVLT retention index loaded on a more circumscribed Retention factor with other measures of retention only (WMS-R Logical Memory Percent Retention, Visual Reproduction Percent Retention). Other factors on which neither of the RAVLT measures loaded were those relating to Verbal Comprehension, Perceptual Organization, or Attention (WAIS-R Digit Span and Arithmetic, WMS-R Mental Control and Visual Span). The robust psychometric characteristics of the RAVLT variables and their demonstrated stability in factor analytic studies of both normal and clinical dementia samples support the gener-

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

alizability of our findings with the RAVLT variables to other similar learning and retention measurement strategies.

Although the search for signature cognitive changes in prodromal AD has largely focused on episodic memory, as was the case in our study, several recent reviews and meta-analyses also suggest that there is decline in other cognitive domains in addition to episodic memory in the few years prior to a dementia diagnosis, including deficits in semantic memory, visuospatial skills, executive functions, and attention and speed of processing (Backman et al., 2004). This widespread decline in cognitive abilities mirrors evidence that multiple brain regions (e.g., medial and lateral temporal lobes, frontal and parietal cortices, cingulate cortex) or connectivity between these regions are impaired in prodromal AD (Small, Mobly, Laukka, Jones, & Backman, 2003). Future studies that more broadly sample cognitive domains beyond episodic memory will be better able to delineate these brain–behavior relationships in MCI and prodromal AD.

Broader conceptualizations of MCI have emerged in recent years to encompass cognitive domains other than episodic memory (Petersen & Morris, 2005), and clinical subtypes that include amnestic and non-amnestic forms, or single or multiple cognitive domains, have been offered. With the advent of these broader classification schemes, diagnostic challenges related to MCI have understandably increased, and neuropsychological assessment of multiple cognitive domains—with sensitive and specific measures of the AD prodrome—will increasingly play a prominent role in resolving these challenges. For example, a pair of recent studies have shown that, when compared to the typical approach to diagnosing MCI (e.g., recall deficit $\geq$1.5 standard deviations; CDR score of 0.5; normal MMSE score), a comprehensive neuropsychological approach to MCI diagnosis results in more robust associations with expected anatomical and stroke risk findings (Jak, Bondi, et al., 2009) as well as better prediction of progression to dementia (Saxton et al., 2009). Use of comprehensive neuropsychological assessment when diagnosing MCI subtypes will help to improve the stability and reliability of diagnosis, as will the use of multiple measurements (e.g., learning and retention measures) within a cognitive domain such as episodic memory (see Jak, Urban, et al., 2009). Our finding that combined learning and retention impairment was superior to isolated learning or retention impairment in predicting progression to AD supports this notion.

Related to this, differences in the classification of some individuals as MCI or normally aging in the current study relative to classification of these individuals within the ADNI reflects the problems associated with the use of different operational criteria across studies (Busse, Hensel, Guhne, Angermeyer, & Riedel-Heller, 2006; Jak, Urban, et al., 2009). The internal consistency of MCI diagnosis or prediction of AD progression based on alterations of the classification criteria (i.e., the cut point at which performance was considered impaired) was not the primary interest of the current study. However, differences in the conversion rates among the three MCI groups identified here, and between these MCI groups and the HC group, suggests that diagnostic schemes that incorporate more than delayed recall and global screening measures will increase sensitivity and reliability in predicting diagnostic outcome and likelihood of conversion to AD. Anchoring such sophisticated diagnostic schemes to underlying brain morphometric changes and prediction of AD progression will also provide much needed improvements in MCI diagnostic procedures.

Despite the potential clinical value of our findings, there are limitations that should be noted. First, with the limited number of learning and memory tests available in the ADNI, it is not possible to compare the relative diagnostic and predictive value of visual versus verbal learning and memory tests. Second, with the large sample sizes afforded by the ADNI, it is possible to observe statistically significant group differences, as we did for the bulk of our comparisons, although the clinical impact of these statistically significant findings may not be as clear cut. Fortunately, for at least a subset of the analyses we were able to provide effect size statistics, the bulk of which showed medium to large effect sizes, bolstering the potential clinical import of the findings. Third, two-year clinical follow-up data was available for only 69% of MCI individuals at the time we conducted this study. This is not uncommon in prospective studies of older adults (Visser, Pluijm, Stel, Bosscher, & Deeg, 2002), particularly with such a large-scale project. Additional follow-up data over a longer time interval could provide clarifying information on the relative progression rates between the LL–HR and HL–LR groups. Nevertheless, despite some dropout, MCI participants with or without follow-up data within each group did not significantly differ in any of the baseline demographic (i.e., age, education level, gender distribution, APOE $\varepsilon 4$ status) or global cognitive (i.e., MMSE score) characteristics. Thus, it seems unlikely that selective attrition occurred.

In conclusion, we provide evidence in support of the use of both learning and retention measures in the diagnosis of MCI. Furthermore, understanding the underlying qualitative feature of memory deficits in MCI can provide critical information for early detection of AD. Individuals with learning or retention impairment appear to be distinguishable not only neuropsychologically but also morphometrically. That is, individuals with learning deficits appear to show a more widespread pattern of gray matter loss, whereas individuals with retention deficits tend to show more focal gray matter loss with largest effects in medial temporal regions and PCC at baseline. Moreover, both learning and retention measures provide good predictive value for longitudinal clinical outcome, although impaired learning had modestly better predictive power than impaired retention. As expected, use of both measures provided the best predictive power. Hence, the conventional practice relying on the use of delayed recall or retention measures only in most MCI diagnostic schemes misses an important subset of individuals with prodromal AD. Overall, our results highlight the importance of including learning measures in addition to retention measures when making a diagnosis of MCI and for predicting clinical outcome. Knowledge of affected memory processes can also help to tailor specific auxiliary mnemonic strategies in cognitive training in MCI populations.

Conflict of interest

Anders M. Dale is a founder and holds equity in CorTechs Labs, Inc., and also serves on the Scientific Advisory Board. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies. The other authors do not have a financial or any other conflict of interest to disclose related to this manuscript.

Uncited references

Andrews-Hanna et al. (2007), Apostolova and Thompson (2008), and Storandt et al. (2006).

Acknowledgments

The authors would like to thank Alain Koyama, Robin G. Jennings, Michele Perry, Chris Pung, and Elaine Wu for downloading and preprocessing the ADNI MRI data.

Funding: This work was supported by the National Institutes on Aging [R01 AG012674, AG031224, K01AG029218, and K24 AG026431]; the National Center for Research Resources [U24 RR021382]; Alzheimer's Association IIRG-07-59343; the Department of Veterans Affairs Advanced Career Development Award

[0007 08-0957]; and the University of California, San Diego (UCSD) Stein Institute for Research on Aging. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and through generous contributions. The Foundation for the National Institutes of Health (www.fnih.org) coordinates the private sector participation of the $60 million ADNI public–private partnership that was begun by the National Institute on Aging (NIA) and supported by the National Institutes of Health. To date, more than $27 million has been provided to the Foundation for National Institutes of Health by Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson & Johnson, Eli Lilly and Co., Merck & Co., Inc., Novartis AG, Pfizer Inc., F. Hoffmann-La Roche, Schering-Plough, Synarc Inc., and Wyeth, as well the Alzheimer's Association and the Institute for the Study of Aging, Food and Drug Administration. Industry partnerships are coordinated through the Foundation for the National Institutes of Health. The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the UCSD. ADNI data are disseminated by the Laboratory of NeuroImaging at the University of California, Los Angeles.

References

Albert, M. S., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of International Neuropsychological Society*, *7*(5), 631–639.

Andrews-Hanna, J. R., Snyder, A. Z., Vincent, J. L., Lustig, C., Head, D., Raichle, M. E., et al. (2007). Disruption of large-scale brain systems in advanced aging. *Neuron*, *56*(5), 924–935.

Apostolova, L. G., & Thompson, P. M. (2008). Mapping progressive brain structural changes in early Alzheimer's disease and mild cognitive impairment. *Neuropsychologia*, *46*(6), 1597–1612.

Arnaiz, E., & Almkvist, O. (2003). Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. *Acta Neurologica Scandinavica Supplementum*, *179*, 34–41.

Axmacher, N., Schmitz, D. P., Weinreich, I., Elger, C. E., & Fell, J. (2008). Interaction of working memory and long-term memory in the medial temporal lobe. *Cerebral Cortex*, *18*(12), 2868–2878.

Backman, L., Jones, S., Berger, A. K., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology*, *19*(4), 520–531.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Bondi, M. W., Monsch, A. U., Galasko, D., Butters, N., Salmon, D., & Delis, D. C. (1994). Preclinical cognitive markers of dementia of the Alzheimer type. *Neuropsychology*, *8*(3), 374–384.

Bondi, M. W., Salmon, D. P., Galasko, D., Thomas, R. G., & Thal, L. J. (1999). Neuropsychological function and apolipoprotein E genotype in the preclinical detection of Alzheimer's disease. *Psychology and Aging*, *14*(2), 295–303.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239–259.

Buckner, R. L. (2004). Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*(1), 195–208.

Buckner, R. L., Head, D., Parker, J., Fotenos, A. F., Marcus, D., Morris, J. C., et al. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage*, *23*(2), 724–738.

Busse, A., Hensel, A., Guhne, U., Angermeyer, M. C., & Riedel-Heller, S. G. (2006). Mild cognitive impairment: long-term course of four clinical subtypes. *Neurology*, *67*(12), 2176–2185.

Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: an empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47.

Q3 Chang, Y. L., Jacobson, M. W., Fennema-Notestine, C., Hagler Jr., D. J., Jennings, R. G., Dale, A. M., et al. (in press). Level of executive function influences verbal memory in amnestic mild cognitive impairment and predicts prefrontal and posterior cingulate thickness. *Cerebral Cortex*.

Choo, I. H., Lee, D. Y., Oh, J. S., Lee, J. S., Lee, D. S., Song, I. C., et al. (in press) Posterior cingulate cortex atrophy and regional cingulum disruption in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging*.

Chua, T. C., Wen, W., Slavin, M. J., & Sachdev, P. S. (2008). Diffusion tensor imaging in mild cognitive impairment and Alzheimer's disease: A review. *Current Opinion Neurology*, *21*(1), 83–92.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cuetos, F., Arango-Lasprilla, J. C., Uribe, C., Valencia, C., & Lopera, F. (2007). Linguistic changes in verbal expression: A preclinical marker of Alzheimer's disease. *Journal of International Neuropsychological Society*, *13*(3), 433–439.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, *9*(2), 179–194.

Dale, A. M., & Sereno, M. I. (1993). Improved localization of cortical activity by combining Eeg and Meg with Mri cortical surface reconstruction—A linear-approach. *Journal of Cognitive Neuroscience*, *5*(2), 162–176.

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968–980.

Fan, Y., Batmanghelich, N., Clark, C. M., & Davatzikos, C. (2008). Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *Neuroimage*, *39*(4), 1731–1743.

Fennema-Notestine, C., Hagler, D. J., Jr., McEvoy, L. K., Fleisher, A. S., Wu, E. H., Karow, D. S., et al. (in press). Structural MRI biomarkers for preclinical and mild Alzheimer's disease. *Human Brain Mapping*.

Fennema-Notestine, C., Ozyurt, I. B., Clark, C. P., Morris, S., Bischoff-Grethe, A., Bondi, M. W., et al. (2006). Quantitative evaluation of automated skull-stripping methods applied to contemporary and legacy images: Effects of diagnosis, bias correction, and slice location. *Human Brain Mapping*, *27*(2), 99–113.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341–355.

Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *Neuroimage*, *9*(2), 195–207.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D. H., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Fujii, T., Okuda, J., Tsukiura, T., Ohtake, H., Suzuki, M., Kawashima, R., et al. (2002). Encoding-related brain activity during deep processing of verbal materials: A PET study. *Neuroscience Research*, *44*(4), 429–438.

Greene, J. D., Baddeley, A. D., & Hodges, J. R. (1996). Analysis of the episodic memory deficit in early Alzheimer's disease: Evidence from the doors and people test. *Neuropsychologia*, *34*(6), 537–551.

Grober, E., Hall, C. B., Lipton, R. B., Zonderman, A. B., Resnick, S. M., & Kawas, C. (2008). Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *Journal of International Neuropsychological Soc*, *14*(2), 266–278.

Grober, E, & Kawas, C. (1997). Learning and retention in preclinical and early Alzheimer's disease. *Psychology and Aging*, *12*(1), 183–188.

Hamalainen, A., Grau-Olivares, M., Tervo, S., Niskanen, E., Pennanen, C., Huuskonen, J., et al. (2008). Apolipoprotein E epsilon 4 allele is associated with increased atrophy in progressive mild cognitive impairment: A voxel-based morphometric study. *Neurodegenerative Diseases*, *5*(3–4), 186–189.

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., et al. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *Neuroimage*, *32*(1), 180–194.

Hannula, D. E., & Ranganath, C. (2008). Medial temporal lobe activity predicts successful relational memory binding. *Journal of Neuroscience*, *28*(1), 116–124.

Happaney, K., Zelazo, P. D., & Stuss, D. T. (2004). Development of orbitofrontal function: Current themes and future directions. *Brain and Cognition*, *55*(1), 1–10.

Harris, M. E., Ivnik, R. J., & Smith, G. E. (2002). Mayo's Older Americans Normative Studies: Expanded AVLT recognition trial norms for ages 57 to 98. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 214–220.

Hart, R. P., Kwentus, J. A., Harkins, S. W., & Taylor, J. R. (1988). Rate of forgetting in mild Alzheimer's-type dementia. *Brain and Cognition*, *7*(1), 31–38.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry*, *140*, 566–572.

Ivnik, R., Malec, J., Smith, G., Tangalos, E., Petersen, R. C., Kokmen, E., et al. (1992). Mayo's Older Americans Normative Studies: Updated AVLT norms for ages 56 to 97. *The Clinical Neuropsychologist*, *6*(Suppl.), 83–107.

Jack, C. R., Jr., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., et al. (2008). The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging*, *27*(4), 685–691.

Jak, A. J., Bondi, M. W., Delano-Wood, L., Wierenga, C., Corey-Bloom, J., Salmon, D. P., et al. (2009). Quantification of five neuropsychological approaches to defining mild cognitive impairment. *American Journal of Geriatric Psychiatry*, *17*(5), 368–375.

Jak, A. J., Urban, S., McCauley, A., Bangen, K. J., Delano-Wood, L., Corey-Bloom, J., et al. (2009). Profile of hippocampal volumes and stroke risk varies by neuropsychological definition of mild cognitive impairment. *Journal of International Neuropsychological Society*, 1–8.

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024

Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., et al. (2006). Reliability in multi-site structural MRI studies: Effects of gradient non-linearity correction on phantom and human data. *Neuroimage*, *30*(2), 436–443.
Kobayashi, Y., & Amaral, D. G. (2003). Macaque monkey retrosplenial cortex: II. Cortical afferents. *Journal of Comparative Neurology*, *466*(1), 48–79.
Kobayashi, Y., & Amaral, D. G. (2007). Macaque monkey retrosplenial cortex: III. Cortical afferents. *Journal of Comparative Neurology*, *502*(5), 810–833.
Koenig, P., Smith, E. E., Moore, P., Glosser, G., & Grossman, M. (2007). Categorization of novel animals by patients with Alzheimer's disease and corticobasal degeneration. *Neuropsychology*, *21*(2), 193–206.
Leube, D. T., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. (2001). Differential activation in parahippocampal and prefrontal cortex during word and face encoding tasks. *Neuroreport*, *12*(12), 2773–2777.
Leube, D. T., Weis, S., Freymann, K., Erb, M., Jessen, F., Heun, R., et al. (2008). Neural correlates of verbal episodic memory in patients with MCI and Alzheimer's disease—A VBM study. *International Journal of Geriatric Psychiatry*, *23*(11), 1114–1118.
Lucas, J. A. (2005). Disorders of memory. *Psychiatric Clinics of North America*, *28*(3), 581–597, 594
Makris, N., Meyer, J. W., Bates, J. F., Yeterian, E. H., Kennedy, D. N., & Caviness, V. S. (1999). MRI-Based topographic parcellation of human cerebral white matter and nuclei. II. Rationale and applications with systematics of cerebral connectivity. *Neuroimage*, *9*(1), 18–45.
Mayes, A. R., & Montaldi, D. (1999). The neuroimaging of long-term memory encoding processes. *Memory*, *7*(5–6), 613–659.
McEvoy, L. K., Fennema-Notestine, C., Roddey, J. C., Hagler, D. J., Jr., Holland, D., Karow, D. S., et al. (2009). Alzheimer disease: Quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology*, *251*(1), 195–205.
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*(7), 939–944.
Mickes, L., Wixted, J. T., Fennema-Notestine, C., Galasko, D., Bondi, M. W., Thal, L. J., et al. (2007). Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. *Neuropsychology*, *21*(6), 696–705.
Modrego, P. J. (2006). Predictors of conversion to dementia of probable Alzheimer type in patients with mild cognitive impairment. *Current Alzheimer Research*, *3*(2), 161–170.
Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., et al. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: A unified account based on multiple trace theory. *Journal of Anatomy*, *207*(1), 35–66.
Moss, M. B, Albert, M. S., Butters, N., & Payne, M. (1986). Differential patterns of memory loss among patients with Alzheimer's disease, Huntington's disease, and alcoholic Korsakoff's syndrome. *Archives of Neurology*, *43*(3), 239–246.
Moulin, C. J., James, N., Freeman, J. E., & Jones, R. W. (2004). Deficient acquisition and consolidation: Intertrial free recall performance in Alzheimer's disease and mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *26*(1), 1–10.
Parsons, M. W., Haut, M. W., Lemieux, S. K., Moran, M. T., & Leach, S. G. (2006). Anterior medial temporal lobe activation during encoding of words: FMRI methods to optimize sensitivity. *Brain and Cognition*, *60*(3), 253–261.
Pengas, G., Hodges, J. R., Watson, P., & Nestor, P. J. (in press). Focal posterior cingulate atrophy in incipient Alzheimer's disease. *Neurobiology of Aging*.
Petersen, R. C., Doody, R., Kurz, A., Mohs, R. C., Morris, J. C., Rabins, P. V., et al. (2001). Current concepts in mild cognitive impairment. *Archives of Neurology*, *58*(12), 1985–1992.
Petersen, R. C., & Morris, J. C. (2005). Mild cognitive impairment as a clinical entity and treatment target. *Archives of Neurology*, *62*(7), 1160–1163.
Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–308.
Powell, H. W., Koepp, M. J., Symms, M. R., Boulby, P. A., Salek-Haddadi, A., Thompson, P. J., et al. (2005). Material-specific lateralization of memory encoding in the medial temporal lobe: Blocked versus event-related design. *Neuroimage*, *27*(1), 231–239.
Powell, M. R., Smith, G. E., Knopman, D. S., Parisi, J. E., Boeve, B. F., Petersen, R. C., et al. (2006). Cognitive measures predict pathologic Alzheimer disease. *Archives of Neurology*, *63*(6), 865–868.
Rabin, L. A., Pare, N., Saykin, A. J., Brown, M. J., Wishart, H. A., Flashman, L. A., et al. (2009). Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Neuropsychology, Development, and Cognition: Section B, Aging, Neuropsychology and Cognition*, *16*(3), 357–376.
Rey, A. (1941). L'examen psychologique dans les cas d'encephalopathie traumatique. *Archives de Psychologie*, *28*, 286–340.
Rosen, A. C., Gabrieli, J. D., Stoub, T., Prull, M. W., O'Hara, R., Yesavage, J., et al. (2005). Relating medial temporal lobe volume to frontal fMRI activation for memory encoding in older adults. *Cortex*, *41*(4), 595–602.
Rosen, V. M., Sunderland, T., Levy, J., Harwell, A., McGee, L., Hammond, C., et al. (2005). Apolipoprotein E and category fluency: Evidence for reduced semantic access in healthy normal controls at risk for developing Alzheimer's disease. *Neuropsychologia*, *43*(4), 647–658.
Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology*, *60*, 257–282.
Saxton, J., Snitz, B. E., Lopez, O. L., Ives, D. G., Dunn, L. O., Fitzpatrick, A., et al. (2009). Functional and cognitive criteria produce different rates of mild cognitive impairment and conversion to dementia. *Journal of Neurology Neurosurgery and Psychiatry*, *80*(7), 737–743.
Scahill, R. I., Schott, J. M., Stevens, J. M., Rossor, M. N., & Fox, N. C. (2002). Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences the United States of America*, *99*(7), 4703–4707.
Shankle, W. R., Romney, A. K., Hara, J., Fortier, D., Dick, M. B., Chen, J. M., et al. (2005). Methods to improve the detection of mild cognitive impairment. *Proceedings of the National Academy of Sciences the United States of America*, *102*(13), 4919–4924.
Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *Institute of Electrical and Electronics Engineers Transactions on Medical Imaging*, *17*(1), 87–97.
Small, B. J., Mobly, J. L., Laukka, E. J., Jones, S., & Backman, L. (2003). Cognitive deficits in preclinical Alzheimer's disease. *Acta Neurologica Scandinavica Supplementum*, *179*, 29–33.
Smith, G. E., Ivnik, R. J., Malec, J. F., & Tangalos, E. G. (1993). Factor structure of the Mayo Older Americans Normative Sample (MOANS) core battery: Replication in a clinical sample. *Psychological Assessment*, *5*(1), 121–124.
Smith, G. E., Ivnik, R. J., Malec, J. F., Kokmen, E., Tangalos, E. G., & Kurland, L. T. (1992). Mayo's Older Americans Normative Studies (MOANS): Factor structure of a core battery. *Psychological Assessment*, *4*(3), 382–390.
Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience*, *27*, 279–306.
Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2006). Longitudinal course and neuropathologic outcomes in original vs. revised MCI and in pre-MCI. *Neurology*, *67*(3), 467–473.
Strange, B. A., Otten, L. J., Josephs, O., Rugg, M. D., & Dolan, R. J. (2002). Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. *Journal of Neuroscience*, *22*(2), 523–528.
Stuss, D. T, & Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society B-Biological Sciences*, *362*(1481), 901–915.
Twamley, E. W., Ropacki, S. A., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of International Neuropsychological Society*, *12*(5), 707–735.
Visser, M., Pluijm, S. M., Stel, V. S., Bosscher, R. J., & Deeg, D. J. (2002). Physical activity as a determinant of change in mobility performance: the Longitudinal Aging Study Amsterdam. *Journal of American Geriatrics Society*, *50*(11), 1774–1781.
Weingartner, H., Kaye, W., Smallberg, S. A., Ebert, M. H., Gillin, J. C., & Sitaram, N. (1981). Memory failures in progressive idiopathic dementia. *Journal of Abnormal Psychology*, *90*(3), 187–196.
Weintrob, D. L., Saling, M. M., Berkovic, S. F., & Reutens, D. C. (2007). Impaired verbal associative learning after resection of left perirhinal cortex. *Brain*, *130*(Pt 5), 1423–1431.
Whitwell, J. L., Petersen, R. C., Negash, S., Weigand, S. D., Kantarci, K., Ivnik, R. J., et al. (2007). Patterns of atrophy differ among specific subtypes of mild cognitive impairment. *Archives of Neurology*, *64*(8), 1130–1138.
Whitwell, J. L., Shiung, M. M., Przybelski, S. A., Weigand, S. D., Knopman, D. S., Boeve, B. F., et al. (2008). MRI patterns of atrophy associated with progression to AD in amnestic mild cognitive impairment. *Neurology*, *70*(7), 512–520.

Please cite this article in press as: Chang, Y. -L., et al. Brain substrates of learning and retention in mild cognitive impairment diagnosis and progression to Alzheimer's disease. *Neuropsychologia* (2009), doi:10.1016/j.neuropsychologia.2009.12.024