Send Orders for Reprints to reprints@benthamscience.ae

Current Alzheimer Research, **2015**, *12*, 69-79

Predicting Progression from Cognitive Impairment to Alzheimer's Disease with the Disease State Index

Anette Hall[1,*], Jussi Mattila[2], Juha Koikkalainen[2], Jyrki Lötjonen[2], Robin Wolz[3], Philip Scheltens[4], Giovanni Frisoni[5], Magdalini Tsolaki[6], Flavio Nobili[7], Yvonne Freund-Levi[8], Lennart Minthon[9], Lutz Frölich[10], Harald Hampel[11], Pieter Jelle Visser[4,12] and Hilkka Soininen[1,13]

[1]Institute of Clinical Medicine, Neurology, University of Eastern Finland, Kuopio, Finland; [2]VTT Technical Research Centre of Finland, Tampere, Finland; [3]Biomedical Image Analysis Group, Department of Computing, Imperial Collage London, London, UK; [4]Department of Neurology, Alzheimer Centre, VU University Medical Center, Amsterdam, the Netherlands; [5]IRCCS San Giovanni, Laboratory of Epidemiology and Neuroimaging, Brescia, Italy; [6]Aristotle University of Thessaloniki, Memory and Dementia Centre, 3rd Department of Neurology, "G Papanicolaou" General Hospital, Thessaloniki, Greece; [7]Clinical Neurology, Department of Neuroscience, Ophthalmology and Genetics, University of Genoa, Genoa, Italy; [8]Institution of NVS, Department of Geriatrics, Section of Clinical Geriatrics, Karolinska Institutet, Department of Geriatrics, Karolinska University Hospital, Huddinge, Stockholm, Sweden; [9]Clinical Memory Research Unit, Department of Clinical Sciences, Malmö Lund University, Sweden; [10]Department of Geriatric Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [11]AXA Research Fund & UPMC Chair, Sorbonne Universités, Université Pierre et Marie Curie, Institut de la Mémoire et de la Maladie d'Alzheimer & Institut du Cerveau et de la Moelle épinière (ICM), Département de Neurologie, Hôpital de la Pitié-Salpétrière, Paris, France; [12]Department of Psychiatry and Neuropsychology, University of Maastricht, Maastricht, Netherlands; [13]Department of Neurology, Kuopio University Hospital, Kuopio, Finland

Abstract: We evaluated the performance of the Disease State Index (DSI) method when predicting progression to Alzheimer's disease (AD) in patients with subjective cognitive impairment (SCI), amnestic or non-amnestic mild cognitive impairment (aMCI, naMCI). The DSI model measures patients' similarity to diagnosed cases based on available data, such as cognitive tests, the APOE genotype, CSF biomarkers and MRI. We applied the DSI model to data from the DESCRIPA cohort, where non-demented patients (N=775) with different subtypes of cognitive impairment were followed for 1 to 5 years. Classification accuracies for the subgroups were calculated with the DSI using leave-one-out cross-validation. The DSI's classification accuracy in predicting progression to AD was 0.75 (AUC=0.83) in the total population, 0.70 (AUC=0.77) for aMCI and 0.71 (AUC=0.76) for naMCI. For a subset of approximately half of the patients with high or low DSI values, accuracy reached 0.86 (all), 0.78 (aMCI), and 0.85 (naMCI). For patients with MRI or CSF biomarker data available, theywere 0.78 (all), 0.76 (aMCI) and 0.76 (naMCI), while for clear cases the accuracies rose to 0.90 (all), 0.83 (aMCI) and 0.91 (naMCI). The results show that the DSI model can distinguish between clear and ambiguous cases, assess the severity of the disease and also provide information on the effectiveness of different biomarkers. While a specific test or biomarker may confound analysis for an individual patient, combining several different types of tests and biomarkers could be able to reveal the trajectory of the disease and improve the prediction of AD progression.

Keywords: Alzheimer's disease, cerebrospinal fluid (CSF), computer-assisted diagnosis, dementia, DESCRIPA, magnetic resonance imaging (MRI), mild cognitive impairment (MCI).

1. INTRODUCTION

Alzheimer's disease (AD) is a major healthcare issue and its prevalence is expected to increase with the aging of the population [1]. Pre-dementia patients who exhibit an observable cognitive decline are often defined as having mild cognitive impairment (MCI) [2, 3]. These individuals are at risk of developing AD or other types of dementia disorders, but some remain stable or even improve over time. Different biomarkers related to the decline from MCI to AD have been identified, with some, such as CSF concentration of $A\beta_{42}$, being related directly to the presence of AD pathology and others, like MRI imaging, tracking the caused neuronal injury. However, the combined effects of biomarkers are quite weakly understood, especially in cases with contrary findings [3]. Multiple criteria have been published for diagnostics of MCI or prodromal AD [4-6], and discussion is ongoing on details like combining information from multiple test results or treatment of different populations, such as amnestic and non-amnestic MCI patients or those with subjective

*Address correspondence to this author at the University of Eastern Finland, Institute of Clinical Medicine/Neurology, P.O.Box 1627, 70211 Kuopio, Finland; Tel: +358-40-3552385; Fax: +358 17 162048;
E-mail: anette.hall@uef.fi

1875-5828/15 $58.00+.00

complaints. The potential of biomarkers is well-known, but the translation of the guidelines to the clinical practice in an objective and standardized way is not a straightforward task.

One vital step towards effective treatments and successful drug trials would be early and accurate prediction of progression from MCI to AD, which has become a major interest in AD research [7-13]. The Disease State Index (DSI) is a new statistical technique for analyzing heterogeneous multimodal data. It measures patients' similarity to previously diagnosed cases based on all available data. This is done stepwise by combining single measurements into groups, such as cognitive data or MRI, which in turn are combined into the total DSI. When compared to other classification methods, such as logistic regression or support vector machines and Bayes, DSI is able to distinguish between stable and progressing MCI patients at a comparable accuracy [14-16]. Unlike many classifiers, DSI works well with missing data and does not require searching for optimal parameters. The DSI has been designed for practical implementation as a clinical decision support system. The accompanying data visualization method, Disease State Fingerprint (DSF) presents DSI data in a visual form, enabling quick evaluation of the patient's status [14]. The Predict AD decision support software, which is based on both DSI and DSF, has been shown to improve diagnostic classification accuracy for clinicians, especially when dealing with larger amounts of heterogeneous data [16-18].

The aims of this study are to expand the evaluation of the DSI method from previous studies [14, 19-20] to another large cohort, which contains different subtypes of cognitive impairment. Progression to AD is evaluated using data from the DESCRIPA study [21], including cognitive tests, APOE genotype, $A\beta_{42}$, T-tau and P-tau concentrations from CSF and manifold-based learning (MBL), tensor-based morphometry (TBM) and segmentation of hippocampal volume from MRI. The patients are divided into groups according to their cognitive symptoms, allowing the comparison of factors related to progression to AD between the groups. By integrating multiple biomarker modalities with clinical data, we test how the DSI method works with different subtypes of MCI and whether it reveals any differences between them. We show that the DSI distinguishes between clear and ambiguous cases, achieving higher accuracies for patients with biomarker data resembling typical stable or AD-progressive cases.

2. METHODS

2.1. Patients

The patients were selected from the DESCRIPA study, which is a multicenter study of the European Alzheimer's Disease Consortium on the development of screening guidelines and clinical criteria for the diagnosis of pre-dementia AD [21]. Twenty memory clinics across Europe recruited patients aged 55 years and older with cognitive complaints, but without a diagnosis of dementia or a somatic, psychiatric or neurological disorder that could have caused the cognitive impairment. The medical ethics committee at each center approved the study and all patients provided informed consent. The patients were divided into groups based on neuropsychological tests [22]. If no impairments were found in the tests, the patients were classified as having subjective cognitive impairment (SCI). Patients with memory impairment were classified as amnestic MCI (aMCI) and those with impairments in other cognitive domains as non-amnestic MCI (naMCI) [23]. Patients with impairments in both amnestic and non-amnestic domains were classified as aMCI. Impairment was defined as z-scored test result of -1.5 or lower. The patients had annual follow-ups for 1 - 5 years, where progression to AD-type dementia was assessed according to the NINCDS-ADRDA criteria [24]. Table **1** shows the demographic data for the different cognitive impairment groups. In addition to the Mini-Mental State Examination (MMSE, n=770), which was available for most patients, there were neuropsychological tests (n= 671 – 749, depending on the test), CSF (n=174), APOE genotype (n=498), and MRI available (n=310) for subsets of patients.

2.2. Clinical and Neuropsychological Tests

As the patients were from several countries, the cognitive tests administered varied from center to center. Only the total score of the MMSE was available for all patients. Other neuropsychological tests were divided into domains of learning, memory, language, executive function, and visuoconstruction [21]. For learning and memory the Rey auditory verbal learning test, the CERAD word list or an equivalent test was used. Verbal fluency for names of animals was the test for language in most centers and Trail Making tests A and B were used for executive function. Visuoconstruction was assessed by the Rey-Osterrieth complex figure, the CERAD figures or a similar test. The results were compared to healthy controls according to age, gender and education and a z-score was calculated with the same normative dataset used in each center, if possible [21].

2.3. APOE

APOE genotype was determined on genomic DNA extracted from EDTA blood with the polymerase chain reaction [25]. Both the presence and number of APOE-ε4 alleles were used in this study.

2.4. CSF

CSF samples were analyzed at Sahlgrenska University Hospital, Mölndal in Sweden. The concentrations of $A\beta_{42}$, T-tau and P-tau were measured with single parameter enzyme-linked immunosorbent assay (ELISA) (Innotest β-amyloid 1-42; Innotesth TAU-Ag; Innogenetics, Ghent, Belgium). The CSF AD-profile was calculated from $A\beta_{42}$ and T-tau values as $CSF\ AD = \frac{A\beta_{42}}{240 + 1.18\,\mathrm{T\text{-}tau}}$ which provides scores below 1 for CSF profiles typically found in AD patients [26]. Both the raw values and the scalar AD-profile were used in this study.

2.5. MRI

MRI scans were conducted at either 1.0 or 1.5 Tesla in 9 different centers, each with their own scanners and protocols. For details, see Appendix. Three fully automated feature extraction methods were applied to the MRI images: manifold-based learning [27], tensor-based morphometry [28] and

segmentation of hippocampal volume [29]. MBL provides low-dimensional representations of the images that capture the variability in the dataset [27, 30-31]. In this study, pairwise image similarities in a region around the hippocampus and amygdala were used and the first 20 dimensions of the resulting manifold were used as features. Image similarities were measured as Normalized Mutual Information in the region of interest after aligning all MRI images to a template space with 10 mm non-rigid registration [32]. TBM measures regional structural differences that are computed from the deformations fields obtained by registering the MRI images to a common template [33]. TBM used a multi-template approach [28, 34] with thirty templates (10 controls, 10 MCI patients, and 10 AD patients) from the Alzheimer's Disease Neuroimaging Study (ADNI) [35]. Based on recent TBM studies [11, 28], four regions of interest were chosen – the amygdala, gyri hippocampal isetambiens, posterior temporal lobe and lateral ventricle temporal horn. Baseline hippocampal volumes (HCV) were measured using fast and robust multi-atlas segmentation [29, 33].

2.6. Disease State Index

The Disease State Index (DSI) is a statistical method that provides values ranging from 0 to 1, measuring the similarity of the patient's data to a diagnosed population [14]. DSI values close to zero indicate typical stable cases, while values near one are similar to cases that progressed to AD. The DSI method was applied to each patient using leave-one-out cross-validation, with the DESCRIPA cohort itself acting as the diagnosed population excluding the patient under analysis.

DSI computation is based on the differences in the distributions of measurement values between two diagnosed groups of patients – those who progressed to AD (positive group) and those who did not (control group). Similarity of a patient's measurement to the positive group is evaluated using a fitness function that provides a likelihood of the measured feature belonging to an AD-progressing patient. The result from the fitness function provides the measure of similarity for that particular feature.

The fitness as a function of measurement value x, is defined as

$$f(x)=FN(x)/(FN(x)+FP(x))$$

where $FN(x)$ is the false negative error rate and $FP(x)$ the false positive error rate in the training data, when using x as the classification threshold. The fitness function is a monotonically increasing function simplifying its interpretation.

In addition to evaluating the fitness of a patient measurement, the relevance of each feature is also calculated. Relevance is the ability of a feature to separate the two groups from each other; it has a value of 0 if the distributions overlap entirely and 1 if there is absolutely no overlap. The relevance is computed as

$$relevance = sensitivity + specificity - 1,$$

where sensitivity and specificity are obtained by classifying the diagnosed patients.

Patient measurements are combined into a composite DSI value using a weighted average, where the fitness values are weighted according to their relevance:

$$DSI = \frac{\sum relevance \cdot fitness}{\sum relevance}.$$

The process of evaluating fitness and relevance and combining the features to a composite DSI are repeated recursively until an overall DSI value from all available data is obtained for the patient. The DSI values can be interpreted as the percentage of data corresponding to a progressing AD profile, where the profile is defined from the data of diagnosed cases.

The individual measurements, composites of features and the total DSI value can be visualized with the Disease State Fingerprint (DSF). It gives a view of the patient's disease state, showing which diagnostics measures are relevant and whether they correspond to a stable patient, or one progressing to AD. The main purpose of the DSF is to help clinicians quickly interpret and analyze large quantities of heterogeneous patient data. The DSF illustrates each feature as colored boxes, with low DSI values in shades of blue, high values in red and those in the middle area as white. The relevance of each feature is indicated by the size of the box. A more comprehensive description of the DSI method, the DSF visualization and the derivation of the model can be found in Mattila *et al.*, 2011 and its supplementary data [14].

2.7. Data Analysis

The patients were grouped into MCI+SCI (including all cases), MCI (including naMCI and aMCI, excluding SCI), naMCI and aMCI. The SCI group was not analyzed separately, since the rate of progression from SCI to AD was too low (3%) to allow meaningful statistical analysis.

To remove features with only marginal information for classification, feature selection was applied by omitting features with relevance under 0.1, corresponding to average sensitivity and specificity < 0.55. This removed the visuoconstruction test and several of the 20 MBL features from all test groups and trail making A (not included in: aMCI, naMCI, MCI), B (naMCI, MCI), fluency (naMCI), TBM Lateral ventricle temporal horn (naMCI, MCI) and posterior temporal lobe (naMCI) from some groups.

The DSI model was evaluated using leave-one-out cross-validation, where one-by-one each patient is compared to all the other cases in the DESCRIPA cohort. Additionally, we tested each of the twenty centers using the other centers as the training group (leave-one-center-out cross validation). We show the results for the seven centers with data for at least 50 patients, to ensure patient confidentiality and maintain statistical relevance. DSI values under 0.5 were predicted as stable cases and over 0.5 as progressing to AD. This cut-off results in roughly equal specificity and sensitivity, but does not maximize classification accuracy. Classification accuracy is the percentage of patients whose diagnosis was assigned correctly by the DSI method, compared to the actual diagnoses at the end of the follow-up time. Classification accuracy is used interchangeably with prediction accuracy as the DSI predicts the final follow-up outcomes from baseline data.

Performance was measured with AUC (area under the receiver-operator curve (ROC)), prediction accuracy, sensitivity, and specificity. Statistical tests are considered significant if $p < 0.05$.

3. RESULTS

Demographic and clinical data of the groups are presented in Table **1**. The average (standard deviation) age for patients was 70 (7.8) years, 57% were female, and the average length of education was 10.4 (4.2) years. Average follow-up time was 2.5 (0.9) years and progression time to AD 1.1 (0.7) years. Percentage of patients that progressed to AD differed between the groups, being 3% for SCI, 16% for naMCI and 34% for aMCI. Supplementary Table **1** shows the demographic information and Supplementary Table **2** the data availability of the selected centers in the DESCRIPA study. MRI was available for at least some of the patients in 10 centers, CSF in 9 and both in 7 centers.

3.1. Classification Results

Group-specific results for predicting progression to AD-type dementia are shown in Table **2**. The results show that the DSI is able to classify patients as stable or AD-progressive with an accuracy of 0.75 for the whole DESCRIPA dataset (SCI+MCI). For the MCI groups the accuracies are slightly lower, 0.70 for aMCI and 0.71 for naMCI and MCI. The AUCs show a similar trend, with the highest AUC, 0.83, obtained for the whole cohort (SCI+MCI). AUC indicates the overall ability to distinguish between the two populations, independent of classification cutoff (DSI = 0.5). Due to the behavior of the DSI model, the sensitivity and specificity for each group are close to the corresponding accuracy. However, the naMCI group has a slightly lower sensitivity, partly due to the smaller size of the group.

The results were validated by comparison to leave-one-out SVM classification with Matlab fitcsvm and 10-fold cross validated DSI calculations for the whole (SCI+MCI)

Table 1. Demographic and clinical data. Abbreviations: SCI, cases with subjective cognitive impairment; naMCI, cases with non-amnestic MCI; aMCI, cases with amnestic MCI; MCI, cases with mild cognitive impairment, the combination of naMCI and aMCI; SCI +MCI, all cases; MMSE, Mini-Mental State Examination score; APOE4 0/1/2, number of cases with 0, 1 or 2 apolipoprotein ε4 alleles. NOTE: Values are mean (standard deviation) or number (percentage). Neuropsychological tests for learning, memory, fluency, trailmaking A, trailmaking B and figure show the mean of the Z-scored results.

	SCI	naMCI	aMCI	MCI	SCI+MCI
Patients (N)	231	196	348	544	775
Age (years)	67.9 (7.7)	70.5 (7.6)	71.1 (7.5)	70.9 (7.5)	70.0 (7.8)
Gender (female)	119 (52%)	113 (58%)	213 (61%)	326 (60%)	443 (57%)
Education (years)	11.6 (4.1)	9.6 (4.3)	10.1 (4.0)	9.9 (4.1)	10.4 (4.2)
Follow-up time (years)	2.6 (0.9)	2.6 (0.9)	2.4 (0.9)	2.5 (0.9)	2.5 (0.9)
Progressed to AD (N)	7 (3%)	31 (16%)	118 (34%)	149 (27%)	156 (20%)
Time to AD (years)	1.6 (1.3)	1.2 (0.9)	1.1 (0.8)	1.1 (0.8)	1.1 (0.7)
MMSE score	28.5 (1.5)	27.7 (2.0)	26.7 (2.4)	27.0 (2.3)	27.5 (2.2)
Learning	0.0 (0.9)	-0.3 (0.7)	-1.8 (0.8)	-1.3 (1.1)	-0.9 (1.2)
Memory	0.1 (0.9)	-0.3 (0.8)	-2.1 (0.8)	-1.5 (1.2)	-1.0 (1.4)
Language	0.5 (1.4)	-0.5 (1.3)	-0.8 (1.2)	-0.7 (1.3)	-0.4 (1.4)
Trailmaking A	0.3 (0.8)	-1.5 (1.7)	-0.8(1.6)	-1.1 (1.7)	-0.7 (1.6)
Trailmaking B	0.4 (0.7)	-1.8 (1.9)	-1.3 (2.0)	-1.5 (2.0)	-0.9 (1.9)
Figure	0.6 (0.7)	-0.3 (1.3)	0.0 (1.2)	-0.1 (1.2)	0.1 (1.1)
CSF Aβ-42 (pg/ml)	659 (255)	579 (288)	496 (247)	523 (263)	575 (267)
CSF T-Tau (pg/ml)	347 (199)	410 (275)	529 (385)	489 (356)	435 (313)
CSF P-Tau (pg/ml)	60 (27)	78 (49)	83 (54)	78 (49)	71 (43)
CSF AD Profile	1.12 (0.53)	0.95 (0.60)	0.72 (0.49)	0.80 (0.54)	0.92 (0.56)
Hippocampal volume (mm^3)	4398 (517)	4305 (541)	3979 (637)	4110 (620)	4192 (606)
APOE4 0/1/2 (N)	87/54/4 (60%/37%/3%)	79/41/6 (63%/32%/5%)	130/77/20 (57%/34%/9%)	209/118/26 (59%/33%/7%)	296/172/30 (59%/35%/6%)

Table 2. Classification results of the DSI. Abbreviation: AUC, Area under the receiver-operator curve.

DSI	MCI+SCI	MCI	aMCI	naMCI
AUC	0.83	0.77	0.77	0.76
Accuracy	0.75	0.71	0.70	0.71
Sensitivity	0.76	0.72	0.70	0.65
Specificity	0.75	0.70	0.70	0.72

cohort. SVM classification achieved an accuracy of 0.78, sensitivity 0.14 and specificity of 0.94. The accuracy is slightly higher than what we get with the DSI, but sensitivity is lost. 10-fold cross validated DSI obtained an AUC of 0.82, accuracy of 0.76, sensitivity of 0.79 and specificity of 0.75, which is very similar to what was obtained here with leave-one-out cross validation.

Classification results for each major center using all the other centers as training data can be found in supplementary Table **3**. The combined results for MCI+SCI and MCI match those obtained with leave-one-out cross validation. For aMCI and naMCI the leave-one-center-out results are slightly less predictive than those in Table **2**. Many of the differences between the centers can be explained by a low number of patients or poor data availability.

Table **3** presents the group-specific results for the combined feature sets: cognitive (including MMSE and neuropsychological tests), genetic (including APOE), CSF (including Aβ-42, tau features and AD-profile) and MRI (including MBL, TBM and HCV). While some neuropsychological data was available for all of the patients, APOE, CSF and MRI biomarker data were only available for subpopulations. Table **3** shows the effect of CSF and MRI biomarker data on prediction accuracy for subgroups having additional biomarker data. The first group (CSF/MRI) includes cases that have CSF or MRI data, or both, available and the second includes cases that have both CSF and MRI available in addition to cognitive tests and possibly genetic data.

The results show that for all groups MRI is the most accurate and genetic data the least accurate predictor. According to the paired McNemar test, genetic tests have a lower accuracy than cognitive, CSF and MRI tests when considering all patients. Accuracy of genetic tests is significantly lower than MRI in MCI and aMCI groups and CSF in aMCI group. McNemar's test also shows that prediction accuracy increases in the MCI group when MRI or CSF data are available, compared to having only cognitive data available. The same is true for the aMCI group when both MRI and CSF data are available. The cognitive tests are the most effective predictors of AD progression for the SCI+MCI group, and least effective for aMCI. Overall, the results show that the inclusion of additional biomarker data, over cognitive data alone, improves prediction accuracy.

Prediction accuracy of the DSI should be higher near the edges of the scale, since patients with lower index values resemble typical stable cases more closely and those with high values are very similar to known AD-progressive cases. Here, patients with total DSI values lower than 0.3 or higher than 0.7 are defined as clear stable and progressive cases, respectively. Group-wise prediction accuracy and the percentage of patients assigned as clear cases for all DESCRIPA patients and for those with CSF and/or MRI are shown in Table **4**. Approximately half of the patients were assigned as clear cases and the prediction accuracies for clear cases were higher than for the whole cohort.

3.2. Visualization of the DSI

Fig. (**1**) presents the distributions of total DSI values for stable and progressive patients within different cognitive subgroups (aMCI, naMCI and SCI) when using MCI+SCI as the DSI training group. The distributions clearly differ between groups, with SCI cases having lower and aMCI higher DSI values, for both stable and progressive patients.

Table 3. Accuracy for combined features and cases with biomarker data available. Abbreviations: MRI/CSF, Only cases with MRI or CSF data available; MRI+CSF, Only cases with both MRI and CSF data available; N, Number of available cases.

Accuracy (N)	SCI+MCI	MCI	aMCI	naMCI
Cognitive	0.72 (775)	0.66 (544)	0.65 (348)	0.68 (196)
Genetic	0.62 (498)	0.61 (353)	0.59 (227)	0.65 (126)
CSF	0.72 (174)	0.68 (108)	0.72 (72)	0.67 (36)
MRI	0.76 (312)	0.74 (223)	0.75 (134)	0.72 (89)
MRI/CSF	0.78 (371)	0.76 (259)	0.76 (157)	0.76 (102)
MRI+CSF	0.79 (114)	0.77 (71)	0.85 (48)	0.70 (23)

Table 4. Accuracy for clear (DSI <0.3 or >0.7) cases. Abbreviations: ALL, All available cases; MRI/CSF, Only cases with MRI or CSF data available; MRI+CSF, Only cases with both MRI and CSF data available; % Cases, Percentage of cases assigned DSI values < 0.3 or > 0.7.

Accuracy (% of Cases)	SCI+MCI	MCI	aMCI	naMCI
ALL	0.84 (58%)	0.78 (52%)	0.78 (50%)	0.85 (49%)
MRI/CSF	0.90 (56%)	0.85 (53%)	0.83 (54%)	0.91 (54%)
MRI+CSF	0.97 (52%)	0.89 (39%)	0.91 (46%)	1.00 (39%)

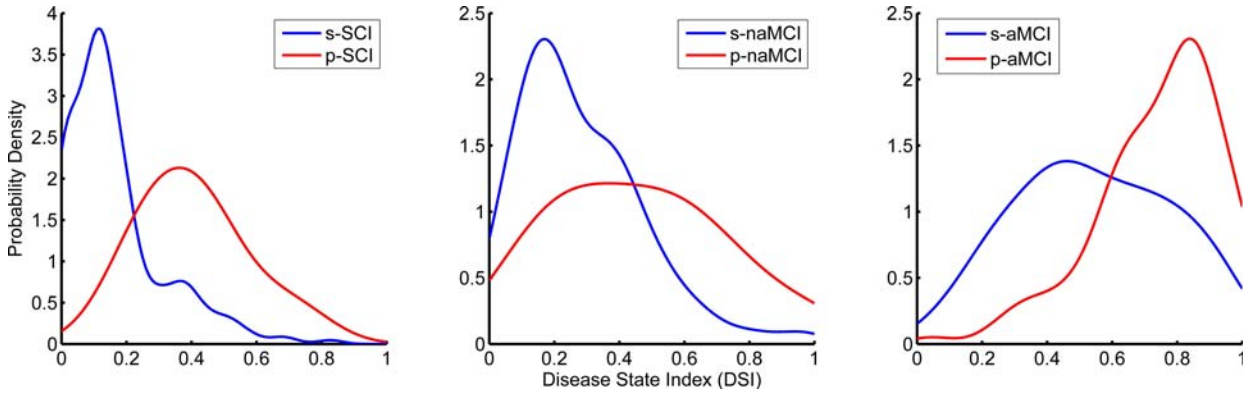

Fig. (1). Probability density distributions of the DSI values for stable and progressive SCI, naMCI and aMCI patients when analyzed in a single group (SCI+MCI). The distributions show the effect of systematic differences on the calculated DSI values. Stable and progressive populations for each group are normalized separately, and do not reflect the relative number of patients.

The DSFs of average stable and progressive DESCRIPA patients for each cognitive subgroup, tested against the MCI+SCI training group, are shown in Fig. (**2**). The values for the average patients were obtained by calculating the mean for each measurement from the corresponding group of cases.

The average stable cases exhibit lower DSI values than the progressive ones. The average stable SCI patient has the lowest total index of 0.13 and the overall profile indicates a very typical stable case. The average stable naMCI patient has a total DSI value of 0.17, while the average stable aMCI patient has a total DSI value of 0.39. The average progressive naMCI patient has a total DSI value of 0.56 and average progressive aMCI patient a total DSI of 0.89. The differences in average indices between the progressive MCI subtypes is due to the DSI measuring similarity to the disease group in the training data, which contains more aMCI cases. That is, aMCI patients are more similar to the disease group – consisting mostly of aMCIs– than naMCI patients, and as such exhibit larger DSI values by definition.

Based on the DSFs, the three cognitive subgroups mainly differ from each other in their neuropsychological test results. The memory and learning tests were originally used to classify the patients, so both stable and progressive aMCI had impairment, while naMCI cases do not. Similarly, all naMCI cases have impairment in fluency or trail making tests. Otherwise, the average stable SCI and naMCI are similar to each other, while stable aMCI cases have slightly higher overall DSI values both in MRI and CSF. The average progressive naMCI patient has lower DSI scores on cognitive tests, MRI and CSF, compared to the average progressive aMCI patient. This indicates that aMCI and naMCI pathologies differ from each other and could be considered separately.

Progressive aMCI and naMCI differ from each other significantly (independent t-test) in learning, memory, CSF amyloid β-42, CSF profile, MBL dimension 7, TBM amygdala and hippocampal volume. Stable aMCI and naMCI differ in MMSE, learning, memory, trailmaking A and B, MBL dimension 7, TBM of gyri hippocampalis et ambiens and hippocampal volume. Stable naMCI differs from stable SCI in all cognitive tests, MBL dimensions 4 and 7 and TBM of posterior temporal lobe and lateral ventricle temporal horn, while stable aMCI differs in all cognitive tests, the CSF AD-profile, MBL dimensions 1 and 4 and hippocampal volume.

4. DISCUSSION

Modern machine learning methods are getting more attention in neurosciences, but their use in clinical practice is still very limited. Several reasons may explain the situation: 1) the applicability of the methods to a certain clinical question has not been validated thoroughly enough using multiple cohorts, 2) available biomarkers might not be accurate enough for affecting clinical decisions and 3) the methods are mathematically too complex and their results too obscure for clinicians who must finally take the responsibility about the diagnosis.

This work tries to answer to these challenges by both validating the DSI method in a large patient cohort, while also showing that good prediction accuracies can be obtained

Fig. (2). DSI Fingerprint data for average measurement values from all stable and progressive SCI, naMCI and aMCI patients tested in the MCI+SCI analysis group. DSI values are shown next to the nodes and visualized also with colors for rapid interpretation. The blue color indicates that the patient's measurement values are similar to those of stable patients and the red color reveals that the data are similar to those cases who progressed to AD during the follow up. The relevance of a test is indicated by the size of the box, with the most important tests shown above those with less relevance. (MBL=Manifold based learning. TBM=Tensor based morphometry. HCV=Hippocampal volume. MMSE=Mini-Mental State Examination. Amnestic=Combination of memory and learning tests. Non-amnestic=Combination of fluency and trail making tests. CSF-profile=$A\beta_{42}$/(240+1.18*T-tau). APOE=Number of APOE4 alleles).

for patients with sufficient biomarker data available. Further, the DSI method allows easy integration of heterogeneous data such as neuropsychological test results and imaging biomarkers, while the related Disease State Fingerprint (DSF) provides a transparent way of showing the data that the prediction is based on.

4.1. Comparison to Previous Studies

The results in this study are comparable to previous studies using similar MCI datasets and the DSI method appears to generalize on several datasets with respect to early AD and MCI cases [7, 9, 11, 14, 36, 37]. A recent review paper by Falahati *et al.* (2014) looked at studies that utilize machine learning and multivariate analysis to predict conversion from MCI to AD [38], focusing also on five selected high-quality MRI only and multimodal studies. In our results of the MCI group the accuracy for MRI was 0.74, while in the selected five MRI only studies the accuracy ranged between 0.68 and 0.75. For multimodal analysis the accuracy in the studies was between 0.68 and 0.69, while in this study we obtained an accuracy of 0.71.

Clear cases with MRI or CSF data were classified accurately (83 - 91%), especially when having both data available (89 - 100%). These results are in line with previous results using the ADNI data, where about 50 % of stable and progressive MCI cases, most of who had CSF or MRI data, were predicted at accuracy of 87 % one year before the clinical diagnosis of AD [19]. These accuracies are very high, when you take into account that the clinical diagnosis, which the predictions are based on, does not always give certainty of an underlying AD pathology.

A recent study using the DESCRIPA dataset compared accuracies of APOE genotype, hippocampal volume and CSF biomarkers for predicting AD progression. They found these biomarkers to be useful for predicting AD in both aMCI and naMCI patients, although with reduced sensitivity for naMCI [23]. They noted that biomarker levels might need different cutoffs when predicting AD-type dementia for the different cognitive impairment groups. In the DSF analysis, the average stable and progressive cases differ from each other notably, suggesting that naMCI should be analyzed separate from aMCI.

Another recent study using the DESCRIPA dataset, comparing different medial temporal lobe measurements, had manual hippocampus volume measurements reaching an AUC of 0.71 and automatic atlas-based hippocampal segmentation an AUC of 0.74 [39]. In comparison, the AUC obtained with the HCV method used in this work was 0.75, clearly on par with these methods.

4.2. Feature-Specific Analysis

The DSI models for predicting progression to AD provide information about features that distinguish the stable MCI cases from those who progress to AD. The relevance of neuropsychological tests depended on which cognitive impairment group was analyzed. In the aMCI group, memory and learning tests were not as predictive as patients were already impaired on these tests, while in naMCI, fluency and trail making tests were not relevant for predicting disease progression. The biomarker model by Jack *et al.* (2013) suggests that function declines closest to the point of progression to AD [40]. Two recent studies have found that functional status and a decline in function are important predictors of progression to AD [7, 41].

With respect to the biomarkers, the differences between the groups were less pronounced. APOE had a small but relevant effect in all groups. Since APOE is a risk factor and does not change as the disease progresses, it is not surprising that it is not as predictive of conversion as disease related changes in CSF or MRI. CSF tau features and the CSF AD-profile proved to be very effective in predicting progression in all groups. Amyloid β-42 concentrations were less relevant, especially for naMCI cases.

Of the MRI analysis methods, MBL performed well in all groups. Although the individual manifolds were not very strong at predicting AD, their combination achieved good classification accuracy. TBM was a good predictor for AD in most groups, with the area around amygdala being clearly the most relevant of the TBM features. Hippocampal volume exhibited high accuracy and relevance for the prediction of AD, making it a very good method for this purpose. It was one of the most relevant features for every group. Overall, the MRI features performed at an expected level compared to previous studies [11, 20, 42].

4.3. Limitations and Future Directions

The DSI analysis was limited by the lack of CSF and MRI biomarker data on many of the patients, especially for progressive SCI and naMCI cases. Another limitation was the follow up period, which varied and was on average 2.5 years. It is possible that some stable cases with a high DSI value actually progressed to AD later or suffered from another neurodegenerative disease.

Many of the features used in this study have a dependence on age, gender, or APOE genotype. Personalization of data analysis could be achieved by comparing patient data only to patients of a similar age, gender, and genotype or by normalizing the data using methods such as Z-scoring, or by correcting the patient data based on measurement variability from healthy elderly controls. A recent study was able to improve the classification accuracy of progression from MCI to AD by personalizing the analysis with a combination of different approaches [43].

The results indicate that aMCI and naMCI pathologies are different. In clinical decision support, there could be a benefit to using two prediction models for disease progression – one for aMCIs and another for naMCIs – tuned to optimally reveal the likelihood of progression in either case. Another option is to use only one model but provide different DSI cut-offs for patients with symptoms referring to amnestic or non-amnestic pathology.

The DSI method has also recently been tested by using training and testing data from different cohorts, in order to provide information on its generalizability [44]. This could be further expanded to cover cohorts with different types of dementia, not only AD, as well as testing it in a clinical setting.

CONCLUSION

The DSI model provides good prediction results for the DESCRIPA dataset, which are also in line with previous studies performed in the same cohort [23, 37, 39]. Despite having patients with different profiles of memory complaints and impairment, the DSI method provided consistent predictions based on the heterogeneous data available from the patients. The model is able to distinguish between clear and ambiguous cases, giving higher prediction accuracies for clear cases, especially if CSF or MRI biomarker data is available.

The results suggest that by combining several different types of tests and biomarkers one can improve the prediction of AD in MCI patients. Although with most patients there are some tests or biomarkers that confound the analysis, a comprehensive data-driven study of the available measurements tends to reveal the state of the disease. The DSI and its accompanying DSF visualization should help clinicians in assessing the overall state of the patient based on the available measurement data.

ABBREVIATIONS

AD = Alzheimer's disease

aMCI = Amnestic MCI

AUC = Area under the receiver-operator curve

DSF = Disease State Fingerprint

DSI = Disease State Index

HCV = Hippocampal volume

MBL = Manifold-based learning

MCI = Mild cognitive impairment

MMSE = Mini-Mental State Examination

MRI = Magnetic resonance imaging

naMCI = Non-amnestic MCI

SCI = Subjective cognitive impairment

TBM = Tensor-based morphometry

CONFLICT OF INTEREST

A. Hall and H. Soininen received funding from the University of Eastern Finland.

J. Mattila, J. Koikkalainen and J. Lötjönen report that VTT Technical Research Centre of Finland owns the patents (U.S. Patent No. 7,840,510, Inventors: JK, JL; PCT/FI2010/050545, pending, Inventors: JM, JL, JK) that cover parts of the methods presented in the paper.

R. Wolz works as a consultant for IXICO Ltd.

F. Nobili has received fees from Bayer-Shering and Eli-Lilly for consultation in 2011-2012.

H. Hampel was supported by Jung-Diagnostics, Hamburg, Germany.

Dr. Visser has served as an advisory board member of Bristol-Myers Squibb and Roche diagnostics. He receives/received research grants from Bristol-Myers Squibb, European Commission 6th and 7th Framework programme, Joint Programme Initiative on Neurodegeneraion, Zon-Mw, the Innovative Medicine Initiative, and Diagenic, Norway.

ACKNOWLEDGEMENTS

This study was partly funded by the European Union 7th Framework Program for research, technological development and demonstration under grant agreements no 601055 and 611005, Predict AD, 5th Framework Program DESCRIPA study QLK-6-CT-2002-02455, the Health Research council of The Academy of Finland (H. Soininen), and strategic funding for UEFBRAIN from University of Eastern Finland (A. Hall, H. Soininen).

H. Hampel was supported by grants of the Katharina-Hardt-Foundation, Bad Homburg, Germany. Harald Hampel is supported by the AXA Research Fund, the Fondation Université Pierre et Marie Curie and the Fondation pour la Recherchesur Alzheimer, Paris, France. The research leading to these results has received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06.

A. Hall, J. Mattila, J. Koikkalainen, J. Lötjönen and H. Soininen designed the research project and participated in the writing of the manuscript. A. Hall conducted the DSI analysis and wrote the first draft of the manuscript. J. Lötjönen, J. Koikkalainen and R. Wolz were responsible for the MRI analysis. P. Scheltens, G. Frisoni, M. Tsolaki, F. Nobili, Y. Freund-Levi, L. Minthon, L. Frölich, H. Hampel, H. Soininen and P.J. Visser were involved in the data collection. All authors took part in reviewing and revising the manuscript.

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Demographics of selected centers in DESCRIPA

Supplementary Table 2. Number of patients with data available from selected centers

Supplementary Table 3. Classification results of the DSI for selected centers and the combined results from all centers

APPENDIX

Scan parameters and MRI protocols used at each DESCRIPA study center [39].

Huddinge

Siemens Avanto 1.5 T, 21 slices, FOV 220 mm, FOV phase 87.1, distance factor 30, phase R>L, slice thickness 5.0 mm, TE: 96 ms, TR: 4000 ms, flip angle 150◦, number of averages 1.

Siemens Symphony 1.5 T, 21 slices, FOV 220 mm, FOV phase 75.0, distance factor 30, phase R>L, slice thickness 5.0 mm, TE: 99 ms, TR: 4100 ms, flip angle 150◦, number of averages 2.

Kuopio

Siemens Vision 1.5 T, T1 3D-scan, MPRAGE OBL; COR>TRA, FOV 250, mat 256x256, 128 slices, TR 9.7ms, TE 4 ms, Slice thickness 2.0mm, no slice gap, flip angle 12◦

Malmö

Siemens Sonata 1.5 T, MPRAGE + lmpr-cor, 144 slices, FOV 250 mm, phase R>L, TR 1970, TE 3.93, distance factor 50, slice thickness 1.5 mm, flip angle 15◦.

Munich

Siemens Magnetom Vision; 1.5 T; MPRAGE; Slice thickness 1.05 mm, TR 11.4, TE 4.4, TI 300; FOV 256*256; Flip angle 8^{o}; number of averages 1.

Thessaloniki

Siemens Expert Plus unit 1.0 T, 3D-MPR: 15 (TR), 7 TE, Flip angle 8^{o}, 250 (Slabth), 1,49Ef thick, 168 Partitions, 250 FOV, 256x192 Matrix, 1 Aquis., ACQ TIME 10, 21min

Brescia

Philips Gyroscan PG 1.0 Tesla: sagittal 3D T1 scan: TR= 20 ms, TE= 5 ms, flip angle= 30°, field of view= 220 mm, acquisition matrix 256 × 256, number of slices= 100/130, slice thickness 1,3 mm.

Genoa

Philips Intera 1.5 Tesla: Sagittal MP RAGE : t1w/3d/tfe, TR 8.5, TE 3.9, Flip 8, thickness 1, no overcontiguous, no gap, matrix 100% 256, fov 256, reconstructed voxel size 1.00/1.00/1.00, rfov 100%, T1 t1w/se, TR 580, TE 15, flip 69, slices 22, thick 5mm, matrix 80% 240, fov 240, rfov 75%, pre (tra)- and postGD-GDPA (tra, sag, cor)

Maastricht

Philip NT, 1.5 T Gyroscan: T1-weighted images obtained in the coronal plane using a 3D-gradient fast field echo (FFE) sequence. TR = 35 ms, TE = 7 ms, FA= 35, FOV= 240 mm, slice thickness = 1.5 mm, matrix size = 256x256, voxelsize = 0.94mm x 0.94mm x 1.5 mm.

Mannheim

Siemens Medical solution Magnetom, Vision plus 1.5 Tesla: T1 MPR 30, TR=11.4 ms, TE=4.4 ms, flipangle=15, FoV=256mm, format 8/8, slices=162, no gap.

VUmc

Siemens Magnetom Impact Expert 1.0 T, 3D scan, 168 slices, FOV 250 mm, matrix 256 × 256; slice thickness 1.5 mm, TE: 7 ms, TR: 15 ms, TI 300 ms, flip angle 15◦.

REFERENCES

[1] Ballard C, Gauthier S, Corbett A, Brayne C, Aarsland D, Jones E. Alzheimer's disease. Lancet 377: 1019-31 (2011).

[2] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, *et al.* International Psychogeriatric Association Expert Conference on mild cognitive impairment. Mild cognitive impairment. Lancet 367: 1262-70 (2006).

[3] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, *et al.* The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7(3): 270-9 (2011).

[4] Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, *et al.* The diagnosis of mild cognitive impairment due to Alzheimer's disease:recommendations from the National Institute on Aging-Alzheimer's Associationworkgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 7(3):270-9 (2011).

[5] Morris JC, Blennow K, Froelich L, Nordberg A, Soininen H, Waldemar G, *et al.* Harmonized diagnostic criteria for Alzheimer's disease:recommendations. J Intern Med 275(3): 204-13 (2014).

[6] Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, *et al.* Advancing researchdiagnostic criteria for Alzheimer's disease: the IWG-2 criteria. Lancet Neurol 13(6): 614-29 (2014).

[7] Cui Y, Liu B, Luo S, Zhen X, Fan M, Liu T, *et al.* Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors. PLoS One 6: e21896 (2011).

[8] Landau SM, Harvey D, Madison CM, Reiman EM, Foster NL, Aisen PS, *et al.* Comparing predictors of conversion and decline in mild cognitive impairment. Neurology 75: 230-38 (2010).

[9] Liu Y, Paajanen T, Zhang Y, Westman E, Wahlund LO, Simmons A, *et al.* Analysis of regional MRI volumes and thicknesses as predictors of conversion from mild cognitive impairment to Alzheimer's disease. Neurobiol Aging 31: 1375-85 (2010).

[10] Spulber G, Simmons A, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, *et al.* An MRI-based index to measure the severity of Alzheimer's disease-like structural pattern in subjects with mild cognitive impairment. J Intern Med 273(4): 396-409 (2013).

[11] Wolz R, Julkunen V, Koikkalainen J, Niskanen E, Zhang DP, Rueckert D, *et al.* Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease. PLoS One 6: e25446 (2011).

[12] Barnes DE, Cenzer IS, Yaffe K, Ritchie CS, Lee SJ. Alzheimer's Disease Neuroimaging Initiative. A point-based tool to predict conversion from mild cognitive impairment to probable Alzheimer's disease. Alzheimers Dement pii: S1552-5260(13)02941-5 (2014).

[13] Gaser C, Franke K, Klöppel S, KoutsoulerisN, Sauer H, *et al.* BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. PLoS One 8(6): e67346 (2013).

[14] Mattila J, Koikkalainen J, Virkki A, Simonsen A, van Gils M, Waldemar G, *et al.* A disease state fingerprint for evaluation of Alzheimer's disease. J Alzheimers Dis 27: 163-76 (2011).

[15] Mattila J, Koikkalainen J, Virkki A, van Gils M, Lötjönen J. Alzheimer's Disease Neuroimaging Initiative. Design and application of a generic clinical decision support system for multiscale data. IEEE Trans Biomed Eng 59: 234-40 (2012a).

[16] Simonsen AH, Mattila J, Hejl AM, Garde E, van Gils M, Thomsen C, *et al.* Application of the PredictAD decision support tool to a Danish cohort of patients with Alzheimer's disease and other dementias. Dement Geriatr Cogn Disord 37(3-4): 207-13 (2014).

[17] Simonsen AH, Mattila J, Hejl AM, Frederiksen KS, Herukka SK, Hallikainen M, *et al.* Application of the Predict AD software tool to predict progression in patients with mild cognitive impairment. Dement Geriatr Cogn Disord 34(5-6): 344-50 (2012).

[18] Liu Y, Mattila J, Ruiz MÁ, Paajanen T, Koikkalainen J, van Gils M, *et al.* Predicting AD conversion: comparison between prodromal AD guidelines and computer assisted Predict AD tool. PLoS One 8(2): e55246 (2013).

[19] Mattila J, Soininen H, Koikkalainen J, Rueckert D, Wolz R, Waldemar G, *et al.* Optimizing the diagnosis of early Alzheimer's disease in mild cognitive impairment subjects. J Alzheimers Dis 32: 969-79 (2012b).

[20] Muñoz-Ruiz M, Hartikainen P, Hall A, Mattila J, Koikkalainen J, Herukka SK, *et al.* Disease State Fingerprint in frontotemporal degeneration with reference to Alzheimer's disease and mild cognitive impairment. JAD 35: 727-39 (2013).

[21] Visser PJ, Verhey FR, Boada M, Bullock R, De Deyn PP, Frisoni GB, *et al.* Development of screening guidelines and clinical criteria for predementia Alzheimer's disease. The DESCRIPA Study. Neuroepidemiology 30(4): 254-65 (2008).

[22] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, *et al.* Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. Lancet Neurol 8: 619-27 (2009).

[23] Vos S, van Rossum IA, Verhey F, Knol DL, Soininen H, Wahlund LO, *et al.* Prediction of Alzheimer's disease in subjects with amnestic and non-amnestic MCI. Neurology 80: 1124-32 (2013).

[24] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work-Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. Neurology 34: 939-944 (1984).

[25] Norberg J, Graff C, Almkvist O, Ewers M, Frisoni GB, Frölich L, *et al.* Regional Differences in Effects of APOE ε4 on Cognitive Impairment in Non-Demented Subjects. Dement Geriatr Cogn Disord 32:135-142 (2011).

[26] Hulstaert F, Blennow K, Ivanoiu A, Schoonderwaldt HC, Riemenschneider M, De Deyn PP, *et al.* Improved discrimination of AD patients using beta-amyloid(1-42) and tau levels in CSF. Neurology 52:1555-62 (1999).

[27] Wolz R, Aljabar P, Hajnal JV, Rueckert D. Manifold learning for biomarker discovery in MR imaging. Machine Learning in Medical Imaging 116-23 (2010).

[28] Koikkalainen J, Lötjönen J, Thurfjell L, Rueckert D, Waldemar G, Soininen H. Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. Neuroimage 56:1134-44 (2011).

[29] Lötjönen J, Wolz R, Koikkalainen J, Julkunen V, Thurfjell L, Lundqvist R, *et al.* Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease. Neuroimage 56: 185-96 (2011).

[30] Belkin M, Niyogi P. Laplacianeigenmaps for dimensionality reduction and data representation. Neural Computation 15:1373-96 (2003).

[31] Wolz R, Aljabar P, Hajnal JV, Lötjönen J, Rueckert D; Alzheimer's Disease Neuroimaging Initiative. Nonlinear dimensionality reduction combining MR imaging with non-imaging information. Med Image Anal 16: 819-30 (2012).

[32] Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigidregistration using free-form deformations: application to breast MR images. IEEE Trans Med Imaging 18(8):712-21 (1999).

[33] Lötjönen JM, Wolz R, Koikkalainen JR, Thurfjell L, Waldemar G, Soininen H, *et al.* Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage 49: 2352-65 (2010).

[34] Brun CC, Lepor N, Pennec X, Lee AG, Barysheva M, Madsen SK, *et al.* Mapping the regional influence of genetics on brain structure variability – a tensor-based morphometry study. NeuroImage 48: 37-49 (2009).

[35] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, *et al.* Neuroimaging Clin N Am 15: 869-77 (2005).

[36] Aguilar C, Westman E, Muehlboeck JS, Mecocci P, Vellas B, Tsolaki M, *et al.* Different multivariate techniques for automated classification of MRI data in Alzheimer's disease and mild cognitive impairment. Psychiatry Res 212(2): 89-98 (2013).

[37] Chincarini A, Bosco P, Gemme G, Esposito M, Rei L, Squarcia S, *et al.* Automatic temporal lobe atrophy assessment in prodromal AD: Data from the DESCRIPA study. Alzheimers Dement 13: 1552-5260 (2013).

[38] Falahati F, Westman E, Simmons A. Multivariate data analysis and machinelearning in Alzheimer's disease with a focus on structural magnetic resonance imaging. J Alzheimers Dis 41(3): 685-708 (2014).

[39] Clerx L, van Rossum I, Burns L, Knol DL, Scheltens P, Verhey F, *et al.* Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment. Neurobiol Aging 34(8):2003-13 (2013).

[40] Jack C, Knopman D, Jagust W, Petersen R, Weiner M, Aisen P, *et al.* Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. Lancet Neurol 12(2): 207-216 (2103).

[41] Gomar JJ, Harvey PD, Bobes-Bascaran MT, Davies P, Goldberg TE. Development and cross-validation of the UPSA short form for the performance-based functional assessment of patients with mild cognitive impairment and Alzheimer disease. Am J Geriatr Psychiatry 19(11): 915-22 (2011).

[42] Muñoz-Ruiz M, Hartikainen P, Koikkalainen J, Wolz R, Julkunen V, Niskanen E, *et al.* Structural MRI in frontotemporal dementia: Comparisons between hippocampal volumetry, tensor-based morphometry and voxel-based morphometry. PLoS One 7: e52531 (2012).

[43] Koikkalainen J, Pölönen H, Mattila J, van Gils M, Soininen H, Lötjönen J. Improved classification of Alzheimer's disease data via removal of nuisance variability. PLoS One 7: e31112 (2012).

[44] Hall A, Muñoz-Ruiz M, Mattila J, Koikkalainen J, Tsolaki M, Mecocci P, *et al.* Generalizability of the Disease State Index Prediction Model for IdentifyingPatients Progressing from Mild Cognitive Impairment to Alzheimer's Disease. J Alzheimers Dis (2014).

Received: July 14, 2014 Revised: November 28, 2014 Accepted: December 04, 2014