Published in final edited form as:
Hum Brain Mapp. 2014 September ; 35(9): 4777–4794. doi:10.1002/hbm.22511.

Analysis of Spatio-temporal Brain Imaging Patterns by Hidden Markov Models and Serial MRI Images

Ying Wang[a,*], **Susan M. Resnick**[b], **Christos Davatzikos**[a], and **for the Baltimore Longitudinal Study of Aging and the Alzheimer's Disease Neuroimaging Initiative1**

[a]Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, USA

[b]Laboratory of Behavioral Neuroscience, National Institute on Aging, Baltimore, MD 21224, USA

Abstract

Brain changes due to development and maturation, normal aging or degenerative disease are continuous, gradual and variable across individuals. To quantify the individual progression of brain changes, we propose a spatio-temporal methodology based on Hidden Markov Models (HMM), and apply it on 4D structural brain MRI series of older individuals. First, regional brain features are extracted in order to reduce image dimensionality. This process is guided by the objective of the study or the specific imaging patterns whose progression is of interest, e.g. the evaluation of Alzheimer's-like patterns of brain change in normal individuals. These regional features are used in conjunction with HMMs, which aim to measure the dynamic association between brain structure changes and progressive stages of disease over time. A bagging framework is used to obtain models with good generalization capability, since in practice the number of serial scans is limited. An application of the proposed methodology was to detect individuals with the risk of developing MCI, and therefore it was tested on modeling the progression of brain atrophy patterns in older adults. With HMM models, the state-transition paths corresponding to longitudinal brain changes were constructed from two completely independent datasets, the Alzheimer's Disease Neuroimaging Initiative and the Baltimore Longitudinal Study of Aging. The statistical analysis of HMM-state paths among the normal, progressive MCI and MCI groups indicates that, HMM-state index 1 is likely to be a predictor of the conversion from cognitively normal to MCI, potentially many years before clinical symptoms become measurable.

Keywords

High-dimensionality spatio-temporal analysis; Brain pattern analysis; MRI series; Mild cognitive impairment; Early MCI predictor

[1]Part data used in the preparation of this article was obtained from the Baltimore Longitudinal Study of Aging (BLSA), and part data were from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. 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 paper.

Correspondence to: Dr. Ying Wang, Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104, Telephone: +1 215 662 7348, Fax: +1 215 614 0266, ying.wang@uphs.upenn.edu.

We have no conflict of interest to declare.

1. Introduction

Longitudinal changes in brain structure and function might be related to growth, aging, or progression of disease. The associations of brain changes with clinically apparent changes vary significantly. For example, normal aging is accompanied by an evolving process of brain atrophy and functional change. Neurodegenerative disorders, such as MCI, involve even more complex brain changes with accelerated volume loss in specific brain regions. Compared with cognitively normal (CN) individuals, MCI individuals have higher risk of conversion to Alzheimer's disease (AD), with an approximately 15% transition rate annually (Petersen et al., 1999). Furthermore, AD shows the gradual accumulation of amyloid plaques and neurofibrillary tangles, followed by structural and functional changes (Wenk, 2003; Bredesen et al., 2006). In this paper, we aim to develop a general method that can efficiently capture continuously evolving brain changes and can ultimately produce individual-based indicators that might relate to disease progression or normal aging. In particular, we examined and applied the proposed method for the early detection of MCI.

High-dimensionality pattern regression methods have played an increasingly important role in characterizing brain change from images in recent years (Formisano et al., 2008; Duchesne et al., 2009; Franke et al., 2010; Wang et al., 2010a). Compared with classification studies to identify individuals into discrete categories, such as disease and non-disease (Golland et al., 2002; Fan et al., 2007; Kloppel et al., 2008; Vemuri et al., 2009; Plant et al., 2009; Davatzikos et al., 2009), pattern regression methods have been found very promising for describing continuously evolving processes, and have been applied to individual-based diagnosis or prediction studies. However, only a baseline scan of each sample was used in these regression studies, which rendered them unavailable to the temporal aspects of a specific individuals' progression. Alternatively, longitudinal analysis of brain image series, which capture the temporal dynamics of brain change, has also been investigated (Sullivan et al., 2002; Resnick et al., 2003; Chtelat et al., 2005; Misra et al., 2009; Driscoll et al., 2009; Sluimer et al., 2010). Most of these studies have been based on group analysis, where the objective was to identify group differences in serial scans. However, as the rate of progression varies across subjects, it is important to derive the stage of progression individually.

Thus, one of the primary goals of this study is to investigate a novel way to characterize the longitudinal progression of brain changes on an individual basis. To achieve it, we propose a 4D serial image analysis method, which also captures the dynamic relationship between clinical appearances and structural brain changes latent in longitudinal MRI series. Due to the high dimensionality of brain images and relatively small size of samples typically available, it is necessary to extract a relatively small set of brain regions that are correlated with anatomical changes that are of interest, prior to evaluating imaging patterns. In this study, we are interested in quantifying the longitudinal progression of brain atrophy patterns in older adults. An adaptive algorithm was adopted to extract regional brain patterns of atrophy that distinguish cognitively normal elderly from AD patients, which was originally used in (Davatzikos et al., 2009). Specially, brain voxels are clustered into regions by measuring a relatively homogeneous correlation between brain image information and clinical variables at each voxel.

Besides the extraction of descriptive and robust brain features, in order to analyze the time-varying changes of brain, an efficient tool for spatio-temporal analysis is necessary. Hidden Markov Models (HMMs) have been found to be among the most powerful approaches for representing dynamic signals (Rabiner, 1989; Moeslund et al., 2006). In particular, HMM is considered as a kind of dynamic Bayesian network, which makes the model well-suited for handling uncertainties of changing signals. Furthermore, given the trained HMM with probabilistic parameters, the transition state path for each subject can be generated to interpret the spatio-temporal changes in a concise way (Forney, 1973). HMMs have been successfully applied in the pattern recognition field, such as speech recognition and activity analysis (Rabiner, 1989; Moeslund et al., 2006). Additionally, Markov chain has been used to represent disease progression through a series of clinical events. Typical diseases modeled by Markov chain include cancer (Newton et al., 2013) and brain degenerative diseases, such as Alzheimer's disease, Huntington's disease (Fonteijn et al., 2012), and Parkinson's disease (Costin et al., 2013). We therefore hypothesized that HMMs are good models to measure dynamic changes in longitudinal brain image data, as well as represent these changes latent in images with concise state-transition paths.

Due to the relatively small number of training subjects that are typically available, training HMM models might be very sensitive to outliers, even though HMMs have been well validated as an effective spatio-temporal model with good generalization ability (Rabiner, 1989; Moeslund et al., 2006). In order to improve the stability, a bagging strategy is adopted, which computes an ensemble combination of models trained on random sub-samples of an initial training set. In this approach, ensemble HMM models are generated, and state paths for each subject are estimated statistically (Breiman, 1996a, b; Davis and Lovell, 2004).

Another goal of this study is to provide predictors that have the potential to detect individuals in the preclinical phase; for example, to identify older individuals who show more pronounced patterns of brain atrophy, but still seem clinically normal during the observation period. To this end, HMMs based on Alzheimer's-like patterns of brain atrophy in older CN and MCI individuals, were trained. Then the corresponding hidden state-transition paths of subjects from different clinical groups, e.g. normal controls, very mild MCI and MCI patients, were compared and analyzed. A state of HMM might be a predictor of CN to MCI conversion, because the state-transition paths of individuals from the same group were expected to show the tendency to converge to the same state.

The remainder of this paper is structured as follows: We first introduce the bagging-based methodology including brain image preprocessing and regional analysis, the proposed HMM structure and model performance evaluation measures in the Method section. Data section provides more detail about two completely independent datasets, in terms of demographics and image acquisition procedures. Results and analysis section summarizes the results of HMM state transition paths, their relationships with cognitive performance, and predictors of early MCI conversion. Discussion and conclusion section discuss the implications of our findings and provide future perspectives.

2. Methods

Although the proposed methodology based on regional feature extraction and HMM is quite general, we examine it on structural brain MRI series with application to the early detection of MCI in older individuals.

2.1. Image pre-processing

The FMRIB's brain extraction tool (BET) is used to remove the skull from image (Smith S. M., 2002), then manual correction is applied to inaccuracies in brain extraction. Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF) are segmented from each skull-stripped brain MRI (Pham et al., 1999). Each segmented brain tissue image is spatially normalized to a brain atlas (Jacob template) that was aligned with the MNI coordinate space (Kabani et al., 1998; Fan et al., 2008) via HAMMER registration (Shen and Davatzikos, 2002). Then, regional volumetric maps, named RAVENS maps are generated using tissue preserving image warping (Davatzikos et al., 2001), which have been extensively validated by a number of pattern analysis studies (Fan et al., 2007; Davatzikos et al., 2009; Wang et al., 2010a). The values of RAVENS maps are directly proportional to the amount of original tissue volume for each region. Therefore, RAVENS maps give a quantitative representation of the spatial distribution of each tissue type. They are akin to modulated tissue density maps commonly used in voxel-based morphometry (VBM) approaches (Ashburner and Friston, 2000; Good et al., 2001, 2002), except using a different registration method. In summary, measurements of RAVENS maps can provide the regional volumetric measurements of the corresponding tissue maps. Individual intracranial volume (ICV) normalization was applied on RAVENS maps to adjust global differences in intracranial size. Last, normalized RAVENS maps were down-sampled and smoothed to incorporate neighborhood information using an 8mm FWHM Gaussian filter.

2.2. Adaptive regional feature extraction and selection

Compared with voxel-wise features, adaptive regional volumetric measurements generate more robust patterns against measurement noise or image pre-processing errors. In order to generate regions of interest, one common way is to gather tissue voxels with relatively homogeneous characteristics by statistical analysis, such as the correlation between voxel-wise values of images and progressive clinical status (e.g. class labels, cognitive test scores, etc.). A spatially stable Pearson correlation has been extensively used by a variety of imaging studies to relate imaging characteristics to clinical variables, including AD and CN classification (Fan et al., 2007), detection of prodromal MCI or AD (Davatzikos et al., 2009; Misra et al., 2009), and regression for continuous development from CN to MCI to AD (Wang et al., 2010a). However, biases are often introduced in the calculation of the correlations due to the limited number of samples. Herein, we used a robust measure to evaluate the voxel relevance with progressive stages, namely Correlation Confidence (Wang et al., 2010b). This measure is calculated as the ratio of the mean and variance of Pearson correlation coefficients calculated for the same location. By extensively examining voxels and their respective correlation coefficients, correlation confidence offers not only the discriminative ability, but also the consistency of patterns. Outliers can be found via high variance of correlation coefficients, which can pick a single point with an extremely high

correlation coefficient, but significantly different from the average. Under the leave-k-out bagging procedure, correlation confidence at a specific voxel from all training images is calculated as follows:

$$c^i(u)=\frac{\overline{c_n^i(u)}}{Var(c_n^i(u))} \quad (1)$$

where

$$c_n^i(u)=\frac{\sum_{j=1}^{m}(f_{n,j}^i(u)-\overline{f_n^l(u)})(y_{n,j}^i-\overline{y_n})}{\sqrt{\sum_{j=1}^{m}(f_{n,j}^i(u)-\overline{f_n^l(u)})^2\sum_{j=1}^{m}(y_{n,j}^i-\overline{y_n})^2}} \quad (2)$$

$c_n^i(u)$ is the Pearson correlation coefficient between the tissue values $f_n^i(u)$ and clinical variables y_n^i at location *u* of tissue map *i* from the *n*th leave-k-out case, from which *k* images are excluded (totally there are *m*+*k* images, m is the number of samples included). Here, $\overline{f_n^l(u)}$ is the mean of $f_{n,j}^i(u)$ over all samples in the *n*th iteration set, and $\overline{y_n}$ is the mean of all samples. Additional details about region clustering and feature representation can be found in our previous work (Fan et al., 2007; Wang et al., 2010a). We highlight that only baseline scans of the AD and NC participants from the ADNI study were applied to create regional features, which then were applied to the longitudinal scans of BLSA participants.

2.3. Hidden Markov Model

As mentioned earlier, we aim to characterize individual longitudinal progression along a direction, e.g. aging, maturation, or disease progression. Furthermore, we want to increase the understanding of dataset, and therefore provide predictors that reflect the underlying evolving process or more specifically, the clinical stages of a brain disease. The HMM model is a well-documented tool to address these problems, and has been validated in many spatiotemporal pattern recognition tasks (Rabiner, 1989; Moeslund et al., 2006).

An HMM model is a Markov process by unobserved (hidden or latent) states. The state is not directly visible, but can only be observed through a set of stochastic processes that produce the sequence of observations, as shown in Figure 1. The relationship between states and observations is modelled by the probability distributions of observation. In general, the probability density function of the observations is represented by a finite mixture of Gaussian distributions. In this study, the observations of the HMM model are continuous, while the hidden state space is discrete. Specifically, observation sequences are the features extracted from longitudinal brain scans, and the state space corresponds to the progressive stages of an individual's development we aim to estimate.

An HMM is characterized with a set of parameters (N, A, B, π), where N is the number of hidden states in the model, A is the state transition probability distribution, B is the observation probability distribution, and π is the initial state probability distribution. Generally the states are fully connected in HMM model, which means that any state could

be reached from any other state. Here we denote the state at time t as $q_t \in S = \{s_1, s_2, \ldots, s_N\}$.

The state transition probability distribution $A = \{a_{ij}\}$:

$$a_{ij}=P(q_{t+1}=s_j|q_t=s_i), 1 \leq i,j \leq N \quad (3)$$

The observation probability distribution $B = \{b_j(O)\}$:

$$b_j(O)=\sum_{m=1}^{M} c_{jm}\Pi(O,\mu_{jm},\nu_{jm}), 1 \leq i,j \leq N \quad (4)$$

where $O = \{o_1, o_2, \ldots, o_T,\}$ is the observation vector being modeled, T is the number of observations in this sequence, M is the number of Gaussian mixtures, c_{jm} is the mixture coefficient for the mth mixture in state s_j and Π is any Gaussian density with mean vector μ_{jm} and covariance matrix ν_{jm} for the mth mixture component in state s_j. The mixture gains c_{jm} satisfy the stochastic constraint:

$$\sum_{m=1}^{M} c_{jm}=1, 1 \leq i,j \leq N \quad (5)$$

The initial state probability distribution $\pi = \{\pi_i\}$:

$$\pi_i=P(q_1=S_i), 1 \leq i,j \leq N \quad (6)$$

For convenience, we use the compact notation $\lambda = (A, B, \pi)$ to indicate the complete parameter set of the HMM model. Given a set of observation sequences, the parameter learning task in HMMs is to find the best set of λ such that $P(O|\lambda)$ is locally maximized. The Baum-Welch (BW) algorithm is a special case of the Expectation-Maximization (EM) algorithm, and it allows us to derive the maximum likelihood estimates of the parameters of the HMM (L. E. Baum and Weiss, 1970; Welch, 2003). More details about BW algorithm can be found in (Rabiner, 1989).

Once the optimal parameter set λ of HMM model has been estimated, we use it to find a single best state sequence $Q = \{q_1, q_2, \ldots, q_t\}$ for the given observation sequence $O = \{O_1, O_2, \ldots, O_t\}$, such that $P(Q|O, \lambda)$ is maximized. The Viterbi algorithm and posterior decoding are used to achieve this aim. Both algorithms are detailed in (Rabiner, 1989; Forney, 1973).

Specifically in neurodegenerative diseases, brain changes usually accompany irreversible tissue volume decline. For this reason, in this paper we used a left-right HMM, in which the state index of the model either decreases or remains the same as time increases (Rabiner, 1989). Thus the state transition matrix of the left-right HMM has the following form:

$$a_{ij}=0, 1 \leq j \leq i \leq N \quad (7)$$

Clearly the left-right HMM has the desirable property to model uni-directional changes over time, such as the progression of brain atrophy.

After training the HMM model parameters, optimal state paths are obtained for a set of subjects. Then we investigate the association between these state paths and the clinical stages of the subjects. For example, we can analyze and compare the differences of state-transition frequency for different clinical groups, such as the MCI and CN groups. This analysis also provides a prognostic indicator for individuals with the risk of MCI, before the onset of clinical diagnosis.

2.4. Bagging framework

A bagging strategy is used to improve the generalization ability of the model (Breiman, 1996a,b). Within each leave-k-out loop, an HMM model is constructed with the associated parameters (π_n, a_n, b_n), where n is the iteration index. For a given subject, the most likely state path is decoded for each HMM model λ_n. The final state path is calculated by selecting the state that occurs most frequently at each time point.

In this study, our HMM modeling was performed using leave-10-out cross-validation instead of leave-1-out scheme. Obviously, leave-many-out setting offers better generalization ability and more reliable models.

2.5. Evaluation of method across ADNI and BLSA studies

In this paper, we evaluated the proposed method across two independent large datasets, which was important for the generalization ability of the analysis tool, and the robustness of indicators or biomarkers related to the study objective. First, we created regional features of brain changes by using baseline images of the normal controls and AD individuals from the ADNI study; this step effectively reduces the high-dimensionality imaging measurements to a set of regional brain patterns that have been significantly related to the progressing brain atrophy. Aiming to detect the MCI individuals at the early stage, we investigated the longitudinal progression of such brain patterns in the BLSA participants (BLSA study is an ongoing longitudinal study of human aging. Most participants remained cognitively normal, and a few of them had shown the cognitive decline. However, no AD patients were involved. Dataset details available in the next section). Next we built HMM models by using the serial patterns of brain atrophy in BLSA participants, and tested the learned models on longitudinal scans of the MCI patients in ADNI study.

3. Data description

3.1. ADNI participants

Data used in the preparation of this article were obtained from the Alzheimers Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). 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, 5-year public-private partnership. The primary goal of ADNI has been to test whether serial magnetic

resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimers disease (AD). Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the U.S. and Canada. The initial goal of ADNI was to recruit 800 adults, ages 55 to 90, to participate in the research approximately 200 cognitively normal 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.

In our study, we used T1-weighted MRI scans at baseline from 63 CN individuals, and 54 AD patients originally used in (Davatzikos et al., 2009). They were used to adaptively extract spatial brain patterns of atrophy that maximally differentiate two groups. The AD group did not differ significantly from the healthy controls in age at baseline (77.41 ± 7.10 vs 75.23 ± 5.40), but had significantly lower MMSE scores (23.20 ± 2.10 vs 29.2 ± 0.98).

We had downloaded 239 MCI patients from the ADNI web site by April 2009, and had pre-processed their baseline and follow-up exams, which were available in our database. According to their CDR scores during the follow-up period, MCI subjects were divided into two sub-groups: converters (MCI-C), whose diagnosis was MCI at baseline and their global CDR score changed from CDR=0.5 to CDR=1, and non-converters, whose global CDR score remained stable. Aiming to predict early MCI, we had only considered 170 MCI-NC. Data from the MCI patients were followed up for an average period of approximately 1 year with a standard deviation of 6 months (range: 0.5–3 years). In this study, 54 of 170 MCI-NC with 2 or 3 years' follow-ups (2.40 ± 0.50) were selected to verify the performance validation of HMM model, which was trained on the BLSA dataset. Their baseline scans and annual follow-up visits from year 1 to year 2 or 3 were included in this study.

3.2. BLSA participants

The BLSA neuroimaging study is a prospective study investigating structural, functional, and cognitive changes associated with normal aging and cognitive impairment (Resnick et al., 2000). The present study involved longitudinal scans from 136 individuals who had been followed for more than four years with annual imaging, neuropsychological and clinical evaluations. Specifically, it included 102 BLSA participants who have remained clinically normal during the study follow-up period, 15 individuals identified as MCI within the context of prospective BLSA follow-ups, as well as 19 individuals presenting progressive cognitive decline, whose Clinical Dementia Rating (CDR) scores were 0.5, but not diagnosed as MCI yet. Diagnoses of MCI were made by consensus clinical conferences and indicated impairment in at least one cognitive domain (typically memory) without evidence of functional loss in everyday activities. MRI interpretative reports were reviewed to rule out other possible causes of dementia after consensus diagnoses were determined. MRI

serial scans of these 136 participants were analyzed and used to construct a HMM model to describe the gradual development of brain atrophy patterns, and identify individuals who were more likely to convert to MCI versus those who remained normal.

Table 1 illustrates sample characteristics from both ADNI and BLSA datasets. The two 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, and the immediate free recall score (sum of five immediate recall trials) on the California Verbal Learning Test (CVLT) (Delis et al., 1987) to assess verbal learning and immediate recall. Note that the average CVLT score of CDR 0.5 group at last visit was slightly higher than that of first visit. It does not mean the CDR 0.5 group improved clinically. This was largely due to the fact that subjects learned how to take the test. In practice, there is often an increase in CVLT scores, particularly younger ones. Indeed, the individuals from CDR 0.5 group were younger at both baseline and last visit, compared with MCI group, as illustrated in Table 1. However, the average MMSE scores decreased in CDR 0.5 group. For MCI group, both CVLT and MMSE scores decreased over time, and the average scores at last visit (MMSE/CVLT: 25.40/36.40) were much lower than those of the other two groups (27.88/52.95 for CDR 0.5 group and 28.44/53.56 for CN group). This suggests that some of these participants with CDR 0.5 were not preclinical MCI because the scores typically would not improve in MCI group.

3.3. Image acquisition

The ADNI dataset is described in www.adniinfo.org (ADNI, 2004). It includes a sagittal volumetric 3D magnetization-prepared 180 degrees radio-frequency pulses and rapid gradient-echo (MPRAGE) protocol. Standard T1-weighted images were collected 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 slightly variable, but the target values were TE 3.9 ms and TR 8.9 ms. The scans used in this paper had gone through specific correction steps such as gradwarp, B1 calibration, N3 correction and (in-house) skull-stripping (Sled et al., 1998).

The neuroimaging details about the BLSA study can be found in (Resnick et al., 2003). The BLSA protocol included an axial T1-weighted volumetric spoiled gradient recalled (SPGR) series (axial acquisition, TR = 35 ms, TE = 5 ms, flip angle = 45°, voxel dimensions of 0.94 * 0.94 * 1.5 mm slice thickness).

4. Results and analysis

We used the HMM methodology to analyze the continuous progression of brain atrophy, and evaluated this methodology on two independent datasets, from the ADNI and BLSA studies. A general outline of the spatio-temporal modeling and analysis is shown in Figure 2.

4.1. Regional patterns of brain atrophy

Recent work suggests that brain patterns of atrophy are observed in some cognitively normal individuals with healthy aging, and that the frequency of these patterns often increases in MCI individuals (Davatzikos et al., 2009, 2011). These findings suggest that gradual brain changes over long time might eventually lead to clinical conversion to MCI or AD. Thus, it

is particularly interesting to investigate the progression of brain atrophy patterns for individuals with normal aging or mild MCI. We also expect that statistical findings based on such spatio-temporal analysis would provide prognostic indicators to identify individuals with a risk of developing MCI or AD, automatically.

To calculate the sparse patterns of brain atrophy, we used the baseline MRI scans of AD/CN participants from the ADNI study to adaptively extract and select a set of regional brain features. As described in Section 2.2, a grouping relevance value (Correlation Confidence) was associated with every regional feature, with higher value corresponding to the area that was more discriminative in differentiating AD and CN. We retained the top 20 brain regions for HMM modelling. This choice results in feature vectors with a smaller size, which provides the potential to obtain more reliable HMM models with better generalization ability. For each leave-10-out training setting under the bagging procedure, we generated a regional map with the selected brain atrophy patterns, and assigned the corresponding ranking scores to them, accordingly. Figure 3 shows the most representative patterns of brain changes with AD development against normal aging, obtained by averaging regional maps from different leave-10-out experiments.

The spatial distribution of atrophy were primarily located at the medial temporal lobe structures, such as entorhinal cortex and bilateral hippocampus, which are highly associated with clinical progressing of MCI, and also are involved in early pathology with AD (Petersen et al., 1999; Tabert et al., 2006). Less extensive areas of atrophy were found elsewhere in superior temporal gyrus, frontal lobe, cingulate region and precuneus. To some degree, these findings are in concordance with most of the existing literature about the distribution of brain atrophy (Blennow et al., 2006). The selected regional patterns were then used for extracting features for the MCI group of the ADNI study and for all longitudinal MRI sequences from the BLSA study.

4.2. Left-right HMM modeling

The number of states is a very important parameter of an HMM model. From a clinical viewpoint, because brain atrophy due to aging or disease evolves gradually, a higher number of states would be preferable in order to map more subtle clinical progression with brain changes. However, due to the limited size of training samples, it would be practically impossible to optimize the parameters of the model reliably if the number of states is large. In this study, we assume the five-state HMM is a reasonable choice considering both computational modeling with limited samples and clinical problem of early MCI detection. The possible clinical transitions before severe MCI would be: keeping normal without cognitive decline, normal with cognitive decline, keeping declining to early MCI, and a flexible stage in between. A 5-state left right HMM was therefore adopted in the present application (Figure 4). The HMM indices appear in decreasing order with progressive disease stages. In other word, state index 5 corresponds to the normal brain, while low index such as state 1 represents more severe clinical stages. Specifically, observation density function b_j is represented by a single Gaussian distribution for each state $s_j, j = \{1,2,3,4,5\}$ in this study.

In our experiments, we built HMM models on the regional feature sequences from the BLSA study, and examined them on MCI individuals of the ADNI study. As has already been mentioned, to address the issue of robustness, we applied the bagging procedure to repeatedly build HMM models.

4.3. State transition path

Given the specific structure of left-right HMM, it is reasonable to assume that the declining state index values correspond to progressive stages of brain atrophy. We are interest in how HMM states transited with the different levels of pathology severity. We therefore analyzed and compared the state transition paths from three subgroups in the BLSA study, i.e., CN, MCI and CDR 0.5 groups, respectively. Figure 5 shows the state transitions of all subjects in the three groups.

In the case of CN group, even though most normal individuals presented continuous declining states, they still remained at the relatively high state index values, from 5 to 2 (top of Figure 5). However, for the subgroups with relatively high risk of developing more severe impairments, such as the MCI group, we found that the majority of individuals converted to states with the lowest index 1 during the observation period, and they even had relatively low HMM states at the beginning (bottom left of Figure 5). The CDR 0.5 group had more heterogeneous and complex progression, as shown on bottom right of Figure 5. Some individuals of this group remained in relatively healthy states comparable to CN individuals, while others transitioned to state 1 analogous to MCI individuals.

To further investigate the performance of HMM models, we also examined serial scans of MCI individuals from the ADNI study by applying the HMM models constructed by BLSA participants. As shown in Figure 6, we found a high prevalence of state 1, largely from the beginning of the observation period. This is not unexpected because the ADNI study focuses on MCI to AD progression, and involves MCI individuals with relatively pronounced cognitive decline at baseline. In contrast, the BLSA is a prospective study of normal aging, and its MCI individuals tend to have relatively mild cognitive decline as detected at very early stages. This suggests that our HMM model is capable of capturing the pathologic development of MCI, and presumably AD, at an early stage, as illustrated by results from the BLSA participants.

Overall, our experimental results support that HMM models derived from adaptive regional parcellation of brain can quantify dynamic patterns of brain atrophy, which seem to correspond with cognitive decline.

4.4. Relationship between HMM states and cognitive performance

Although the decoded HMM states are anticipated to be highly associated with clinical stages of brain changes with atrophy as well as cognitive decline, we did not fully understand how they match each other. To determine the relationship between the hidden states of HMMs and cognitive performance, we examined the respective frequency of state transitions from the three subgroups of the BLSA dataset mentioned in Section 3.2.

As the MCI and CDR 0.5 groups are presenting some extent of cognitive impairment clinically from the healthy beginning of recruitment, higher transition numbers are anticipated for them than CN individuals. The left plot of Figure 7 provides a normalized summary of the state-transition number from three groups. However, only half of the participants from both cognitive impairment groups (MCI/CDR0.5 groups: 53.33% / 52.63%) showed state transitions, while the CN group had a similar proportion (46.08%) of transitions. We further examined those stable individuals without any transitions from the three groups, and observed that they followed the diverse initial HMM states. As shown in the right plot of Figure 7, a larger proportion (75%) of the stable MCIs remained at state 1, and none of them had relatively high state values, e.g. 4 and 5. For the CDR 0.5 group, a relatively balanced distribution was measured in terms of state 1 and 'healthier' states 3 to 5. In contrast to the two cognitive impairment groups, the CN group showed a much larger proportion (40.4%) of individuals remaining in state 5.

Additionally, we performed statistical analysis of the state transition count for MCI group from the ADNI study. As shown in Figure 8, although the large majority of MCI individuals did not show any state transition from baseline, those who remained stable showed a dominated distribution on state 1 (Figure 6). As we discussed in Section 4.3, the 5-state HMM models built using BLSA participants can characterize the progression with MCI in the early phrase. Compared with BLSA participants, MCI patients from the ADNI study had been diagnosed to be the worse clinical stages at baseline. Therefore, it is reasonable that most ADNI participants showed state index 1 without state transitions.

4.5. Predictor of future progression to MCI

Due to the increasingly high incidence of MCI and AD after the age of 65, it is desirable to provide a predictor that identifies early stages of the disease. We therefore investigated the HMM state indexes at each individual's last visit for the three subgroups showing different severities of cognitive impairment. The normalized count of last state index values for the respective subgroups are shown in Figure 9.

Notably, the MCI group from both BLSA and ADNI studies showed a distribution with a single peak on state 1. For the BLSA CDR 0.5 group, we observed a high prevalence of state 1 and a decrease in the number of subjects with increasing state numbers. For the CN group, we observed significant variability in last state values that covered the full range of HMM states. The majority of CN elderly demonstrated states ranging from 2 to 5, although a few individuals were in state 1. The convergence of the MCI and CDR 0.5 groups to state 1, which was not observed for the CN group, indicates that state 1 can be considered as a potential predictor of MCI, namely MCI-like state.

Since state 1 has emerged from our experiments as a likely predictor of MCI, we were particularly interested in 17 CN individuals in the BLSA dataset, i.e. subjects who converted to state 1 even though they did not meet diagnostic criteria for cognitive impairment. In order to investigate their respective progression of brain changes, we first analyzed the state paths for these subjects (Figure 10).

Interestingly, all seventeen individuals had a relatively low state index at the beginning of the study. Some of them started from state 3 or 2 and then converted to state 1. Others started from state 1 and remained in this state. This finding implies that these subjects had relatively pronounced brain atrophy at the first visit.

We searched for further support for the validity of our predictor by comparing the clinical variables of these 17 subjects (Group 1) to those of the healthiest subjects (Group 2). Group 2 consisted of CN subjects who remained in state 5 during the course of study. Importantly, the two groups were comparable with respect to baseline age. We observed that Group 1 subjects showed relatively worse cognitive performance with lower CVLT scores (baseline/last visit: 49.19/47.94), compared to Group 2 subjects (baseline/last visit: 51.47/54.74). Furthermore, the difference between first and last visit scores was negative for Group 2, while it was positive for Group 1. A negative difference is generally the indicator of cognitive decline, while a positive difference is likely due to the 'learning effect' in neuropsychological testing and is generally observed in healthy individuals with repeated test exposure. Interestingly, the commonly used MMSE scores were comparable for these two groups. However, the standard deviations of average MMSE scores for Group 1 subjects at both first and last visits (std: 1.88/1.54) were greater than those of Group 2 subjects (std of baseline/last visit: 0.85/0.62). This is in agreement with the well documented finding that the individuals at risk to progress to MCI are quite heterogeneous (Petersen et al., 1999). Additional data is presented in Table 2.

This analysis further bolstered our confidence in the estimative power of our predictor.

4.6. Relationship between HMM states and longitudinal SPARE-AD scores

SPARE-AD (Spatial Pattern of Abnormalities for Recognition of Early AD) score has been validated as good imaging biomarkers of early AD, and indicated the presence of AD if positive, and otherwise if negative (Davatzikos et al., 2009; Fan et al., 2007). We are therefore interested in the relationship between HMM states and SPARE-AD scores. Both SPARE-AD study and the proposed method used the same image preprocessing procedures and had been evaluated by the same datasets. Differently, SPARE-AD score was calculated by a pattern classification method described in (Fan et al., 2007). In particular, a specific classifier was derived by utilizing MRI T1-weighted scans of 66 cognitively individuals (mean age ± S.D., 75.18 ± 5.39), and 56 AD patients (77.40 ± 7.02) from the ADNI cohort, and then was applied to the longitudinal scans of BLSA participants (Davatzikos et al., 2009).

As each individual has both HMM state path and longitudinal SPARE-AD series, we examined the linear correlation for several subgroups in the BLSA study. As showed in Table 3, all groups were highly correlated, 0.58 of correlation in MCI group, 0.68 of correlation in CDR 0.5 group, and 0.51 of correlation in Group 1. Totally it is 0.63. Note that HMM models were constructed by BLSA participants, and measured the progression of brain atrophy before severe MCI. It explains why Group 1 has relatively lower correlation.

4.6.1. Prediction of MCI conversion—SPARE-AD scores at each time point were considered as an index sequence to represent the dynamics of brain changes over time.

Apparently HMM state path also can explain the dynamics by involving time series analysis. It is interesting to compare their predictive capabilities on abnormal brain changes, particularly MCI or AD conversion in CN subjects.

For individual SPARE-AD scores at multi-time points, we fitted a curve to it, and then automatically estimated the conversion point of AD on the curve. Observation of graphs of year vs. SPARE-AD score revealed that the shape closely resembled a sigmoidal function, which has a 'S' pattern with a rapid changing point. This is line with that support vector machine attempt to discriminate one class training samples from the other. The sigmoidal function used to fit the SPARE-AD curves is defined as follows:

$$f(x,\beta)=\beta_1\left(\frac{1}{1+e^{\beta_3}(x+\beta_4)}+\beta_2\right) \quad (8)$$

Visual inspection of these curves shows that strong conversion begins to occur near the point where the slope is half of the maximum slope occurring at the inflection point. In this study we consider the point with this slope value to be the point of conversion. If that slope is zero, we took the point of conversion to be the last year of data. In addition, if the estimated conversion point is outside the range of available data, the we considered the conversion occurred at the first or last year of the follow-up periods, depending on the average score of all data point individually. If the average score is greater than zero, the conversion point is considered as the first year, and if the negative average score is less than zero, we assume the subject has not converted yet, but consider the last year of available data in a conservative way as conversion point.

We also calculated the conversion point of HMM state path. Given state 1 indicated as MCI status, it is reasonable to take state 2 as the beginning of conversion from normal control to MCI. If a subject remained as state 1 since the first visit, we considered its conversion time point as the first year of data. If the states transited from index 3 to 1 directly, the point of conversion is considered to be the half year before the time point of being state 1. In keeping with the conversion estimation of SPARE-AD scores, the conversion point of HMM state path is the last year of available data if the subject did not transit to state 2 during the follow-up period. Figure 11 shows an example of the estimated curves of SPARE-AD scores and HMM state path for two typical subjects, along with the estimated points of conversion.

Given the definition and experimental design of SPARE-AD scores and HMM states, SPARE-AD score sequence captured the conversion of AD while HMM state path predicted MCI from CN subjects. For the same individual, the conversion points decided by HMM state path should be early than that calculated by SPARE-AD scores. To address this point, we built the median SPARE-AD score curve and the mean HMM state path based on the calculated points of AD/MCI conversion from all abnormal subjects above, respectively. The difference was described in Figure 12, where the average MCI conversion obtained by HMM showed 3 years earlier than SPARE-AD index.

4.6.2. Brain region selection and HMM model generalizability—As we mentioned before in Result section, 20 top-ranked brain regions were selected from all adaptively

generated brain clusters. The proposed method was 4D spatio-temporal analysis approach rather than single baseline image based classification methods. In order to optimize the number of feature input to HMM modeling, we have outlined the correlation coefficients of time-varying HMM state path with SPARE-AD score series with increasing number of brain regions from 5 to 40 (Figure 13). In order to balance the model complexity and generalization ability, we selected 20 features as it achieved the highest correlation score.

4.7 Analysis of mean and variance images from five HMM-state groups

Based on the HMM state path, an individual brain image can be viewed as an observation of a HMM state, and thus totally we had five groups of brain images corresponding to the five HMM states. We generated the mean and variance images for each HMM-state group (Figure 14 for grey matter and Figure 15 for white matter). As stated above, lower HMM state indices indicated more severe cognitive impairment stages. We found that the mean images through 5 stages demonstrated a gradual progression of brain tissue loss in the Hippocampus, Caudate Nucleus and Thalamus (regions labeled as white). Similarly, tissue shrinkage was also observed by the decreasing red areas through the frontal and temporal lobes from HMM-state group 5 to group 1.

Difference exists within each group of brain images, due to the fact that the transition from a healthy state to a disease state progresses gradually, and brain images from each HMM-state group involves the progression of cognitive impairment from one level to the next. Also, for brain regions showing tissue loss across the 5 HMM states, moderate variance (shown in green color in Figure 14 & 15) was identifiable within each HMM-state group. This indicates the changes seen within group were homogeneous. Referring back to the mean images, the changes seen between groups supported that the five HMM-state groups differed from each other and corresponded to progressive stages of cognitive impairment.

5. Discussion and conclusion

A variety of neuroimaging studies have examined longitudinal brain changes associated with growth, aging, or disease progression via group analysis (Sullivan et al., 2002; Resnick et al., 2003; Driscoll et al., 2009; Sluimer et al., 2010). In this study, we introduced a novel method, regional feature-based HMM, to investigate longitudinal progression of brain changes on an individual basis, aiming to develop early markers of progression of structural brain changes. We applied the proposed method to the analysis of spatio-temporal progression of brain atrophy patterns in individuals with normal cognition and relatively mild MCI. We investigated the potential of the proposed model to predict conversion from normal cognition to MCI or AD. A contribution of this study is that the proposed method is developed and evaluated by using images from two independent datasets with complementary strengths. Highly discriminative regions were obtained from the CN and AD participants of the ADNI study, who were exactly the same as in our previous classification work (Fan et al., 2008 and Davatzikos et al., 2009). In particular, the latter work generated SPARE-AD scores, which have been used to evaluate the performance of HMM model. It would be fair to compare two models when using the same training data. Feature vector sequences were extracted from the same regions of images from a completely independent

longitudinal dataset (BLSA), and were subsequently used for training the HMM model. The learned HMM models were applied to the MCI patients from the ADNI dataset. To follow the BLSA annual follow-up period, 54 of 234 ADNI MCI individuals with longer annual follow-up scans (2 or 3 years) had been used to validate the HMM model. Our hypothesis is that there is difference between early MCI and late MCI. It is thus desirable that the detected imaging biomarkers/indicators specifically characterize early MCI only, rather than "already" MCI together. The validation experiments on already MCI data (ADNI) showed their longitudinal progress remained in the lowest HMM state 1 while early MCI (BLSA) showed progression from relatively higher states to the lowest state. These results supported that our HMM indicators could capture the subtle characters of early MCI. This evaluation scheme across different studies is a good indicator of the clinical applicability and generalization ability of the method (Kloppel et al., 2008; Davatzikos et al., 2009). However, most previous studies on high-dimensional pattern classification (Golland et al., 2002; Fan et al., 2007; Vemuri et al., 2009; Misra et al., 2009; Plant et al., 2009), pattern regression (Ashburner, 2007; Duchesne et al., 2009; Franke et al., 2010; Wang et al., 2010a), or longitudinal analysis for a variety of neurodegenerative diseases (Sullivan et al., 2002; Resnick et al., 2003; Wang et al., 2010b) were restricted to a single dataset, which was insufficient to demonstrate the generalization ability of these approaches. Herein, this spatio-temporal scheme was applied and evaluated on the ADNI and BLSA imaging datasets. In view of the consistent experimental results, our method is expected to be a well generalized tool, with the potential to contribute to other imaging studies that aim to capture any kind of spatio-temporal changes on an individual basis.

The AD-like spatial patterns of brain atrophy have been validated as features by our previous studies (Fan et al., 2007; Davatzikos et al., 2009), as well as confirmed by others (Kloppel et al., 2008; Vemuri et al., 2009). However, due to limited number of training samples, only a few top-ranked brain regions were used in this study for the HMM modeling, in order to obtain a stable spatio-temporal model. Although few features have been used, these regions are determined from the data, and primarily located in the medial temporal lobe, which fits with the known patterns of AD development from both clinical and neuroimaging research studies (Wenk, 2003; Bredesen et al., 2006; Davatzikos et al., 2009). Considering the importance of extraction of discriminative and consistent patterns with brain changes, further work will focus on effective feature extraction for longitudinal data. Also a 4D segmentation algorithm is necessary to provide better temporal consistency and to extract more robust features. Recent studies on 4D segmentation have begun to emerge in the literature (Xue et al., 2006; Shi et al., 2010), and provide potentially powerful pre-processing tools for longitudinal brain images, which can lead to more stable HMM models.

Notably, the training samples for HMM modeling were older adults, even though some of them were identified as having very mild impairment based on CDR scores of 0.5. Therefore, the HMM model might represent progressive brain changes from normal aging to the very early stages of cognitive impairment. Based on the state-transition paths indicating the continuous progression of spatial atrophy patterns, we examined the frequency of state transitions for the subgroups with different levels of cognitive impairment. Our experimental results indicate that most relatively mild MCI individuals in the BLSA study showed rapid state transitions, while those who were stable were mostly on state 1. On the other hand, the

vast majority of MCI individuals from the ADNI study were in state 1 at the first time point. We believe that this is due to the differences in subject recruitment in these studies. The ADNI MCI individuals, due to the focus of the study, had more severe impairment than the BLSA MCI individuals, who were initially recruited as CN and developed MCI during the course of this prospective longitudinal study (some were called retrospectively MCI). As expected, the vast majority of CNs had progressive state transition over time, but still remained in the relatively healthy stages. Our analysis on state transition paths showed that a decrease in HMM state value corresponds to the decline of cognitive performance. Furthermore, analysis of the relationship between HMM states and SPARE-AD score index also demonstrated that HMM has strong potential to predict early conversion of MCI.

Although it is not possible to know if HMM states correspond to underlying stages of pathology, state 1 is consistently observed for the individuals with CDR score 0.5 or above. More interestingly, it has been also observed for a small proportion of individuals who are cognitively normal. These individuals performed significantly worse on neuropsychological tests and showed greater decline over time. This finding is consistent with previous studies that reported that some CN individuals remain stable for a long period before converting to MCI or AD, while others show a continuous decline in memory and cognition (Wenk, 2003; Bredesen et al., 2006; Blennow et al., 2006). Further follow-up of the entire BLSA cohort will allow us to verify the predictive ability of state 1.

While the regional feature based HMM model is a promising approach to evaluate the spatio-temporal progression of brain changes individually, several important improvements could be proposed. First, automatically learning the number of states in HMM, as the literature suggested (Brand, 1999), would be possible when larger datasets are available for training in the future. That would provide a more adaptive model that fits the data better. Secondly, it would be interesting to investigate multi-modality based spatio-temporal analysis. In fact, the functional imaging patterns, CSF biomarkers and clinical assessments can provide complementary information, which leads to better understanding of brain changes associated with various studies. For example, test scores of cognitive performance (Petersen et al., 2001; Luis et al., 2003), computerized scores of MRI scans (Davatzikos et al., 2009), or PET imaging patterns (Fan et al., 2008) could be combined. However, a challenging problem is how to integrate the diverse patterns, especially without the availability of larger datasets. To deal with this problem, we plan to combine multi-modality patterns at the level of the model rather than the voxel level or decision level. For example, an extension of HMM with more complex structure linking to multiple pattern sequences will be emphasized in future work for higher predictive power.

Acknowledgments

The authors thank Evi Parmpi for her help with handing MRI datasets.

This study was financially supported by the NIH grant R01AG14971, by N01-AG-3-2124, and by the NIA intramural program.

Part data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical

Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). 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 University of California, San Diego. ADNI data are disseminated by the Laboratory for NeuroImaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation.

References

ADNI. 2004. http://www.loni.ucla.edu/ADNI/.

Ashburner J. A fast diffeomorphic image registration algorithm. NeuroImage. 2007; 38(1):95–113. [PubMed: 17761438]

Ashburner J, Friston K. Voxel-based morphometry: the methods. NeuroImage. 2000; 11:805–821. [PubMed: 10860804]

Blennow K, de Leon M, Zetterberg H. Alzheimer's disease. Lancet. 2006; 368(9533):387–403. [PubMed: 16876668]

Brand M. Structure learning in conditional probability models via an entropic prior and parameter extinction. Neural Computation. 1999; 11:1155–1182.

Bredesen D, Rao R, Mehlen P. Cell death in the nervous system. Nature. 2006; 443(7113):796–802. [PubMed: 17051206]

Breiman L. Bagging predictors. Machine Learning. 1996a; 26(2):123–140.

Breiman, L. Statistics Department. University of California; 1996b. Out-of-bag Estimation. Technical Report.

Chtelat G, Landeau B, Eustache F, Mzenge F, Viader F, de la Sayette V, Desgranges B, Baron JC. Using voxel-based morphometry to map the structural changes associated with rapid conversion in MCI: A longitudinal MRI study. NeuroImage. 2005; 27(4):934–946. [PubMed: 15979341]

Costin H, Geman O. Parkinson's Disease Prediction Based on Multistate Markov Models. International Journal of Computers Communications & Control. 2013; 8(4):525–537.

Davatzikos C, Bhatt P, Shaw LM, Batmanghelich KN, Trojanowski JQ. Prediction of MCI to AD conversion, via MRI, CSF biomarkers, and pattern classification. Neurobiology of Aging. 2011; 32(12):2322, e19–e27. [PubMed: 20594615]

Davatzikos C, Genc A, Xu D, Resnick S. Voxel-based morphometry using the RAVENS maps: Methods and validation using simulated longitudinal atrophy. NeuroImage. 2001; 14:1361–1369. [PubMed: 11707092]

Davatzikos C, Xu F, An Y, Fan Y, Resnick SM. Longitudinal progression of Alzheimer's-like patterns of atrophy in normal older adults: the SPARE-AD index. Brain. 2009; 132(8):2026–2035. [PubMed: 19416949]

Davis RIA, Lovell BC. Comparing and evaluating HMM ensemble training algorithms using train and test and condition number criteria. Formal Pattern Analysis & Applications. 2004; 6:327–335.

Delis DC, Kramer JH, Kaplan E, Ober B. California verbal learning test - research edition. The Psychological Corporation. 1987

Driscoll I, Davatzikos C, An Y, Wu X, Shen D, Kraut M, Resnick SM. Longitudinal pattern of regional brain volume change differentiates normal aging from MCI. Neurology. 2009; 72(22): 1906–1913. [PubMed: 19487648]

Duchesne S, Caroli A, Geroldi C, Collins DL, Frisoni GB. Relating one-year cognitive change in mild cognitive impairment to baseline MRI features. NeuroImage. 2009; 47(4):1363–1370. [PubMed: 19371783]

Fan Y, Resnick SM, Wu X, Davatzikos C. Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study. NeuroImage. 2008; 41(2): 277–285. [PubMed: 18400519]

Fan Y, Shen D, Gur RC, Gur RE, Davatzikos C. Compare: Classification of morphological patterns using adaptive regional elements. IEEE Transactions on Medical Imaging. 2007; 26:93–105. [PubMed: 17243588]

Folstein MF, Folstein SE, McHugh PR. Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research. 1975; 12(3):189–198. [PubMed: 1202204]

Fonteijn, Hubert M.; Clarkson, Matthew J.; Modat, Marc; Barnes, Josephine; Lehmann, Manja; Ourselin, Sebastien; Fox, Nick C.; Alexander, Daniel C. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. NeuroImage. 2012; 60(3):1880–1889. [PubMed: 22281676]

Formisano E, De Martino F, Valente G. Multivariate analysis of fMRI time series: classification and regression of brain responses using machine learning. Magnetic Resonance Imaging. 2008; 26(7): 921–934. [PubMed: 18508219]

Forney GD. The viterbi algorithm. Proceedings of the IEEE. 1973; 61(3):268–278.

Franke K, Ziegler G, Klppel S, Gaser C. Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: Exploring the influence of various parameters. Neuroimage. 2010; 50(3):883–892. [PubMed: 20070949]

Golland, P.; Fischl, B.; Spiridon, M.; Kanwisher, N.; Buckner, RL.; Shenton, ME.; Kikinis, R.; Dale, A.; Grimson, WEL. Discriminative analysis for image-based studies; International Conference on MICCAI. Vol. 2488 of Lecture Notes in Computer Science; 2002. p. 508-515.

Good C, Johnsrude I, Ashburner J, Henson R, Friston K, Frackowiak R. A voxel-based morphometric study of ageing in 465 normal adult human brains. NeuroImage. 2001; 14:21–36. [PubMed: 11525331]

Good CD, Scahill RI, Fox NC, Ashburner J, Friston KJ, Chan D, Crum WR, Rossor MN, Frackowiak RSJ. Automatic differentiation of anatomical patterns in the human brain: Validation with studies of degenerative dementias. NeuroImage. 2002; 17(1):29–46. [PubMed: 12482066]

Kabani N, MacDonald D, Holmes CJ, Evans A. A 3D atlas of the human brain. Neuroimage. 1998; 7(4):S717.

Kloppel S, Stonnington CM, Chu C, Draganski B, Scahill RI, Rohrer JD, Fox NC, Jack CR Jr, J A, Frackowiak RSJ. Automatic classification of MRI scans in Alzheimer's disease. Brain. 2008; 131(3):681–689. [PubMed: 18202106]

L. E. Baum T, Petrie GS, Weiss N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. Ann. Math. Statist. 1970; 41(1):164–171.

Luis C, Loewenstein D, Acevedo A, Barker W, Duara R. Mild cognitive impairment: directions for future research. Neurology. 2003; 61(4):438–444. [PubMed: 12939414]

Misra C, Fan Y, Davatzikos C. Baseline and longitudinal patterns of brain atrophy in MCI patients, and their use in prediction of short-term conversion to AD: Results from ADNI. NeuroImage. 2009; 44(4):1415–1422. [PubMed: 19027862]

Moeslund TB, Hilton A, KrÄuger V. A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding. 2006; 104(2):90–126.

Newton, Paul K., et al. Spreaders and Sponges Define Metastasis in Lung Cancer: A Markov Chain Monte Carlo Mathematical Model. Cancer research. 2013; 73(9):2760–2769. [PubMed: 23447576]

Petersen R, Smith G, Waring S, Ivnik R, Tangalos E, Kokmen E. Mild cognitive impairment: Clinical characterization and outcome. Neurology. 1999; 56(3):303–308.

Petersen R, Stevens J, Ganguli M, Tangalos E, Cummings J, DeKosky S. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). report of the quality standards subcommittee of the american academy of neurology. Neurology. 2001; 56(9): 1133–1142. [PubMed: 11342677]

Pham D, Prince J. Adaptive fuzzy segmentation of magnetic resonance images. IEEE Trans. Medical Image. 1999; 18(9):737–752.

Plant C, Teipel S, Oswald A, Bhm C, Meindl T, Mourao-Miranda J, Bokde AHH, Ewers M. Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease. Neuroimage. 2009; 50:162–174. [PubMed: 19961938]

Rabiner LR. A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE. 1989:257–286.

Resnick SM, Goldsza AF, Davatzikos C, Golski S, Kraut MA, Metter EJ, Bryan RN, Zonderman AB. One-year age changes in MRI brain volumes in older adults. Cerebral Cortex. 2000; 10(5):464–472. [PubMed: 10847596]

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 Apr; 23(8): 3295–3301. [PubMed: 12716936]

Shen D, Davatzikos C. Hammer: Hierarchical attribute matching mechanism for elastic registration. IEEE Transaction on Medical Imaging. 2002; 21:1421–1439.

Shi F, Fan Y, Tang S, Gilmore JH, Lin W, Shen D. Neonatal brain image segmentation in longitudinal MRI studies. Neuroimage. 2010 Jan; 49(1):391–400. [PubMed: 19660558]

Sled J, Zijdenbos A, Evans A. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Transactions on Medical Imaging. 1998; 17(1):87–97. [PubMed: 9617910]

Sluimer JD, Bouwman FH, Vrenken H, Blankenstein MA, Barkhof F, van der Flier WM, Scheltens P. Whole-brain atrophy rate and CSF biomarker levels in MCI and AD: A longitudinal study. Neurobiology of Aging. 2010; 31(5):758–764. [PubMed: 18692273]

Smith SM. Fast robust automated brain extraction. Human Brain Mapping. 2002; 17(3):143–155. [PubMed: 12391568]

Sullivan E, Pfefferbaum A, Adalsteinsson E, Swan G, Carmelli D. Differential rates of regional brain change in callosal and ventricular size: A 4-year longitudinal MRI study of elderly men. Cereb Cortex. 2002; 12:438–445. [PubMed: 11884358]

Tabert M, Manly J, Liu X, et al. Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. Gen. Psychiatry. 2006; 63(8):916–924.

Vemuri P, Gunter JL, Senjem ML, Whitwell JL, Kantarci K, Knopman DS, Boeve BF, Petersen RC Jr, CR J. MRI and CSF biomarkers in normal MCI, AD subjects: Diagnostic discrimination and cognitive correlations. Neurology. 2009; 73:287–293. [PubMed: 19636048]

Wang Y, Fan Y, Bhatt P, Davatzikos C. High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables. NeuroImage. 2010a; 50(4):1519–1535. [PubMed: 20056158]

Wang Y, Resnick SM, Davatzikos C. Spatio-temporal analysis of brain MRI images using Hidden markov models. Medical Image Computing and Computer-Assisted Intervention. 2010b; Vol. 6362:160–168. [PubMed: 20879311]

Welch LR. Hidden markov models and the BaumWelch algorithm. IEEE Information Theory Society Newsletter. 2003

Wenk G. Neuropathologic changes in Alzheimer's disease: potential targets for treatment. Journal of Clinical Psychiatry. 2003; 64:7–10. [PubMed: 12934968]

Xue Z, Shen D, Davatzikos C. Classic: Consistent longitudinal alignment and segmentation for serial image computing. NeuroImage. 2006; 30:388–399. [PubMed: 16275137]

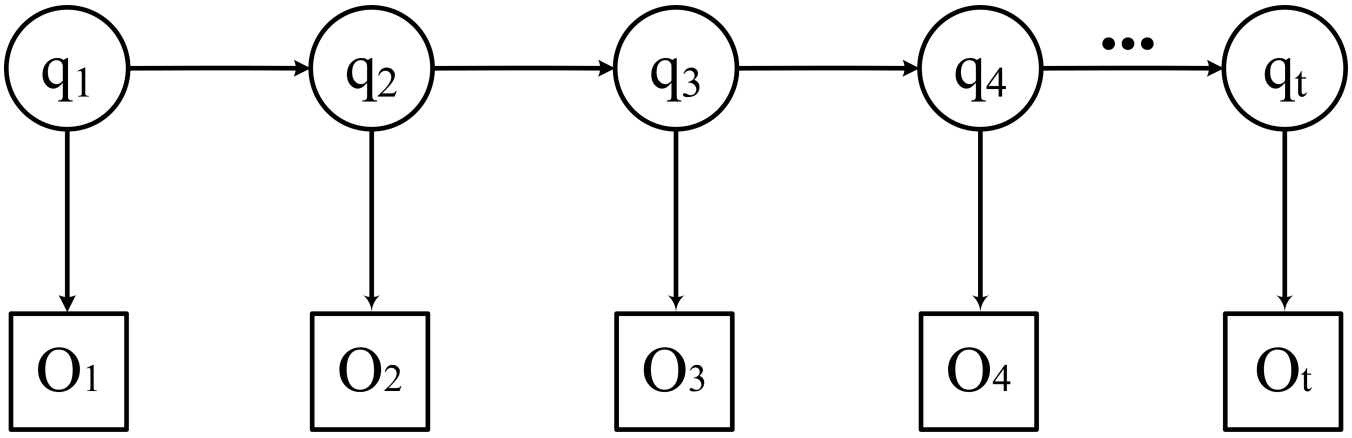


Figure 1.

HMM structure. The random variable q_t is the hidden state at time t (with the model from the above diagram, $q_t \in S = \{s_1, s_{2,\ldots}, s_n\}$). The random variable O_t is the observation vector at time t. The arrows in the diagram denote conditional dependencies.

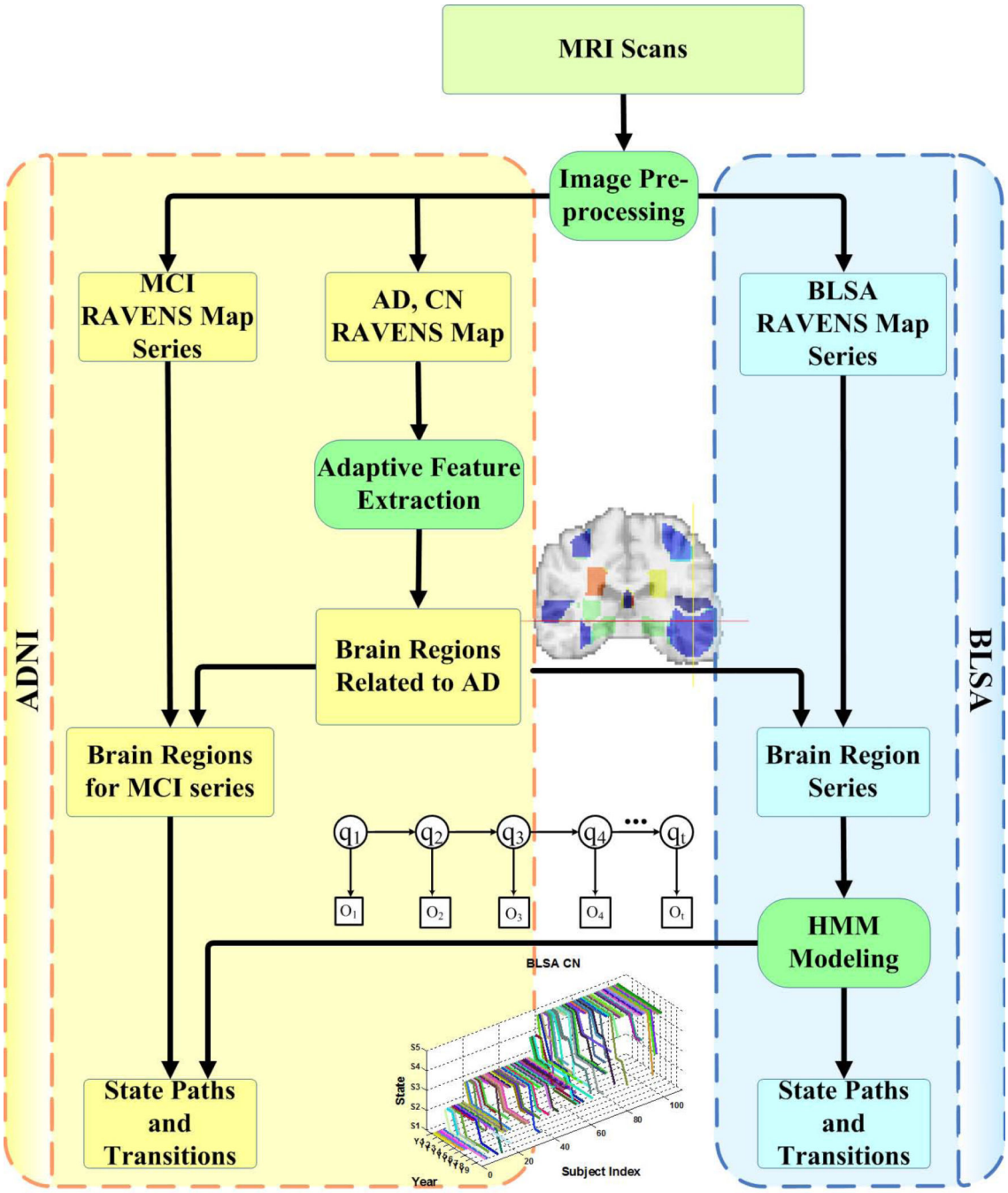


Figure 2.

A methodological framework for spatio-temporal analysis of brain patterns, validated by two independent datasets, ADNI and BLSA.

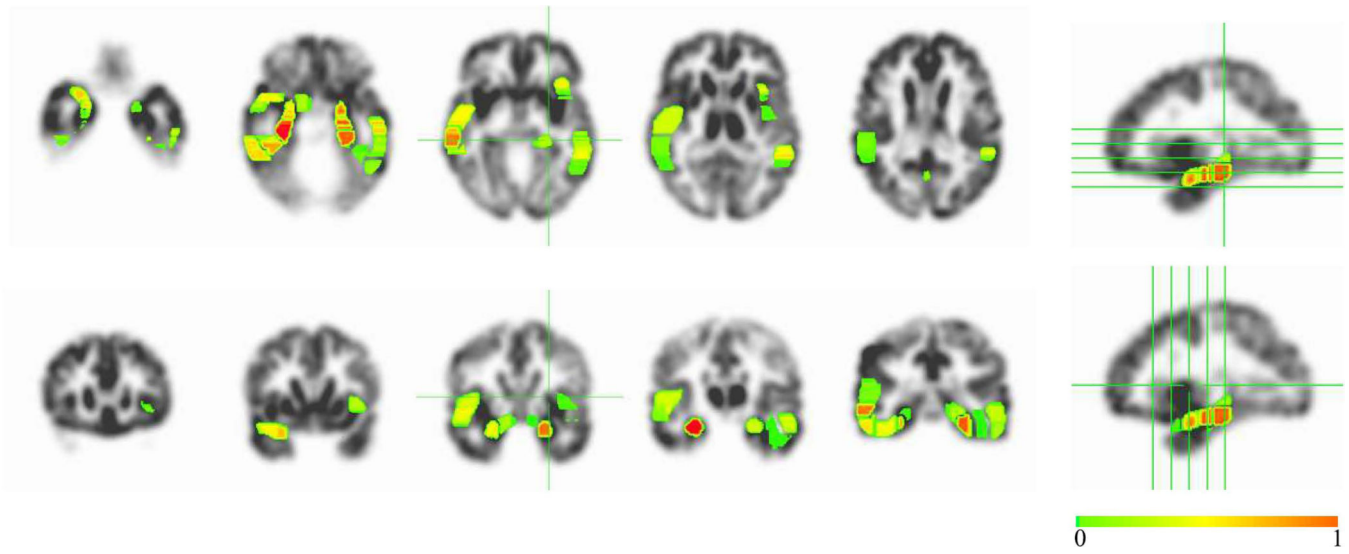
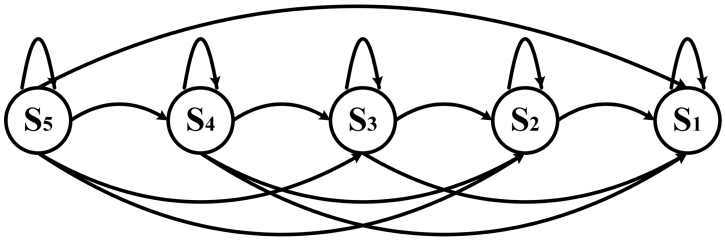


Figure 3.

Examples of top-ranked regions signifying differences between normal aging and AD in the ADNI study. By averaging all regional maps from different subgroups under the bagging procedure, the values of Correlation Confidence for the top 20 regions were normalized to the unit interval. The map was color-coded so that stronger relevance was coded by 'warmer' color (with the highest set for red).



$$\begin{array}{c} \\ S_5 \\ S_4 \\ S_3 \\ S_2 \\ S_1 \end{array} \begin{array}{c} \begin{array}{ccccc} S_5 & S_4 & S_3 & S_2 & S_1 \end{array} \\ \begin{bmatrix} a_{55} & a_{54} & a_{53} & a_{52} & a_{51} \\ 0 & a_{44} & a_{43} & a_{42} & a_{41} \\ 0 & 0 & a_{33} & a_{32} & a_{31} \\ 0 & 0 & 0 & a_{22} & a_{21} \\ 0 & 0 & 0 & 0 & a_{11} \end{bmatrix} \end{array}$$

Figure 4.

State transition diagram (left) and matrix (right) of the left-to-right HMM with 5 states. The hidden state $q_t \in \{s_1, s_2, s_3, s_4, s_5\}$. The arrows in the diagram denote conditional dependencies.

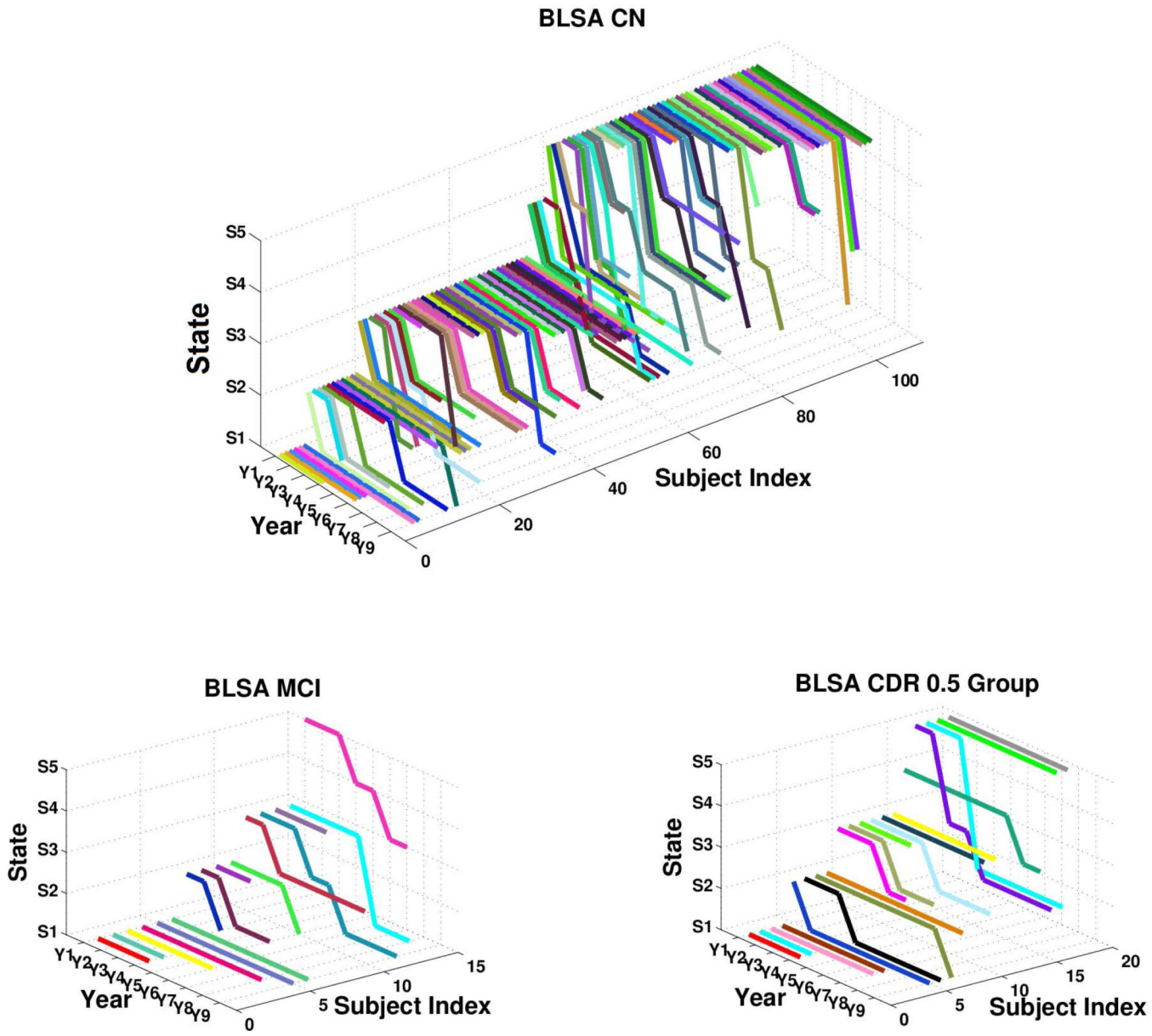


Figure 5.
State paths of the CN, MCI and CDR 0.5 groups from the BLSA study, respectively.

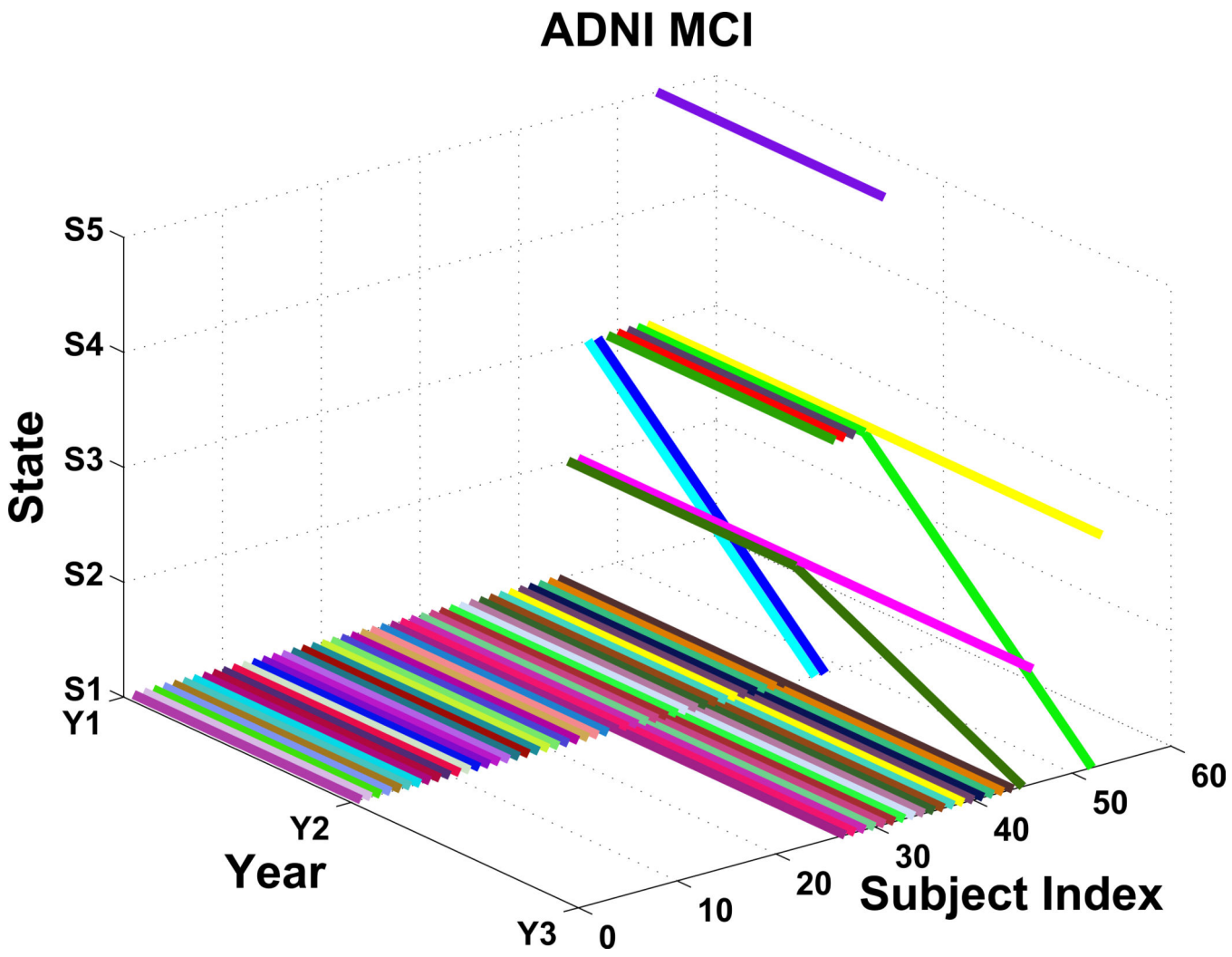


Figure 6.
State paths for the MCI patients of the ADNI study.

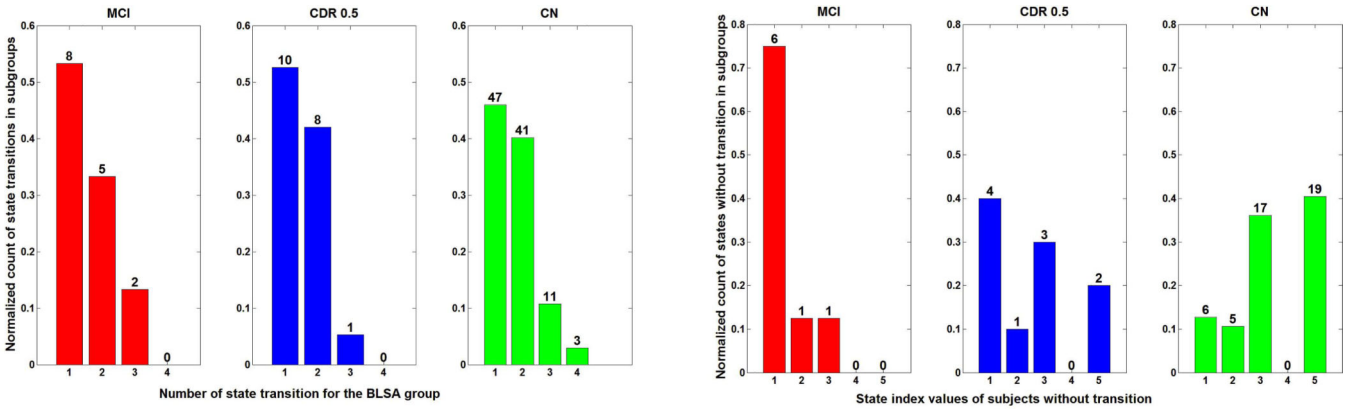


Figure 7.

Left: Normalized count of state transition of the MCI, CDR 0.5 and CN groups from the BLSA study; Right: Percentage of stable individuals with their state index values from the corresponding groups. The raw number is on top of each bin.

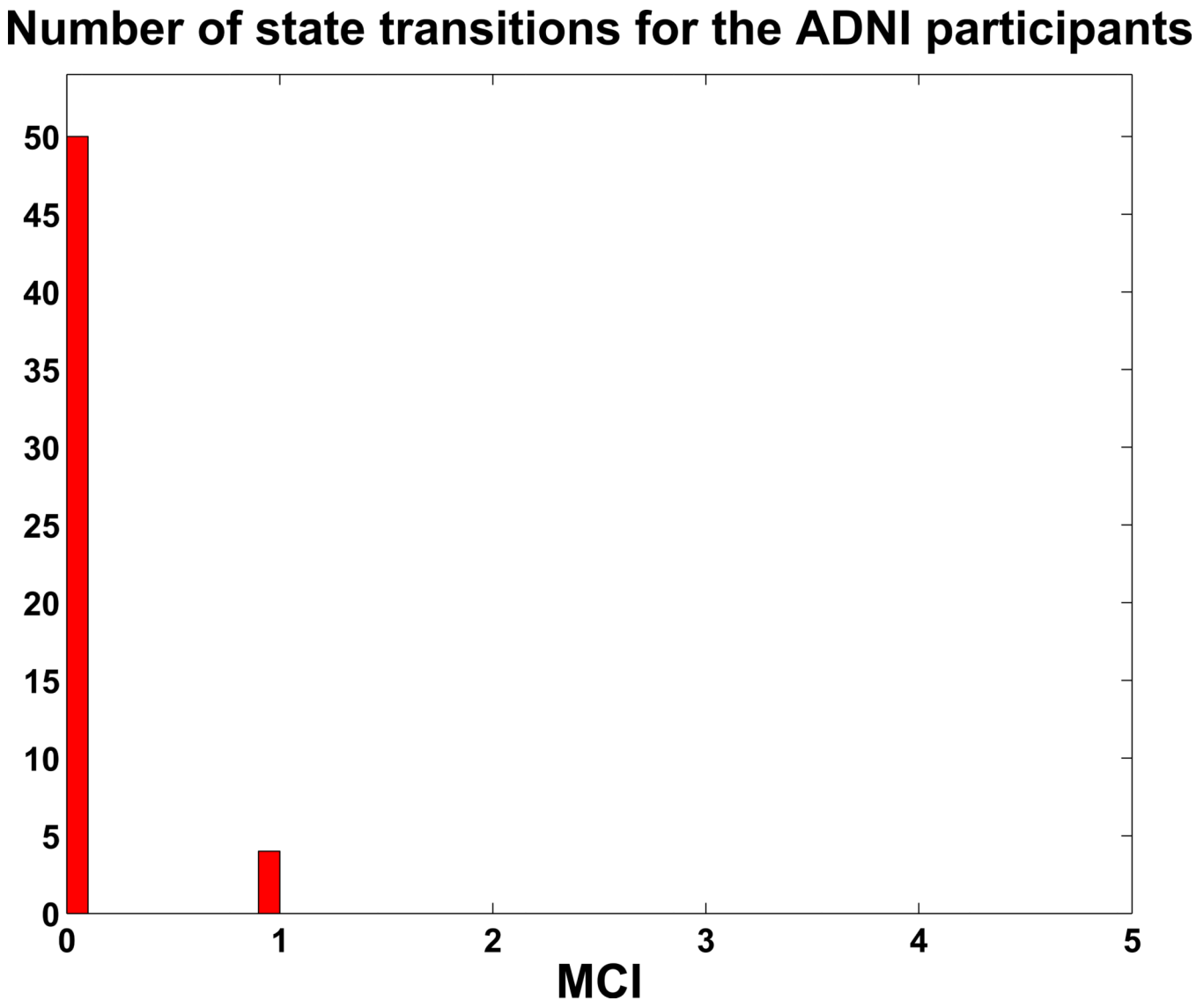


Figure 8.
The number of state transitions of MCI individuals from the ADNI study.

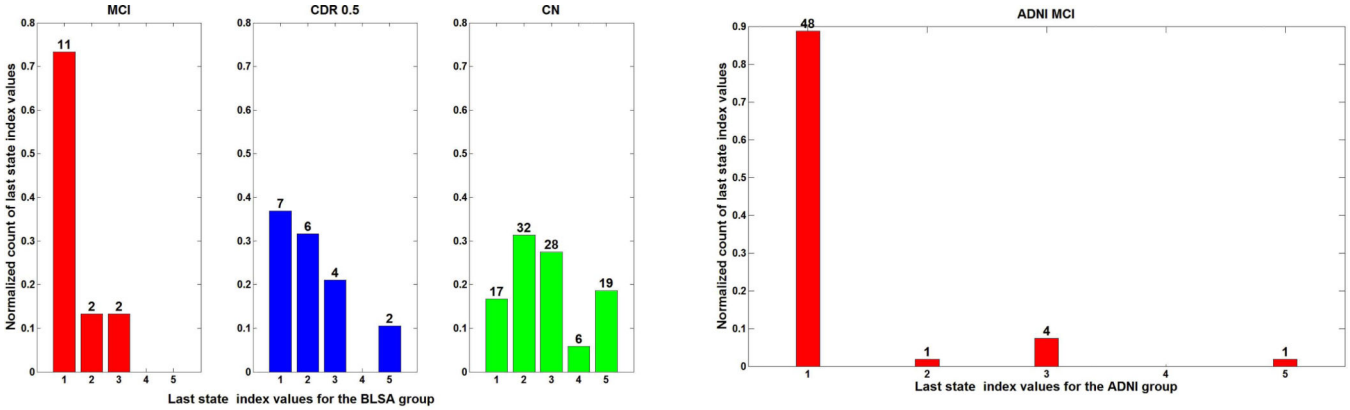


Figure 9.

Normalized count of the last state index during follow-up period from the three subgroups of the BLSA study (left) and the MCI group from the ADNI study (right) while the raw number showed on top of each bin.

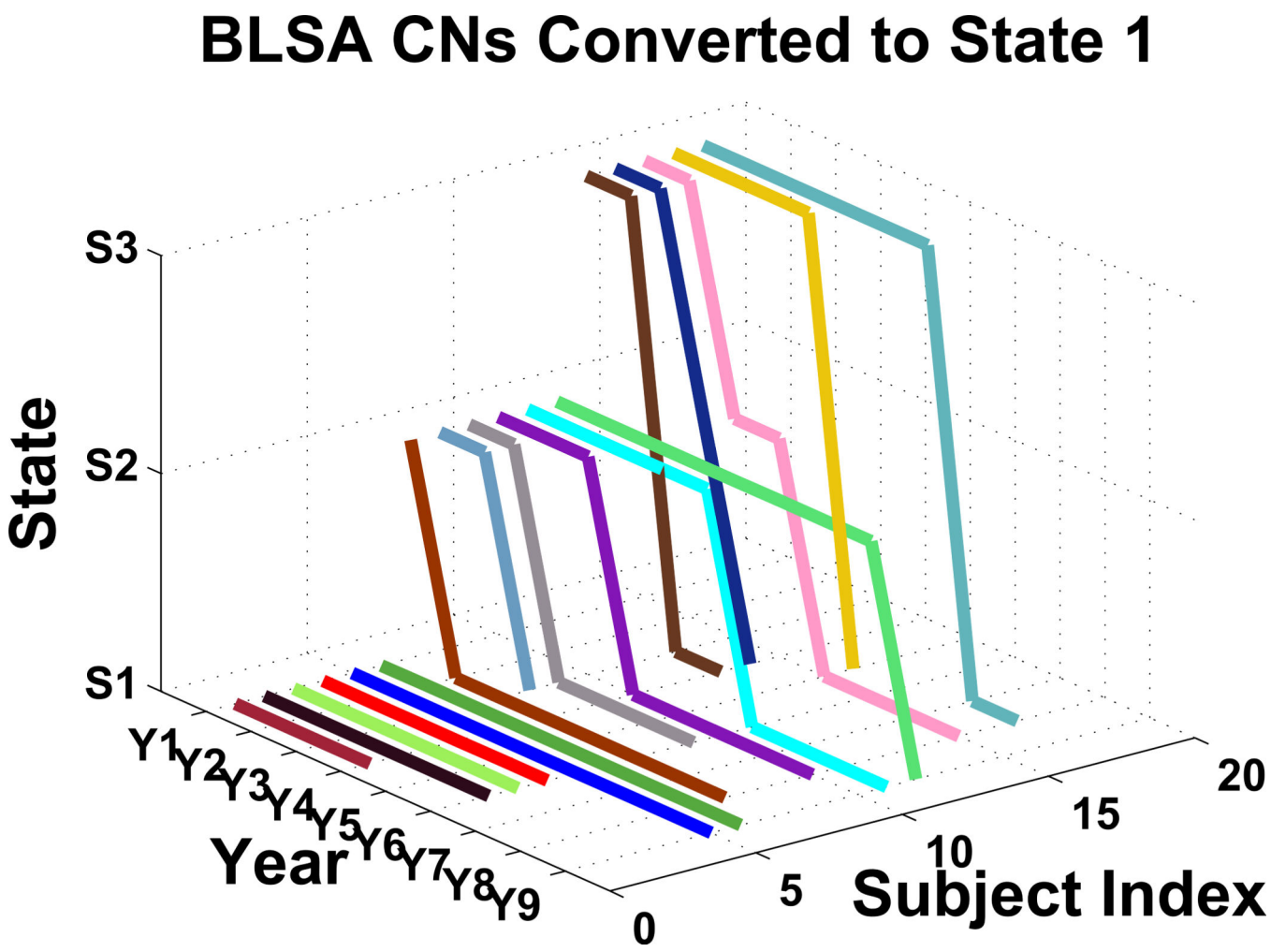


Figure 10.
State paths of the CNs with state conversion to 1 during the course of the BLSA study.

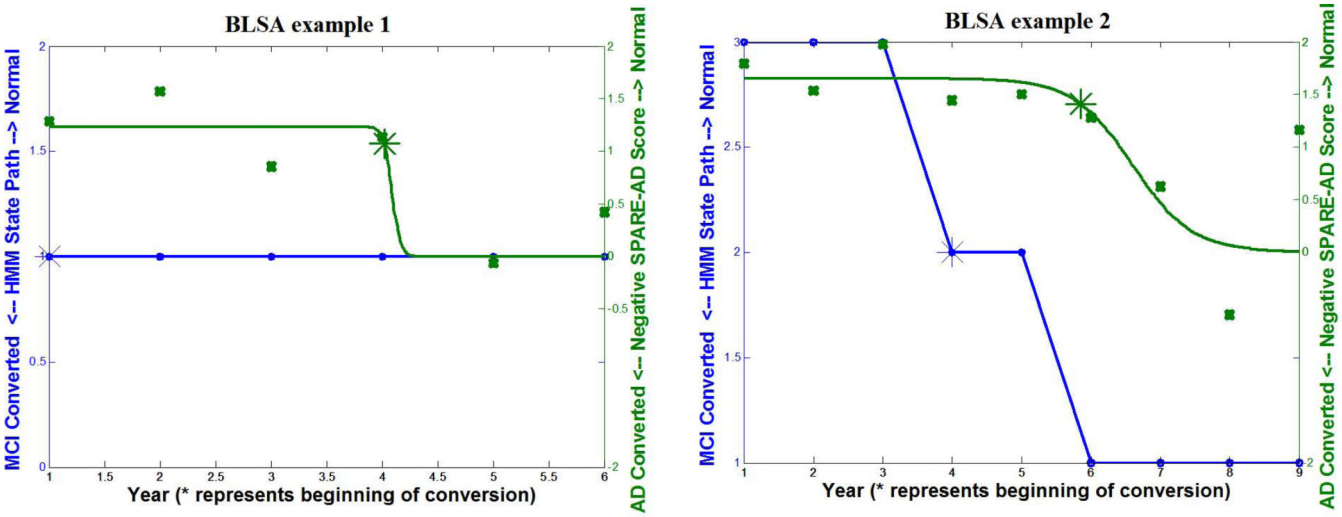


Figure 11.

Examples of the estimated SPARE-AD score curve and HMM state path from the BLSA study.

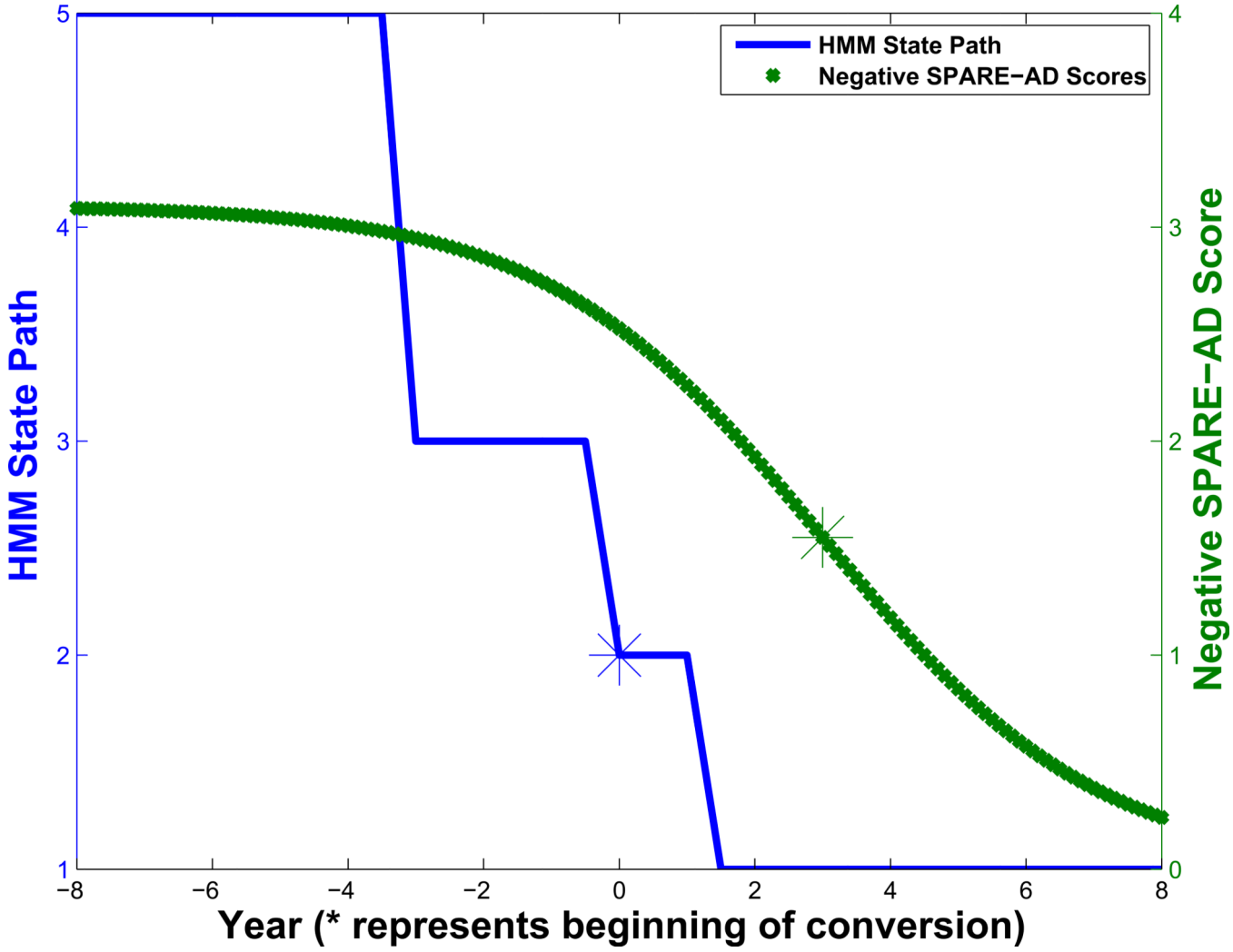


Figure 12.
The median HMM state path and SPARE-AD score sequence (time axis in this figure was normalized to HMM state path).

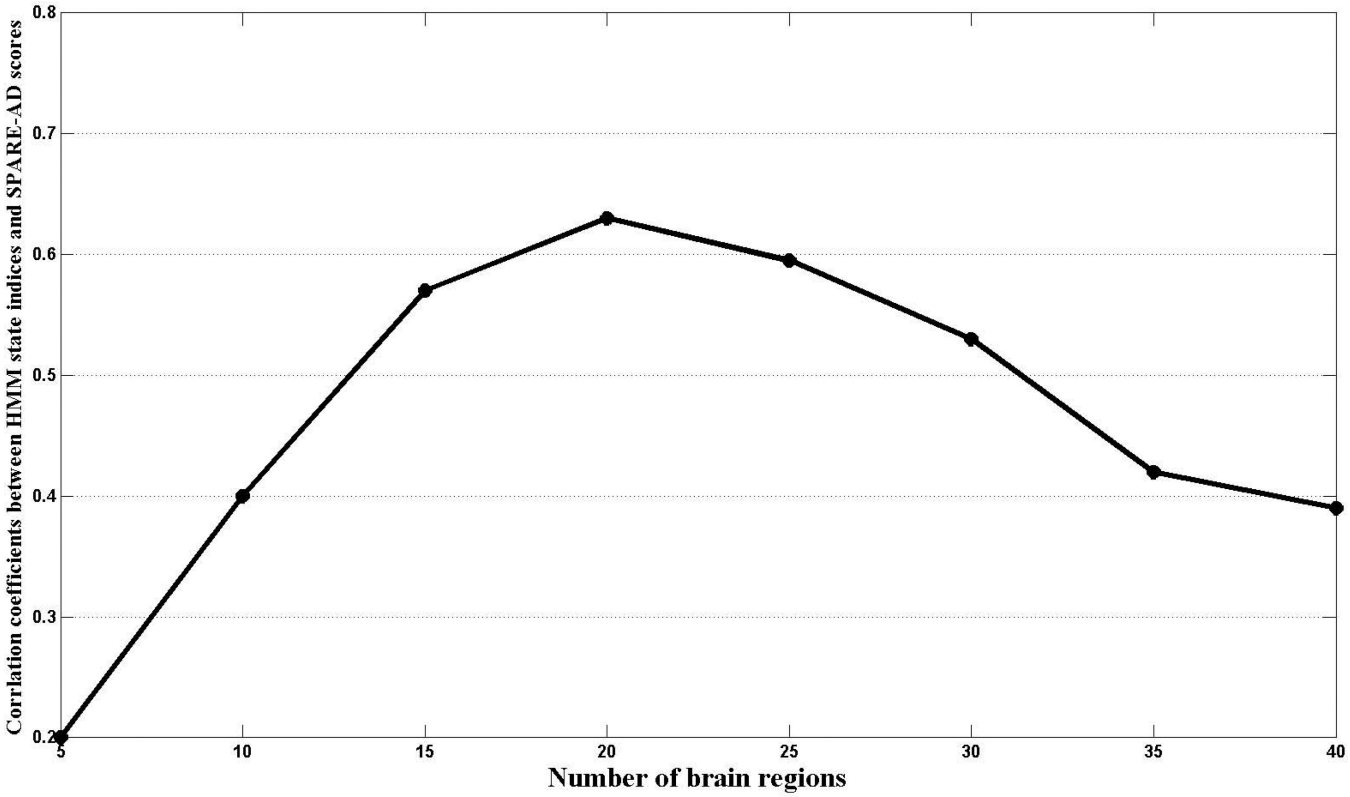


Figure 13.

The relationship between HMM state indices and SPARE-AD scores with increasing number of regional brain features.

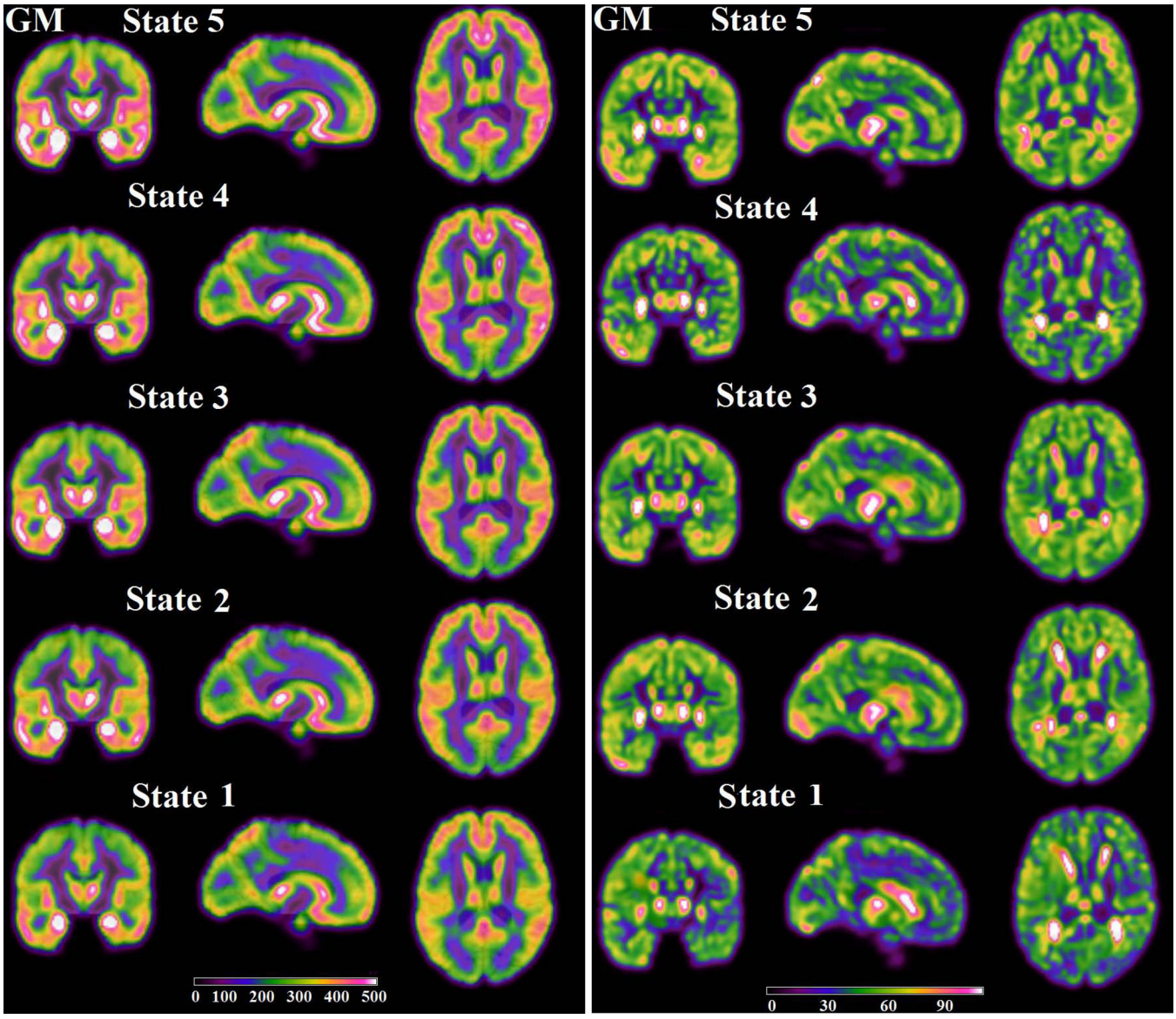


Figure 14.
The mean (right) and variance (left) images of brain grey matter for the five HMM-state groups, respectively.

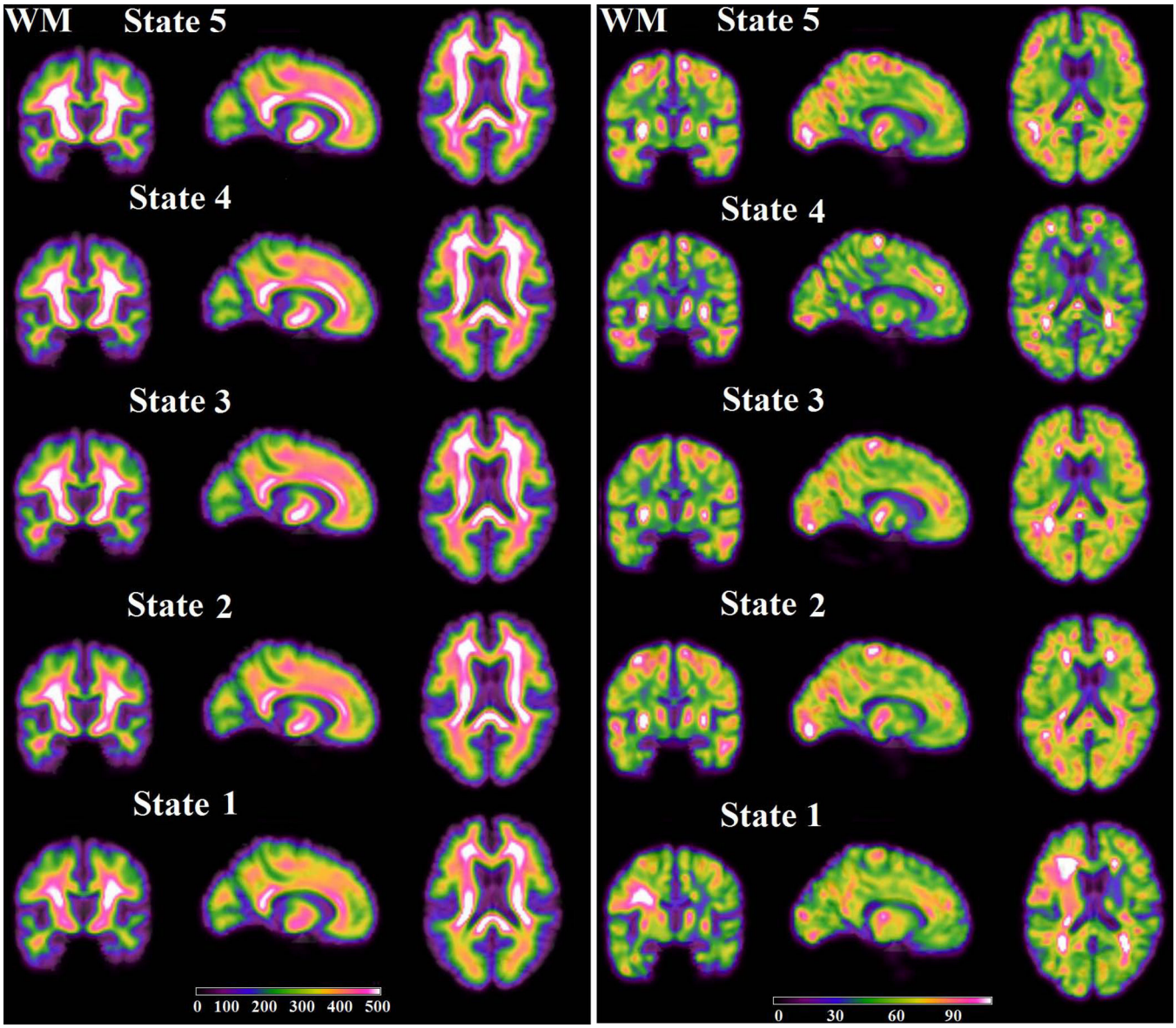


Figure 15.
The mean (right) and variance (left) images of brain white matter for the five HMM-state groups, respectively.

Table 1

Characteristics of the BLSA and ADNI participants in the present study

ADNI Group	MCI	AD	CN
No. of subjects	54	54	63
Total scans	130	54	63
Sex (M/F)	35/19	23/31	33/30
Baseline age, mean (SD)	76.13 (7.70)	77.41 (7.10)	75.23 (5.40)
Age at last visit, mean (SD)	77.86 (7.79)	*	*
MMSE at first visit, mean (SD)	26.41 (2.23)	23.20 (2.10)	29.2 (0.98)
MMSE at last visit, mean (SD)	24.15 (4.89)	*	*
Follow-up period (year), mean (SD)	2.40 (0.50)	*	*
BLSA Group	**MCI**	**CDR 0.5**	**CN**
No. of subjects	15	19	102
Total scans	94	138	753
Sex (M/F)	10/5	13/6	56/46
Baseline age, mean (SD)	77.12 (7.20)	72.54 (7.45)	70.29 (7.52)
Age at last visit, mean (SD)	82.90 (7.10)	79.23 (6.56)	76.79 (7.61)
MMSE at first visit, mean (SD)	27.21 (2.55)	28.56 (1.38)	28.91 (1.30)
MMSE at last visit, mean (SD)	25.40 (3.04)	27.88 (2.26)	28.84 (1.17)
CVLT at first visit, mean (SD)	44.73 (10.91)	52.05 (10.56)	53.61 (10.78)
CVLT at last visit, mean (SD)	36.40 (11.66)	52.95 (14.79)	53.56 (14.78)
Follow-up period (year), mean (SD)	5.27 (2.09)	6.26 (1.85)	6.38 (1.71)

Table 2

Characteristics of two CN subgroups from the BLSA study: Group 1 for individuals converted to state 1, and Group 2 for participants who remained in state 5 during the observation period.

BLSA CNs	Group 1	Group 2	CN
No. of subjects	17	19	102
Total scans	112	128	753
Sex (M/F)	13/4	6/13	56/46
Baseline age, mean (SD)	69.75(8.08)	69.72 (6.72)	70.29 (7.52)
Age at last visit, mean (SD)	76.19 (9.07)	76.40 (6.18)	76.79 (7.61)
MMSE at first visit, mean (SD)	28.67 (1.88)	29.05 (0.85)	28.91 (1.30)
MMSE at last visit, mean (SD)	28.59 (1.54)	29.41 (0.62)	28.84 (1.17)
CVLT at first visit, mean (SD)	49.19 (10.04)	51.74 (11.76)	53.61 (10.78)
CVLT at last visit, mean (SD)	47.94 (10.77)	54.74 (15.06)	53.56 (14.78)

Table 3

Correlation table of the HMM states and SPARE-AD indexes of each subgroup (BLSA study): MCI, CDR 0.5 group, and Group 1 including 17 individuals converted to state 1. Abnormal group was constituted by all three groups.

BLSA	MCI	CDR 0.5 group	Group 1	Abnormal group
Correlation	−0.58	−0.68	−0.51	−0.63