Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index

Christos Davatzikos[1], Feng Xu[1], Yang An[2], Yong Fan[1], and Susan M. Resnick[2]

[1]Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania,

[2]Laboratory of Personality and Cognition, National Institute on Aging

Running Title: AD-like brain atrophy in normal aging

Total number of words in text, including tables, figure legends and references: 7,014
Total number of words in abstract: 294

Correspond to:

Christos Davatzikos

Professor

Department of Radiology

University of Pennsylvania

3600 Market Street, Suite 380

Philadelphia, PA 19104

USA

Tel: 215-349-8587

Fax: 215-614-0266

E-Mail: christos@rad.upenn.edu

Abstract

A challenge in developing informative neuroimaging biomarkers for early diagnosis of Alzheimer's disease (AD) is the need to identify biomarkers that are evident before the onset of clinical symptoms, and which have sufficient sensitivity and specificity on an individual patient basis. Recent literature suggests that spatial patterns of brain atrophy discriminate among AD, mild cognitive impairment (MCI) and cognitively normal (CN) older adults with high accuracy on an individual basis, thereby offering promise that subtle brain changes can be detected during prodromal AD stages.

Here, we investigate whether these spatial patterns of brain atrophy can be detected in CN and MCI individuals and whether they are associated with cognitive decline. Images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) were used to construct a pattern classifier that recognizes spatial patterns of brain atrophy that best distinguish AD patients from CN on an individual person basis. This classifier was subsequently applied to longitudinal MRI scans of CN and MCI participants in the Baltimore Longitudinal Study of Aging (BLSA) neuroimaging study. The degree to which AD-like patterns were present in CN and MCI subjects was evaluated longitudinally in relation to cognitive performance.

The oldest BLSA CN individuals showed progressively increasing AD-like patterns of atrophy, and individuals with these patterns had reduced cognitive performance. MCI was associated with steeper longitudinal increases of AD-like patterns of atrophy, which separated them from CN (ROC area under the curve equal to 0.89).

Our results suggest that imaging-based spatial patterns of brain atrophy of AD, evaluated with sophisticated pattern analysis and recognition methods, may be useful in discriminating among CN individuals who are likely to be stable versus those who will

show cognitive decline. Future prospective studies will elucidate the temporal dynamics of spatial atrophy patterns and the emergence of clinical symptoms.

Keywords: early diagnosis, Alzheimer's Disease, mild cognitive impairment, brain structure, aging, SPARE-AD

Introduction

Alzheimer's disease (AD) poses significant medical, social, and socioeconomic challenges, as it is the most common dementia, with incidence rates doubling every five years after the age of 65. Although there are currently no disease-modifying treatments, many potential treatments are being tested, some of which may have significant side-effects. Thus, it is critical to identify biomarkers that identify early stages of the disease and facilitate effective and well-targeted treatment before significant neuronal damage.

Neuroimaging measures have been playing a central role in the search for biomarkers of AD that can be used for early diagnosis and monitoring of disease progression and response to therapy. Recent studies have focused on individuals with mild cognitive impairment (MCI), who have higher rates of conversion to AD (as high as 15%/year) than CN individuals (Petersen et al., 1999). Many investigators consider MCI to be early AD, as it has been shown that many MCI individuals have similar patterns of atrophy and β-amyloid deposition as AD patients (Chetelat et al., 2002; Fan et al., 2008; Karas et al., 2004; Klunk et al., 2006; Rowe et al., 2007; Ziolko et al., 2006), albeit some MCI individuals remain clinically stable over time and, consistently, some also present normal structural brain profiles (Fan et al., 2008). Given the high rates of conversion from MCI to AD and the abundant neuropathology already evident in MCI post-mortem (Mufson et al., 1999; Scheff et al., 2006), greater emphasis should be placed on identifying those CN individuals who present evolving AD-like patterns of brain atrophy and might be relatively more likely to progress to MCI and AD. Identification of such individuals at a very early stage before the onset of clinical symptoms may lead to more effective intervention of pharmacological treatments for AD as these become available.

Several longitudinal studies of normal aging have measured brain changes through regions of interest (ROI) and voxel-based analysis (Convit et al., 2000; Convit et al., 1997; Golomb et al., 1993; Mueller et al., 1998; Resnick et al., 2001; Resnick et al., 2003; Sullivan et al., 2002) and have increased our understanding about how different brain regions change in normal aging populations. Although total brain or ROI volumes may be reduced with aging and AD, their inter-individual variations and overlap across populations render them insufficient diagnostic tools for individuals, especially at early disease stages. The development of high-dimensional pattern classification methods in recent years (Fan et al., 2007b; Kloppel et al., 2008; Lao et al., 2004; Vemuri et al., 2008) offers the potential to obtain highly sensitive and specific neuroimaging biomarkers from individuals, rather than groups, which has great importance for early diagnosis and for individual patient management. These methods use sophisticated pattern analysis algorithms that are trained to identify patterns of normal or abnormal structure and function (Davatzikos et al., 2005a), which are used for classification at the individual level. We have shown previously that spatial patterns of brain atrophy discriminate between CN and AD with high accuracy (areas under the ROC curve: 0.965) (Fan et al., 2008).

Here, we investigate whether these spatial patterns of brain atrophy can distinguish among CN individuals and whether these patterns are associated with cognitive decline. The current study is the first, to our knowledge, to utilize high-dimensional pattern classification to evaluate the progression of abnormal patterns of brain atrophy in a prospectively followed cognitively normal cohort of older adults. We first train the classifier to recognize spatial patterns of brain atrophy that distinguish AD from CN individuals in the Alzheimer's Disease Neuroimaging Initiative (ADNI) cohort. The classifier produces an algorithm for determination of a quantitative value for each

individual, which we refer to as the *SPARE-AD* index (*Spatial Pattern of Abnormality for Recognition of Early AD*). More positive SPARE-AD implies a more AD-like pattern of brain atrophy, and more negative SPARE-AD implies a more normal pattern of brain morphology. The ADNI classifier is then applied to MRI scans of CN and MCI participants from the Baltimore Longitudinal Study of Aging (BLSA) to determine the presence and longitudinal progression of these patterns via longitudinal progression of the SPARE-AD index. Finally, the cognitive performance of CN individuals displaying abnormal patterns of brain atrophy is compared to CN individuals displaying normal brain structure. Since the BLSA is a prospective study, it provides the opportunity to detect very early stages of AD.

Material and Methods

BLSA participants

The BLSA is a prospective longitudinal study of aging. Its neuroimaging component, currently in its 14th year, has followed 158 individuals (age 55 -85 years at enrollment) with annual or semi-annual imaging and clinical evaluations. The neuroimaging substudy of the BLSA is described in detail in (Resnick et al., 2000; Resnick et al., 2003). Exclusionary criteria at initial evaluation were: CNS disease (epilepsy, stroke, bipolar illness, previous diagnosis of dementia), severe cardiovascular disease (myocardial infarction, coronary artery disease requiring angioplasty or bypass surgery), severe pulmonary disease, or metastatic cancer. The current study used longitudinal data from 109 BLSA participants that have remained cognitively normal up to September 2007. It also used longitudinal data from 15 individuals that were diagnosed with MCI over the course of the BLSA neuroimaging study. A diagnosis of MCI was assigned by consensus conference if a participant had deficits in either a single

cognitive domain (usually memory) or had more than one cognitive deficit but did not have functional loss in activities of daily living. Participants were evaluated at the consensus conference if their Blessed Information Memory Concentration (Blessed et al., 1968) score was greater than 3 or if their informant or subject Clinical Dementia Rating (CDR) (Morris, 1997) score was 0.5 or above. The demographic characteristics of the BLSA participants in this study are shown in Table 1.

The BLSA and neuroimaging studies are approved by the local institutional review boards, and all participants gave written informed consent prior to each assessment.

ADNI participants

The Alzheimer's Disease Neuroimaging Initiative (ADNI) is described in www.adni-info.org. The goal of ADNI is to recruit 800 adults, ages 55 to 90, approximately 200 CN older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information see www.adni-info.org. The data of all ADNI participants used in the current study have been described previously (Fan et al., 2008). Briefly, MRI scans from 66 CN individuals (mean age ± std. deviation 75.18±5.39), and 56 AD patients (77.40±7.02) were analyzed and used to construct a classifier to discriminate between CN and AD. The MMSE scores (mean ± std. deviation) of the CN and AD groups at baseline were 29.1±1.0, and 23.1±1.8, respectively. The groups were relatively well-balanced in terms of sex (50%, 57% women in CN and AD, respectively).

Image acquisition

We used T1-weighted MR images to measure regional patterns of brain atrophy. The image acquisition parameters have been described in (Resnick et al., 2001) for BLSA, and in www.adni-info.org for ADNI. Briefly, the BLSA protocol included an axial T1-

weighted volumetric spoiled gradient recalled (SPGR) series (axial acquisition, TR = 35ms, TE = 5ms, flip angle = 45^{o}, voxel dimensions of .94 X .94 X 1.5 mm slice thickness). The ADNI protocol included a sagittal volumetric 3D MPRAGE with 1.25 × 1.25 mm in-plane spatial resolution and 1.2 mm thick sagittal slices (8° flip angle). TR and TE values of the ADNI protocol were somewhat variable, but the target values were TE ~ 3.9ms and TR ~ 8.9ms.

Image analysis

Images were first preprocessed according to previously validated and published techniques (Goldszal et al., 1998). The pre-processing steps included 1) alignment to the AC-PC plane; 2) removal of extra-cranial material (skull-stripping); 3)Tissue segmentation into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF), (Pham and Prince, 1999); 4) High-dimensional image warping (Shen and Davatzikos, 2002) to a standardized coordinate system, a brain atlas (template) that was aligned with the Montreal Neurologic Institute coordinate space (Kabani et al., 1998); 5) formation of regional volumetric maps, named RAVENS maps (Davatzikos et al., 2001; Goldszal et al., 1998; Shen and Davatzikos, 2003), using tissue preserving image warping (Goldszal et al., 1998). RAVENS map intensity values quantify the regional distribution of GM, WM, and CSF, with one RAVENS map for each tissue type. In particular, RAVENS values in the template's (stereotaxic) space are directly proportional to the volume of the respective structures in the original brain scan. Therefore, regional volumetric measurements and comparisons are performed via measurements and comparisons of the respective RAVENS maps. For example, patterns of GM atrophy in the temporal lobe are quantified by patterns of RAVENS decrease in the temporal lobe in the stereotaxic space.

The RAVENS approach has been extensively validated (Davatzikos et al., 2001; Goldszal et al., 1998) and applied to a variety of studies (Beresford et al., 2006a; Beresford et al., 2006b; Driscoll et al., 2007; Gur et al., 2006; Kim et al., 2003; Resnick et al., 2001; Resnick et al., 2004; Resnick et al., 2000; Resnick et al., 2003; Stewart et al., 2006). It bears similarities with the "optimized voxel based morphometry (VBM)" approach (Good et al., 2002), except it uses a highly conforming high-dimensional image warping algorithm that captures finer structural details.

High-dimensional Classification: the SPARE-AD index as a biomarker for AD

We applied a high-dimensional pattern classification approach, which we have published and applied in a number of neuroimaging studies (Davatzikos et al., 2005b; Fan et al., 2008; Fan et al., 2007a; Fan et al., 2005; Lao et al., 2004). This approach considers all brain regions jointly and identifies a minimal set of regions whose volumes jointly maximally differentiate between CN and AD on an individual scan basis. As described in the Introduction, the pattern classification method provides a SPARE-AD index*;* positive SPARE-AD implies AD-like brain structure, and more negative SPARE-AD implies more normal structure. The pattern classifier determined the spatial patterns of brain atrophy that best distinguished AD patients from CN on an individual person basis using the ADNI sample; as anticipated, these patterns tended to reflect regional atrophy in the temporal lobe, posterior cingulate, and other areas known to be affected in AD (Fan et al., 2008).

We first evaluated the frequency of more AD-like SPARE-AD values in CN individuals for different age groups, and compared RAVENS maps of CN individuals in the upper quartile versus remaining 75% of SPARE-AD scores. To illustrate the network of regions contributing to SPARE-AD differences between CN in the top

quartile versus lower quartiles of SPARE-AD, group comparisons were performed via voxel-based statistical analysis software (http://www.fil.ion.ucl.ac.uk/spm/software/spm5) of respective RAVENS maps that were normalized by an approximation to the total intra-cranial volume (ICV), so that spatial patterns of atrophy are examined without the confounding effect of head size. This approximation to ICV was obtained by summing the volumes of GM, WM, ventricular, as well as of CSF within the sulci of the cortex that are contained within the outer brain boundary defined by the skull stripping algorithm. The approximation to ICV, which correlates (r=0.93) with the more traditional approach to definition of ICV, was employed in this analysis for consistency with the approach used in the development of the ADNI classifier. RAVENS maps were smoothed prior to statistical analysis using 8 mm full-width at half-maximum (FWHM) smoothing kernel.

Next, we evaluated the longitudinal progression of SPARE-AD in normals and in MCI by applying the classifier developed on the ADNI sample to all longitudinal MRI scans of the BLSA CN and MCI individuals, thereby allowing us to follow the evolution of the SPARE-AD index with increasing age. Mixed-effects models were used to estimate individual SPARE-AD rates of change, defined as annual changes in SPARE-AD scores. Mixed models with cognitive status (MCI versus CN) as a predictor were used to test the difference in rates for MCI versus CN.

Cognitive evaluations and associations with SPARE-AD

To determine the relationship between SPARE-AD progression and cognitive performance, we examined the SPARE-AD index values and rates of change in the SPARE-AD index in relation to performance on tests of mental status and memory. From the battery of neuropsychological tests administered to participants in conjunction

with each imaging evaluation, we selected four measures for analysis. The four measures used in the current analyses were the total score from the Mini-Mental State Exam (MMSE) (Folstein et al., 1975) to assess mental status, the immediate free recall score (sum of five immediate recall trials) and the long-delay free recall score on the California Verbal Learning Test (CVLT) (Delis et al., 1987) to assess verbal learning and immediate and delayed recall, and the total number of errors from the Benton Visual Retention Test (BVRT) (Benton, 1974) to assess short-term visual memory. We focused on these measures because changes in new learning and recall are among the earliest cognitive changes detected during the prodromal phase of AD (Grober et al., 2008). To examine relationships between SPARE-AD and cognitive performance, CN individuals in the highest quartile of SPARE-AD index values (most AD-like) were compared with the remaining sample. Relationships between SPARE-AD and the four measures of cognitive performance were examined by t-tests for unadjusted analyses and by analysis of covariance for analyses adjusted for age and sex.

Results

Prevalence of AD-Like SPARE-AD in CN

Table 2 summarizes the SPARE-AD index values as a function of age decade for the CN BLSA participants, including all 818 scans of the 109 CN. In particular, in Table 2 we have summarized the age distribution of positive SPARE-AD scores, as well as of top 25% SPARE-AD scores. We report each scan individually (numbers not in parentheses) and numbers of subjects (numbers in parentheses) that fall in each age group. For the latter, we used average SPARE-AD and average age. Percentages are reported relative to the number of scans (subjects) in the respective age range. Overall, CN below the age of 80 had negative SPARE-AD. However, more AD-like scores were

more frequent in older individuals, even though these individuals had normal cognition by clinical consensus criteria. Fig. 1 plots the mean SPARE-AD score of each of the 109 CN individuals against mean age over each participant's follow-up period. The Pearson correlation between mean SPARE-AD score and mean age is 0.43 ($p < .0001$), which is highly significant. Although the quadratic term does not reach significance, a Box-Cox transformation (with lambda = -3) provides the best fit to the data ($r = 0.44$; $p < 0.001$) and indicates the presence of a nonlinear association.

Spatial patterns of atrophy

In order to visually investigate the spatial pattern of regional volumetric differences between the CNs with the highest SPARE-AD scores (the top quartile, referred to as CN_high) and the CNs with SPARE-AD scores in the lower 75% (referred to as CN_low), we performed voxel-wise analysis of the GM and WM RAVENS maps. Fig. 2 shows regions where the CN_high showed less gray and white matter volumes, respectively, compared to the CN_low subjects. Significant decreases in tissue volumes in the more AD-like CN were evident primarily in the temporal lobe. We note that the classifier used to derive the SPARE-AD score uses regions from the temporal lobe, the cingulate and the insula, as described in {Fan, 2008 #2464}, because those are the regions that best discriminate between AD and CN. Therefore, it is reasonable that the group differences observed herein are primarily located in these regions. Fig. 2 therefore should be interpreted as a visual representation the atrophy patterns that lead to the SPARE-AD scores described in our results.

Because regions of abnormal white matter on T1 images appear dark and are typically segmented as GM, we also examined regions of GM that appeared to have greater volumes in the CN_high SPARE-AD group. The CN_high group appeared to have more GM tissue around the ventricles (Fig. 3, left), especially the posterior

periventricular regions, which are WM regions that tend to present small vessel pathology in older individuals.

Longitudinal Progression of SPARE-AD in CN

The rates of longitudinal progression of the SPARE-AD index of the 109 CNs are shown in Fig. 4, as a function of mean age during each person's follow-up period. Because the number of available follow-up scans varied considerably across individuals, mixed-effects regression was used to estimate all rates of change and shows significant increases in rates of SPARE-AD with age ($p < 0.001$). For illustration, the rates are shown as a function of age, with a Pearson correlation between rates of SPARE-AD change and mean age of 0.45 ($p < .0001$), with a significant quadratic effect ($p < 0.05$). The increasing longitudinal rates of SPARE-AD change with age are consistent with the nonlinearly increasing SPARE-AD values with age shown in the cross-sectional analysis in Fig. 1.

Longitudinal rate of change of SPARE-AD in MCI

The rates of SPARE-AD change of the 15 participants who were diagnosed with MCI over the course of the study are plotted in Fig. 5 against the age of the participants. Most of these individuals showed relatively rapid increase of the SPARE-AD index, even though their SPARE-AD scores were largely in the negative range (only 14 of the 97 scans had positive scores consistent with their relatively mild stage of impairment). Mixed-effects models were used to compare rates of change in CN versus MCI. Rates of change were significantly greater in MCI compared with CN (estimate = 0.15, SE = 0.017, $p < 0.0001$).

Separability of CN and MCI on an individual subject basis, based on rate of SPARE-AD change.

For individual patient management, it is important to be able to determine whether the individual is likely to remain stable or convert to MCI. We constructed a receiver operating characteristic (ROC) curve, by varying the threshold applied to the rate of SPARE-AD change and assigning individuals with rates of change higher than the threshold to MCI, and vice versa for CN. Fig. 6 shows the ROC curve, which achieved an area under the curve (AUC) equal to 0.89.

Relationship between cognitive performance and SPARE-AD in CN.

Cross-sectional analyses of the four cognitive measures (CVLT Sum of Immediate Free Recall, CVLT Delayed Free Recall, and BVRT errors) in relation to SPARE-AD were performed using subgroups determined from the mean SPARE-AD, the SPARE-AD at first imaging visit, and the SPARE-AD at the last imaging visit. Cognitive performance between SPARE-AD groups was compared by t-test for unadjusted index values, and by analysis of covariance for SPARE-AD adjusted for age and sex. As shown in Table 3, using mean SPARE-AD and mean cognitive performance unadjusted for age and sex, both short and long-term verbal memory scores were significantly lower ($p < 0.01$) in CN_high compared with CN_low. Using SPARE-AD groupings and cognitive scores from the first visit, MMSE in addition to CVLT performance was significantly lower ($p < 0.01$) in the CN_high compared with CN_low groups. Cognitive performance did not differ between groups using data from the last follow-up, and only the results for the MMSE at the first visit remained significant after adjusting for age and sex.

We also divided individuals by the rate of SPARE-AD change, yielding a high versus low change group. Again, CN with high rates of SPARE-AD increase (the top

25% rate of SPARE-AD change values) showed significantly lower CVLT learning and memory performance compared with CN with low rates of SPARE-AD change (Table 4). In addition, CN with high rates of SPARE-AD change had poorer performance on the MMSE, a measure of mental status, during the last visit. However, these findings were no longer significant after adjustment for age and sex.

Relationship between cognitive performance and SPARE-AD in MCI.

In the group of 15 participants diagnosed as MCI over the course of the study, the individuals with positive SPARE-AD, as well as the ones with relatively higher rate of SPARE-AD change, showed relatively worse cognitive performance. (We did not examine quartiles, due to the small sample size.) Mixed effects regression showed a significant association between the SPARE-AD index and MMSE scores of individuals diagnosed with MCI (estimate = -1.23, SE = 0.33, $p < 0.001$, Figure 7).

Discussion

A variety of neuroimaging studies have examined brain structure, as well as its longitudinal change, in cognitively normal samples and in MCI and AD via group analyses. We introduced the use of support vector machine learning approaches for classification of cognitively normal and impaired individuals at an individual level, as opposed to investigating group differences (Davatzikos et al., 2008, epub 2006; Fan et al., 2007b; Lao et al., 2004). The potential of this approach for individual classification and diagnosis has been confirmed recently by others (Duchesne et al., 2008; Kloppel et al., 2008; Vemuri et al., 2008). Our current study builds upon a computer-based pattern classification method constructed in (Fan et al., 2008) to detect spatial patterns of brain atrophy that distinguish between AD patients and CNs on an individual basis. In this study, we applied the classification algorithm that distinguished between AD patients

and CN subjects in the ADNI to a different sample of CN and MCI subjects from the BLSA. This approach generates a CN-like (negative) and AD-like (positive) SPARE-AD index of spatial atrophy patterns. We examined the frequency and longitudinal progression of AD-like spatial atrophy patterns in the BLSA cohort of cognitively normal elderly, as well as in relatively mild MCI individuals.

Our results indicate that although the vast majority of CNs have negative SPARE-AD and remain relatively stable over time, the proportion of individuals showing more AD-like, even positive, SPARE-AD increases with age. Comparisons of CN groups showing relatively higher SPARE-AD and CNs with relatively lower SPARE-AD revealed differences in spatial atrophy patterns consistent with the pattern of atrophy characteristic of AD. A strength of this study is that we examined SPARE-AD patterns in CN individuals who have been followed prospectively and remained clinically normal during the study follow-up period. Despite the lack of clinically evident impairment, CN individuals in the upper quartile of SPARE-AD, compared with the remaining CN individuals, had significantly lower performance on tests of mental status and immediate and delayed verbal memory. Declines in verbal episodic memory are among the earliest cognitive changes preceding a diagnosis of dementia, by as much as an average of seven years when investigated within the context of a prospective study (Grober et al., 2008), and the most robust gray matter differences contributing to the SPARE-AD classifier involved temporal lobe structures, which are critical for maintenance of intact memory performance. Moreover, individuals with steeper increases in the rate of SPARE-AD had lower cognitive performance. The majority of associations between cognitive performance and SPARE-AD index did not hold after adjusting for age, indicating overlap in the factors mediating spatial atrophy change and cognitive change. This is not unexpected, since cognitive decline occurs in parallel with brain tissue loss in aging populations, and age-adjustment may remove the relationship

of interest. More sophisticated dynamic modeling of longitudinal data and statistical approaches that avoid age-adjustment in larger samples may be necessary to determine whether SPARE-AD has a robust association with cognitive performance.

Notably, cross-sectional relationships between the SPARE-AD index and cognitive scores were evident for the mean values and those at the first but not last visit. The absence of associations for the last visits when participants are older is consistent with postmortem findings that AD neuropathology may show a different pattern in the oldest-old (Giannakopoulos et al., 1995). Future longitudinal follow up of these individuals, half of which are also enrolled in the BLSA autopsy study, will further elucidate the predictive value of high SPARE-AD or high rate of SPARE-AD change for AD neuropathology. However, our results indicate that the SPARE-AD index might potentially be an important early biomarker of AD progression, even before symptoms come to clinical attention. The longitudinal stability of the SPARE-AD index in CN with more negative scores, as indicated by low rates of change, indicates that SPARE-AD might be a relatively objective tool, which will assist in the evaluation of structural phenotypes associated with AD and aid in the discrimination of CN who are likely to remain stable versus those who are at greatest risk for memory impairment.

A much larger proportion of the MCI individuals showed high rate of SPARE-AD change, as expected. The MCI group also showed a relatively large and uniform spread in the range of ~0.1 to ~0.5/year, which indicates a rather rapid progression of AD-like brain atrophy. This agrees with the well documented finding that MCI individuals are quite heterogeneous, and that some will convert to AD in the following years whereas others will remain stable for a long period. As these MCI individuals were identified within the context of prospective BLSA follow-ups rather than referrals to memory clinics, they are initially studied during very early stages of impairment and have

relatively mild MCI. This is in agreement with the fact that most MCIs had negative SPARE-AD albeit many had high rates of change, indicating that rate of change may be a stronger predictor of conversion to AD. Further follow up of the entire BLSA cohort will allow us to evaluate the predictive value of the SPARE-AD and its rate of change in MCI converters to AD.

Our ability to distinguish between MCI and CN using a single value, namely the rate of SPARE-AD change, is very promising. In addition, CNs with high rates of SPARE-AD change showed lower cognitive performance; thus, CNs that were "misclassified" as MCI based on SPARE-AD index might actually develop MCI in the near future. However, we did not find any particular relationship between the exact year of conversion and the SPARE-AD. Some MCI subjects converted at low (negative) SPARE-AD values, and others at higher values after years of SPARE-AD increase. However, what was common in most MCI subjects was that they had high rates of SPARE-AD change. In view of the importance of the accurate estimation of rate of SPARE-AD change, future work in our group will emphasize the use of robust image analysis methods for estimation of rate of change. 4D segmentation and warping methods (Shen and Davatzikos, 2004; Xue et al., 2006) that have recently appeared in the literature promise to provide the foundation of future longitudinal analyses.

In Figures 2 and 3, voxel-based comparisons of more AD-like and CN-like CN individuals demonstrated greater amounts of gray matter in the periventricular regions for the more AD-like compared to CN-like group. While more AD-like CN showed the expected AD-like patterns of brain atrophy, primarily in the temporal lobe, the increase in estimated gray matter in periventricular regions highlights regions of greater white matter abnormalities. These findings are consistent with a role of increased vascular pathology underlying the progression to AD. It is important to note that the periventricular white matter signal abnormalities were not used by the classifier in

stratifying the subjects, since the classifier constructed from the ADNI AD and CN individuals (Fan et al., 2008) incorporated only temporal, prefrontal, and posterior parietal cortical regions. Therefore, subjects presenting AD-like atrophy patterns, i.e. more positive SPARE-AD scores, were identified based solely on their cortical atrophy in those regions. The observation that CN in the upper quartile of SPARE-AD scores also showed periventricular leukoareosis suggests that subjects developing AD-like atrophy developed vascular pathology in parallel. These findings are consistent with recent evidence that vascular pathology and AD-type neuropathology act in an additive manner to increase the risk for clinical dementia(Schneider et al., 2004; Troncoso et al., 2008), perhaps by increasing the likelihood that a person will cross the clinical threshold for a diagnosis of dementia. However, concurrent analysis of quantitative measures of progression of vascular disease in combination with measures of atrophy is necessary to better understand whether there might be any causal relationship between these two pathologies, or whether they simply develop in parallel.

Another contribution of the current study is that it evaluates the stability of pattern classification methods across two different large studies, which is important for the clinical applicability and generalization ability of these analysis tools across different clinics as biomarkers of AD. In particular, the CN and AD participants of the ADNI study were used to construct a classifier that recognizes AD-like patterns of brain atrophy (Fan et al., 2008), and was then applied to the BLSA, a completely independent longitudinal study of normal aging. Previous reports employing similar methods have been restricted to single studies (Davatzikos et al., 2008, epub 2006; Fan et al., 2008; Vemuri et al., 2008) and therefore do not test the generalization ability of these classifiers as biomarkers of AD. However recent studies testing similar methods across sites have begun to emerge (Kloppel et al., 2008). These studies suggest that pattern

classification methods are likely to be helpful tools in diagnosis of dementia and prognosis of its progression.

One limitation in interpreting our findings is that we do not have a gold standard for evaluation of the meaning of the positive SPARE-AD score, although we hypothesize that increasing spatial atrophy patterns will correspond to increasing AD pathology. However, our results suggest that future studies should investigate the temporal dynamics of associations between spatial patterns of atrophy, vascular disease, and neuropathology in leading to memory impairment and dementia. Prospective imaging studies, such as the BLSA neuroimaging study, in combination with autopsy assessment of neuropathology will provide important information on the temporal relationships among these cognitive and brain changes in older adults.

Acknowledgments

This study was supported in part by NIH funding sources N01-AG-3-2124 and R01-AG14971, by a grant by the Institute for the Study of Aging, and by the Intramural Research Program of the National Institute on Aging, NIH. The authors would like to thank Ms. Paraskevi Parmpi for help with data analysis and document preparation.

Group	MCI	CN
No. of subjects	15	109
Sex no. of males	10	60
Baseline age, mean (S.D.)	77.0 (7.2)	68.8 (7.7)
Age at last visit, mean (S.D.)	82.9 (7.1)	75.6 (8.1)
MMSE at first visit, mean (S.D.)	27.2 (2.5)	28.9 (1.3)
MMSE at last visit, mean (S.D.)	25.4 (3.0)	28.8 (1.2)

Table 1. Characteristics of the BLSA participants in the current study.

Age	50~59	60~69	70~79	80~89	>=90	Total
Total Number of scans (subjects)	27 (3)	304 (42)	329 (44)	146 (20)	12 (0)	818 (109)
SPARE-AD > 0						
# of scans (subjects)	0 (0)	4 (1)	5 (0)	22 (3)	3 (0)	34 (4)
% scans (subjects)	0%(0%)	1.32%(2.38%)	1.52%(0%)	15.07%(15%)	25.00%(0%)	4.16%(3.67%)
SPARE-AD in upper quartiles						
# of scans (subjects)	0 (0)	33 (6)	79 (10)	86 (11)	6 (0)	204 (27)
% of scans (subjects)	0%(0%)	10.9%(14.3%)	24%(22.7%)	58.9% (55%)	50% (NaN)	24.9% (24.8%)

Table 2. Statistics of the SPARE-AD for the total of 818 scans of all 109 CN. Numbers outside parentheses indicate results obtained by treating each scan as an individual measurement, and numbers in parentheses indicate results obtained by finding the average SPARE-AD and age of all scans of a given individual, and then using the average value as a single measurement of that individual. The total number of scans per group is shown at the top row; each subject has multiple scans, one per visit. The total number of subjects per group is shown at the top row in parentheses. Percentages were calculated relative to the total number of scans (subjects) in each age category.

	Mean scores	Mean Scores Adjusted (for sex and age)	First Visit Scores	First Visit Scores Adjusted	Last Visit Scores	Last Visit Scores Adjusted
CVLT List A Sum	0.0075	0.5155	0.0020	0.1797	0.3357	0.3754
CVLT Long Delay Free	0.0067	0.3150	0.0067	0.2674	0.1817	0.8458
BVRT Errors	0.0930	0.9728	0.3505	0.8944	0.5994	0.5592
MMSE	0.1719	0.9666	0.0092	0.0174	0.5203	0.8606

Table 3. P-values of cross-sectional differences in cognitive performance between the CN subjects having the top 25% of the SPARE-AD scores and the remaining 75%. Column labels indicate the scores used in the analysis, e.g. “mean scores” indicates the p-value is based on mean cognitive performance and mean SPARE-AD index over time. Covariates in adjusted comparisons are baseline age and sex. CVLT California Verbal Learning Test; BVRT Benton Visual Retention Test; MMSE Mini-Mental State Exam.

	Mean Cognitive Score	Mean Cognitive Score - Adjusted	First Visit Cognitive Score	First Visit Cognitive Score - Adjusted	Last Visit Cognitive Score	Last Visit Cognitive Score - Adjusted
CVLT List A Sum	0.0295	0.4784	0.0250	0.3792	0.0190	0.2780
CVLT Long Delay Free	0.0448	0.5288	0.0455	0.4476	0.0066	0.1018
BVRT Errors	0.3987	0.4906	0.6139	0.6453	0.3302	0.9097
MMSE	0.2020	0.8316	0.3475	0.8306	0.0579	0.1901

Table 4. P-values of differences in cognitive performance between the CNs with high rates of SPARE-AD change and CNs with low rates of SPARE-AD change (remaining 75%). Covariates in adjusted comparisons are baseline age and sex. CVLT California Verbal Learning Test; BVRT Benton Visual Retention Test; MMSE Mini-Mental State Exam.

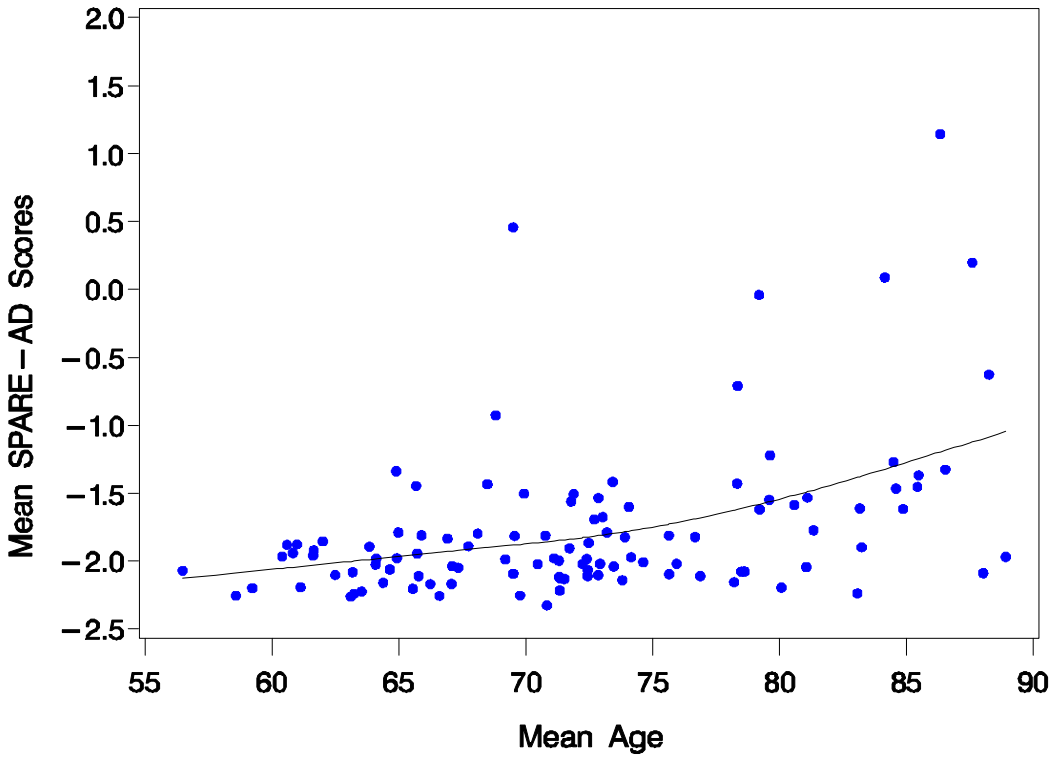
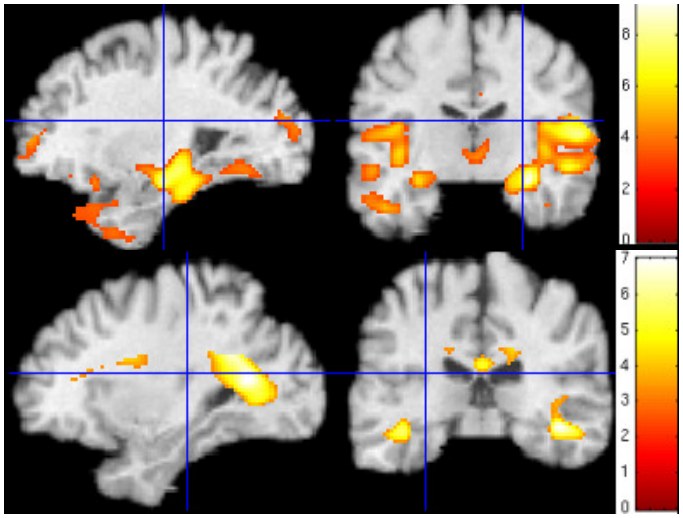


Fig. 1. Mean SPARE-AD scores of each of the 109 CN individuals plotted against mean age over their follow-up period.



(b)

Fig. 2. T-statistic voxel-wise maps of RAVENS GM (top) and WM (bottom) comparing the 75% of CNs with the lowest SPARE-AD scores (CN-like CNs) minus those with the top 25% SPARE-AD (relatively more AD-like CNs).

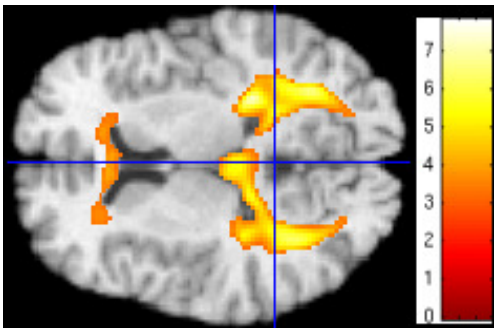


Fig. 3. Regions in which CNs with the top 25% (highest) SPARE-AD scores had significantly higher GM RAVENS maps than the bottom 75%, indicating increased peri-ventricular abnormal WM tissue that appears gray in T1-weighted images and is segmented as GM.

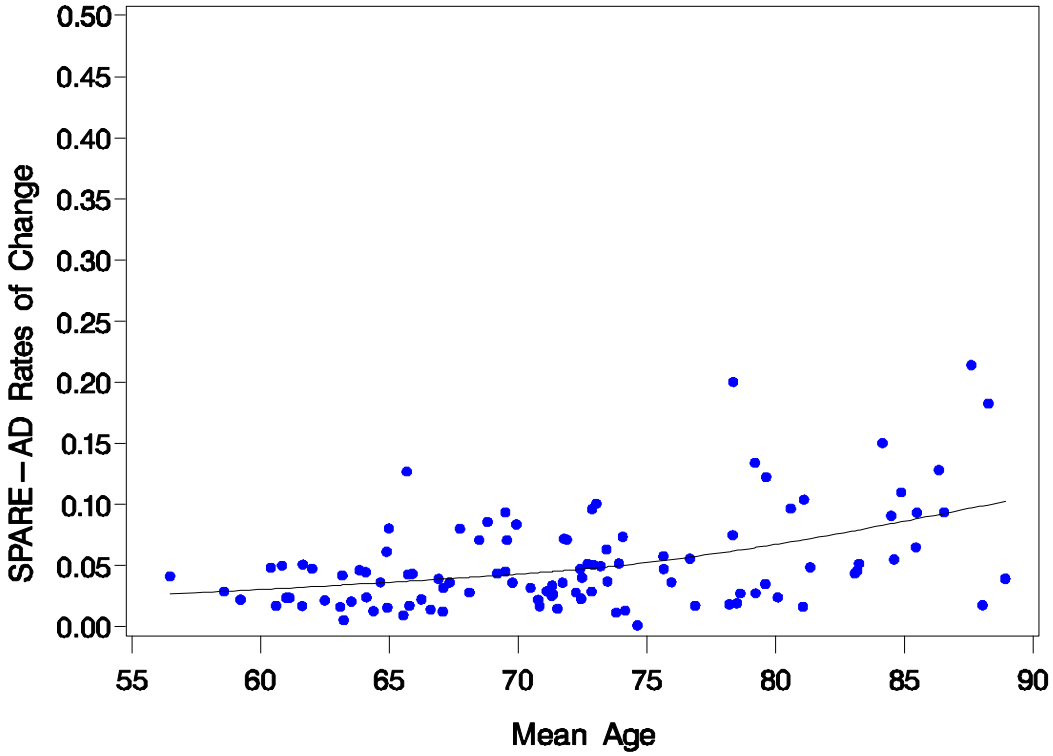


Fig. 4. Rate of SPARE-AD change as a function of average age during follow-up period, for the 109 CN individuals

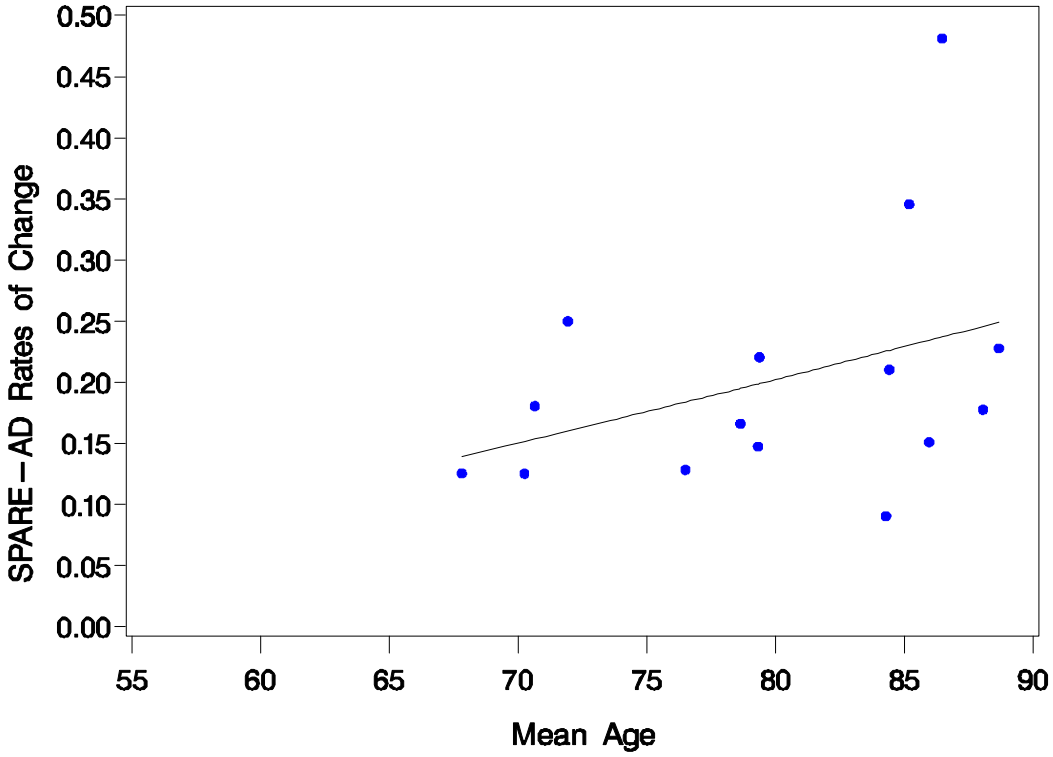


Fig. 5. Rate of SPARE-AD change as a function of age in 15 MCI participants.

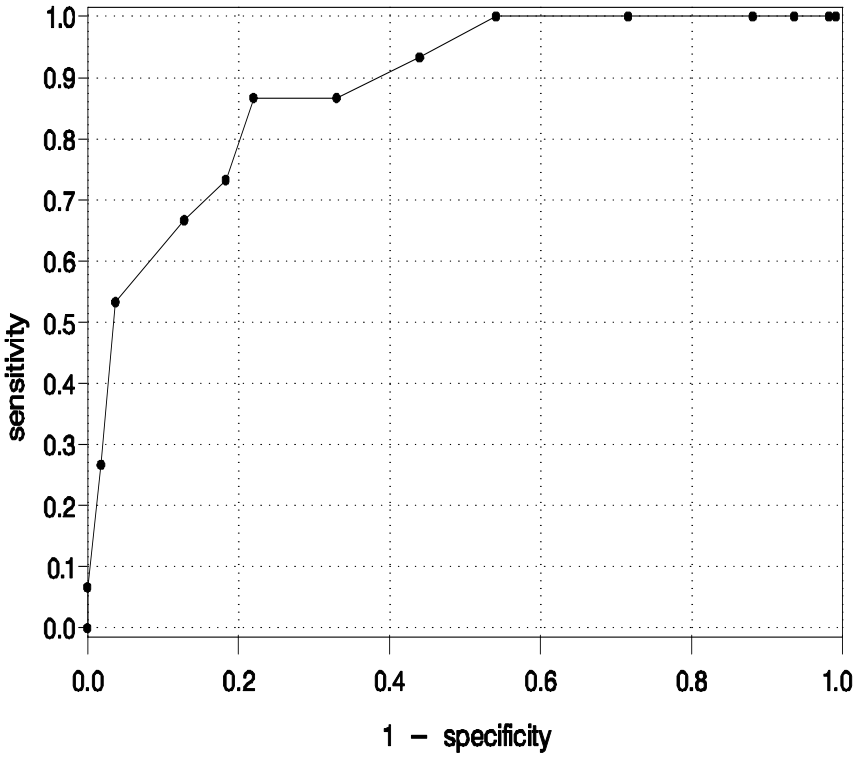


Fig. 6. ROC curve of individual subject classification to CN or MCI, based on the rate of SPARE-AD change. AUC=0.885

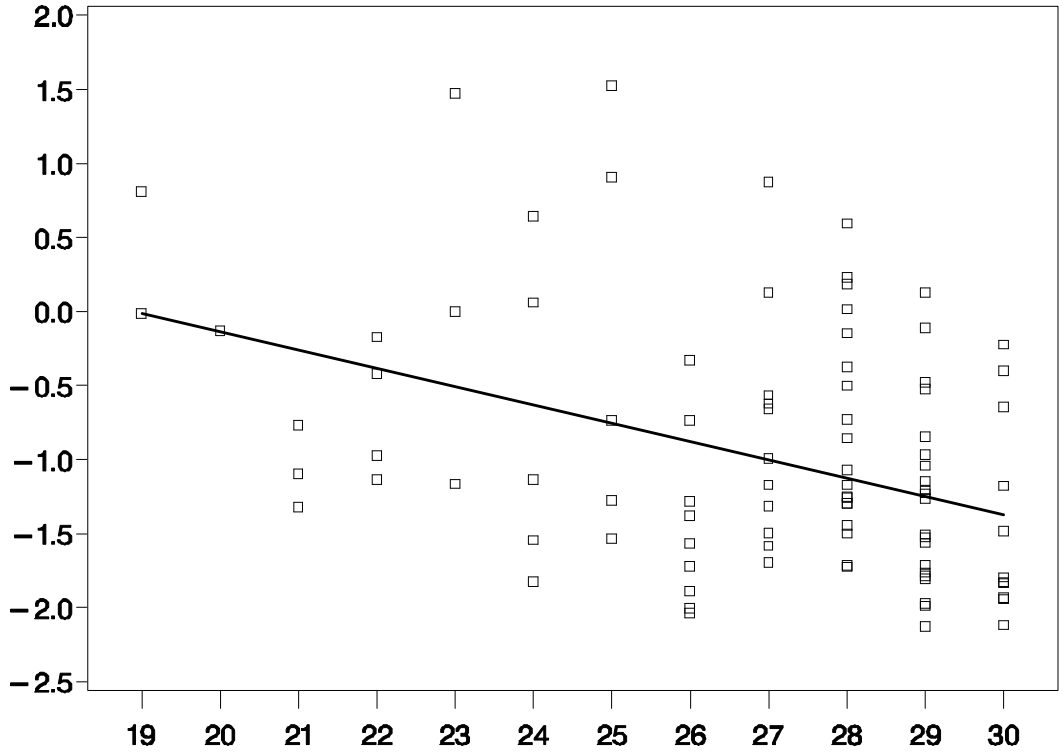


Fig. 7. SPARE-AD index values plotted against MMSE scores for all MCI individuals.

The linear effect ($p < 0.001$) was determined using mixed-effects regression

References

Benton A. Revised Visual Retention Test. New York: Psychological Corporation, 1974.

Beresford T, Arciniegas D, Alfers J, Clapp L, Martin B, Beresford H, et al. Hypercortisolism In Alcohol Dependence and Its Relation to Hippocampal Volume Loss. Journal of Studies on Alcohol 2006a; 67: 861-867.

Beresford TP, Arciniegas DB, Alfers J, Clapp L, Martin B, Du Y, et al. Hippocampus Volume Loss Due to Chronic Heavy Drinking. Alcoholism: Clinical and Experimental Research 2006b; 30: 1866-1870.

Blessed G, Tomlinson B, Roth M. The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. British Journal of Psychiatry 1968; 114: 797-811.

Caffo B, Chen S, Stewart W, Bolla K, Yousem D, Davatzikos C, et al. Are brain volumes based on magnetic resonance imaging mediators of the associations of cumulative lead dose with cognitive function? American Journal of Epidemiology 2008; 167: 429-437.

Chetelat G, Desgranges B, de la Sayette V, Viader F, Eustache F, Baron J-C. Mapping gray matter loss with voxel-based morphometry in mild cognitive impairment. Neuroreport 2002; 13: 1939-1943.

Convit A, de Asis J, de Leon MJ, Tarshish CY, De Santi S, Rusinek H. Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer's disease. Neurobiology of Aging 2000; 21: 19-26.

Convit A, De Leon MJ, Tarshish C, De Santi S, Tsui W, Rusinek H, et al. Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease. Neurobiology of Aging 1997; 18: 131-8.

Davatzikos C, Fan Y, Wu X, Shen D, Resnick SM. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging. Neurobiology of Aging 2008, epub 2006; 29: 514-523.

Davatzikos C, Genc A, Xu D, Resnick SM. Voxel-Based Morphometry Using the RAVENS Maps: Methods and Validation Using Simulated Longitudinal Atrophy. NeuroImage 2001; 14: 1361-1369.

Davatzikos C, Ruparel K, Fan Y, Shen D, Acharyya M, Loughead J, et al. Classifying spatial patterns of brain activity for lie-detection. Neuroimage 2005a; 28: 663-668.

Davatzikos C, Shen DG, Wu X, Lao Z, Hughett P, Turetsky BI, et al. Whole-brain morphometric study of schizophrenia reveals a spatially complex set of focal abnormalities. JAMA Archives of General Psychiatry 2005b; 62: 1218-1227.

Delis D, Kramer J, Kaplan E, Ober B. California Verbal Learning Test - Research Edition. New York: The Psychological Corporation, 1987.

Driscoll I, Davatzikos C, An Y, Wu X, Shen D, Kraut M, et al. Longitudinal brain changes in cognitively impaired and unimpaired older adults. Society of neuroscience. San Diego, CA, 2007.

Duchesne S, Caroli A, Geroldi C, Barillot C, Frisoni GB, Collins DL. MRI-Based Automated Computer Classification of Probable AD Versus Normal Controls. IEEE Transactions on Medical Imaging 2008; 27: 509-520.

Fan Y, Batmanghelich N, Clark CM, Davatzikos C, the Alzheimer's Disease Neuroimaging Initiative. Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional

pattern classification, predict subsequent cognitive decline. Neuroimage 2008; 39: 1731-1743.
Fan Y, Rao H, Hurt H, Giannetta J, Korczykowski M, Shera D, et al. Multivariate examination of brain abnormality using both structural and functional MRI. Neuroimage 2007a; 36: 1189-1199.
Fan Y, Shen D, Davatzikos C. Classification of Structural Images via High-Dimensional Image Warping, Robust Feature Extraction, and SVM. In: Duncan JS and Gerig G, editors. MICCAI. Vol 3749 / 2005. Palm Springs, California, USA: Springer Berlin / Heidelberg, 2005: 1-8.
Fan Y, Shen D, Gur RC, Gur RE, Davatzikos C. COMPARE: Classification Of Morphological Patterns using Adaptive Regional Elements. IEEE Transactions on Medical Imaging 2007b; 26: 93-105.
Folstein M, Folstein S, McHugh P. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 1975; 12: 189-193.
Giannakopoulos P, Hof PR, Vallet PG, Giannakopoulos A-S, Charnay Y, Bouras C. Quantitative analysis of neuropathologic changes in the cerebral cortex of centenarians. Progress in Neuro-Psychopharmacology and Biological Psychiatry 1995; 19: 577-592.
Goldszal AF, Davatzikos C, Pham D, Yan M, Bryan RN, Resnick SM. An image processing protocol for the analysis of MR images from an elderly population. Journal of Computer Assisted Tomography 1998; 22: 827-837.
Golomb J, deLeon MJ, Kluger A, George AE, Tarshish C, Ferris SH. Hippocampal atrophy in normal aging: an association with recent memory impairment. Archives of Neurology 1993; 50: 967-973.
Good CD, Scahill RI, Fox NC, Ashburner J, Friston KJ, Chan D, et al. Automatic differentiation of anatomical patterns in the human brain: Validation with studies of degenerative dementias. Neuroimage 2002; 17: 29-46.
Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. Journal of the International Neuropsychological Society 2008; 14: 266-78.
Gur R, Davatzikos C, Shen D, Wu X, Fan Y, Hughett P, et al. Whole-brain deformation based morphometry MRI study of schizophrenia. SCHIZOPHRENIA BULLETIN 2006; 31: 408-408.
Kabani N, MacDonald D, Holmes CJ, Evans A. A 3D atlas of the human brain. Neuroimage 1998; 7: S717.
Karas GB, Scheltens P, Rombouts SARB, Visser PJ, Schijndel RAv, Fox NC, et al. Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease. Neuroimage 2004; 23: 708-716.
Kim J-S, Kanaan R, Kaufmann W, Ross C, Calhoun V, Xu D, et al. Abnormal White Matter Organization in Huntington's Disease Evaluated With Diffusion Tensor MRI. ISMRM. Toronto, Canada, 2003.
Kloppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, et al. Automatic classification of MR scans in Alzheimer's disease. Brain 2008; 131: 681-689.
Klunk WE, Mathis CA, Price JC, Lopresti BJ, DeKosky ST. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. Brain 2006; 129: 2805-7.

Lao Z, Shen D, Xue Z, Karacali B, Resnick SM, Davatzikos C. Morphological classification of brains via high-dimensional shape transformations and machine learning methods. Neuroimage 2004; 21: 46-57.
Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. International Psychogeriatrics 1997; 9 Suppl 1: 173-6; discussion 177-8.
Mueller EA, Moore MM, Kerr DC, Sexton G, Camicioli RM, Howieson DB, et al. Brain volume preserved in healthy elderly through the eleventh decade. Neurology 1998; 51: 1555-1562.
Mufson EJ, Chen E-Y, Cochran EJ, Beckett LA, Bennett DA, Kordower JH. Entorhinal cortex beta-amyloid load in individuals with mild cognitive impairment. Exp Neurol 1999; 158: 469-90.
Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild Cognitive Impairment: Clinical Characterization and Outcome. Archives of Neurology 1999; 56: 303-308.
Pham DL, Prince JL. Adaptive fuzzy segmentation of magnetic resonance images. IEEE Transactions on Medical Imaging 1999; 18: 737-752.
Resnick S, Davatzikos C, Kraut M, Zonderman A. Longitudinal changes in MRI volumes in older adults. Neurobiology of Aging 2001; 22: 5.
Resnick S, Pham D, Davatzikos C, Kraut M. Sex differences in regional cerebral blood flow: Clinical implications for Alzheimer's disease. Neurobiology of Aging 2004; 25: 263, 2004.
Resnick SM, Goldszal A, Davatzikos C, Golski S, Kraut MA, Metter EJ, et al. One-year age changes in MRI brain volumes in older adults. Cerebral Cortex 2000; 10: 464-472.
Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C. Longitudinal Magnetic Resonance Imaging Studies of Older Adults: A Shrinking Brain. The Journal of Neuroscience 2003; 23: 295-301.
Rowe CC, Ng S, Ackermann U, Gong SJ, Pike K, Savage G, et al. Imaging beta-amyloid burden in aging and dementia. Neurology 2007; 68: 1718-25.
Scheff SW, Price DA, Schmitt FA, Mufson EJ. Hippocampal synaptic loss in early Alzheimer's disease and mild cognitive impairment. Neurobiol Aging 2006; 27: 1372-84.
Schneider JA, Wilson RS, Bienias JL, Evans DA, Bennett DA. Cerebral infarctions and the likelihood of dementia from Alzheimer disease pathology. Neurology 2004; 62: 1148-55.
Shen D, Davatzikos C. HAMMER: Hierarchical attribute matching mechanism for elastic registration. IEEE Transactions on Medical Imaging 2002; 21: 1421-1439.
Shen D, Davatzikos C. Measuring Temporal Morphological Changes Robustly in Brain MR Images Via 4-Dimensional Template Warping. NeuroImage 2004; 21: 1508-1517.
Shen DG, Davatzikos C. Very high resolution morphometry using mass-preserving deformations and HAMMER elastic registration. NeuroImage 2003; 18: 28-41.
Stewart WF, Schwartz BS, Davatzikos C, Shen D, Liu D, Wu X, et al. Past Adult Lead Exposure is Linked to Neurodegeneration Measured by Brain MRI. Neurology 2006; 66: 1476-1484.
Sullivan EV, Pfefferbaum A, Adalsteinsson E, Swan GE, Carmelli D. Differential Rates of Regional Brain Change in Callosal and Ventricular Size: a 4-Year Longitudinal MRI Study of Elderly Men. Cerebral Cortex 2002; 12: 438-445.
Troncoso JC, Zonderman AB, Resnick SM, Crain B, Pletnikova O, O'Brien RJ. Effect of infarcts on dementia in the Baltimore longitudinal study of aging. Annals of Neurology 2008.
Vemuri P, Gunter JL, Senjem ML, Whitwell JL, Kantarci K, Knopman DS, et al. Alzheimer's disease diagnosis in individual subjects using structural MR images: Validation studies. Neuroimage 2008; 39: 1186-1197.

Xue Z, Shen D, Davatzikos C. CLASSIC: Consistent Longitudinal Alignment and Segmentation for Serial Image Computing. Neuroimage 2006; 30: 388-399.
Ziolko SK, Weissfeld LA, Klunk WE, Mathis CA, Hoge JA, Lopresti BJ, et al. Evaluation of voxel-based methods for the statistical analysis of PIB PET amyloid imaging studies in Alzheimer's disease. Neuroimage 2006; 33: 94-102.