

# Morphological factor estimation via high-dimensional reduction

Anon.

**Abstract.** We propose a novel formulation for a morphological factor estimate based on structural MRI able to assist disease state estimate while achieving significant data reduction. We test this methodology within the context of aid to diagnostic in probable Alzheimer’s disease (AD) and prediction of conversion in Mild Cognitive Impairment (MCI). 349 subjects were included in this study. The method consisted in (a) creating a reference MRI features eigenspace using intensity and local volume change data from 149 healthy, young subjects; (b) projecting MRI data from the 75 probable AD, 75 controls (CTRL) and 49 MCI subjects in that space; (c) extracting the high-dimensional classification function; and (d) calculating a single morphological factor based on an attraction model within the discriminative feature space. We used this methodology in leave-one-out experiments to (1) confirm the improvement of this model over other approaches; (2) obtain accuracy estimates for the discrimination of probable AD from CTRL (90%) and the prediction of conversion of MCI subjects to probable AD, on average 1.5 years before the clinical assessment (79.4%); and (3) determine the resolution of the morphological factor in a scan-rescan setting (<1-4%). We estimate that the proposed formulation of the morphological factor is relevant within the context of aid to diagnostic and prediction of future clinical status in probable AD.

**Keywords:** MRI, high-dimensional data reduction, attraction model, Alzheimer’s disease, Mild Cognitive Impairment

## Introduction

A growing body of literature relates the use of machine learning methods to build classification functions from features of interest extracted from medical imaging data (e.g. **MRI**, PET) within the context of aid to clinical diagnostic of probable Alzheimer’s disease (**AD**) and/or the prediction future clinical status for individuals with Mild Cognitive Impairment (**MCI**), a putative precursor to probable AD [1-6]. These techniques have in common the reduction of large, high-dimensional image vectors into smaller features spaces, and the identification of a low-dimensional discriminating function. Authors have reported attempts to further simplify the discriminating function by calculating a single, quantitative scalar measure – the structural abnormality index score [7], the structural-functional biomarker score [8], and the disease evaluation factor [9]. While the first two rely on support vector machine analysis of a feature space composed of grey matter concentration patterns, the latter relies on linear kernel approaches to data reduction and classification of MRI appearance, defined as the combination of T1w intensity and local volume shape characteristics for all voxels within a volume of interest. We propose a different morphological factor formulation, based on this latter methodology, and extensible to other modalities and to other sources of data. We derive the formulation and estimate its efficiency within the context of aid to diagnostic in probable AD by verifying the hypothesis that it accurately describes current and future clinical status.

## Methods

### Subjects

A total of 349 subjects were included in this study. The first cohort, or **reference group**, consisted in 149 young, neurologically healthy individuals from the ICBM database [10], whose scans were used to create a reference feature space of image data.

The second cohort consisted in 150 subjects: 75 patients with a diagnosis of probable AD (**AD group**) and 75 age-matched controls (**CTRL group**) without neurological or neuropsychological deficit. The probable AD subjects were individuals with mild to moderate probable AD [11] recruited among outpatients seen at the Centro San Giovanni di Dio Fatebenefratelli – The National Center for AD (Brescia, Italy) between November 2002 and January 2005. CTRL subjects were taken from an ongoing study of the structural features of normal aging. All subjects were followed a minimum of 3 years after inclusion; this longitudinal clinical evaluation constitutes our reference diagnostic.

The third cohort consisted in 49 MCI subjects taken from a prospective project on the natural history of MCI, carried out in the same memory clinic. Inclusion and exclusion criteria are described elsewhere [12]. All MCI

patients underwent a yearly follow-up visit, consisting of complete clinical and neuropsychological examination, from 1 to 4 years after enrolment. In those individuals that converted to dementia, status was ascertained according to clinical diagnostic criteria for AD [11], subcortical vascular dementia [13], dementia with Lewy bodies [14], and fronto-temporal dementia [15]. Within the larger prospective cohort of 100 MCI patients enrolled from April 2002 to December 2006, we have selected patients retrospectively for this study based on their (a) having been followed clinically a minimum of 48 months after their baseline MR scan; and (b) having remained either stable (**MCI-S group**;  $N = 29$ ) or progressed to probable AD (**MCI-P group**;  $N = 20$ ; mean progression 1.5 yrs; SD 0.7 yrs). The 48 month longitudinal clinical evaluation constitutes our reference diagnostic.

The last subject is a healthy volunteer that acted as human quality control phantoms within the context of the pilot European ADNI project [16](**E-ADNI**). He was scanned three times at the Brescia memory clinic (scan; repeat scan, same session; rescan) on the same day.

Ethics Committees approved the study and informed consent was obtained from all participants.

## Data

The ICBM subjects from the reference group were scanned in Montreal, Canada on a Philips Gyroscan 1.5T scanner (Best, Netherlands) using a T1-weighted fast gradient echo sequence (sagittal acquisition, TR=18 ms, TE=10 ms, 1mm X 1mm X 1mm voxels, flip angle 30°). MRI data for probable AD, CTRL, MCI and E-ADNI subjects were acquired in Brescia, Italy on a single Philips Gyroscan 1.0T scanner (Best, Netherlands) using a T1-weighted fast field echo sequence (sagittal acquisition, TR=25 ms, TE=6.9 ms, 1mm X 1mm X 1,3mm voxels).

## Data processing

A cursory overview of the automated image processing methodology follows; the reader is referred to previous descriptions of the classification methodology by their authors [3] for additional details. Images from all subjects were processed in an identical fashion. Processing included intensity inhomogeneity correction [17], intensity scaling, global and linear registration [18], extraction of a pre-determined volume of interest centered on the medial temporal lobes, and nonlinear registration within the volume of interest towards a common reference target [19], and computation of determinants of the Jacobian of the deformation field [20]. In addition to that pipeline, we integrated nonlocal means denoising [21] after intensity inhomogeneity correction, transformed the intensity into z-score maps, and chose to compute the *log*-determinants of the deformation field.

## Data reduction and feature selection

The first data reduction step was to construct a feature space based on the  $N=149$  reference subjects. To this end, we used Principal Components Analysis (**PCA**) of two high-dimensional image vectors within a volume of interest centered on the medial temporal lobe to generate a low-dimensional feature space for classification: 1) the T1w MRI intensity z-score maps; and 2) the vector of log-determinants. With PCA we moved from a massive amount of data ( $2 \times 149 \times 4E10^5$  voxels) to a lower subspace model of maximum  $N-1$  dimensionality, further restricted by using only the first  $k$  eigenvectors  $\lambda$  that contribute up to a given threshold  $r$  in the description of the total variance of the system:

$$\cdot \quad (1)$$

Once the reference eigenspace has been formed, the reference group data is no longer used. We then proceeded by projecting rasterized vectors of intensity and local volume changes for subjects in the study groups into the reference space. The distribution of eigencoordinates along any principal component for a given population was assessed via quantile plots and Shapiro-Wilke statistics for normal distribution. Following the projection we used a system of supervised linear classifiers to identify the hyperplane that best separated the groups under study (e.g. CTRL vs. probable AD; MCI-S vs. MCI-P). To this end the data was first normalized to guard against variables with larger variance that might otherwise dominate the classification, We employed forward stepwise regression analysis via Wilk's  $\lambda$  method to select the set of discriminating variables  $\{\lambda_i\}$ , with  $f \ll N-1$ , forming the discriminating hyperplane.

## Comparative morphological factor construction

The morphological factor is based on the concept of distance along the restricted set of eigenvectors  $\{\lambda_i\}$ . In our image-based feature space, the distance  $d$  can be calculated in a number of different fashions.

**(a) Manhattan distance.** We propose initially the signed difference between subject eigencoordinates along the eigenvector  $\lambda_F$  and the mean of the CTRL distribution for that eigenvector, denoted  $\mu_F$ ; as this distance increases the likelihood of belonging to the CTRL group decreases:

$$(2)$$

**(b) Euclidean distance.** We propose the Euclidean distance between position  $p_i$  of each subject  $s_i$  and both CTRL and probable AD means along the restricted set of eigenvectors  $\{\lambda_F\}$  in all  $F$  directions, with  $F \ll N-1$ . As the distance to one center decreases, the distance to the second should increase. In Eq. 3 we demonstrate the distance to the mean of the probable AD group:

$$(3)$$

**(c) Weighted distance.** It is possible to weigh each eigenvector by an associated measure of significance, for example Wilk's  $\lambda$  from the stepwise regression analysis [9] or a factor derived from univariate t-tests. While the Wilk's  $\lambda$  is trivially obtained from the regression analysis, an univariate weight such as the Koikkalainen factor formulation [22] entails performing a t-test comparing the group eigencoordinate distributions (e.g. CTRL vs. probable AD; MCI-S vs. MCI-P) for each eigenvector of the restricted set, resulting in the p-value  $p(\lambda_F)$  for that distribution; from these p-values the significance weight  $S_F$  was calculated:

$$(4)$$

The significance increases as the differences between the CTRL and AD groups grows, and reaches zero when there are no statistically significant difference (at the  $p=0.05$  level) between both distributions.

The resulting weighted distance  $D_i$  combines the aforementioned distances (Manhattan, Euclidean) with a weight  $S_F$  (either Wilk's  $\lambda$ , or Koikkaalinen factor) over all eigenvectors  $F$  from the restricted set  $\{\lambda_F\}$  as follows:

$$(5)$$

**(d) Gravitational model.** As our final formulation, we extend the principle of image-based distance to the context of an attraction field that follows Newton's Law of Universal Gravitation, whereby any two elements of mass  $m$  within the feature space will exert upon one another an attractive force that will vary proportionally to the inverse of the square of the distance between them. In our context the force exerted by one group (e.g. CTRL) decreases as the distance between a subject and the center of mass of the CTRL group grows, while the force exerted by the second group (e.g. probable AD) increases as distance decreases between the same subject and the second group's center of mass. In a multiple group scenario, the calculated combined force serves as a quantitative measure of the likelihood of belonging to one of the groups.

In such a classical formulation the force between any subject  $s_i$  with mass  $m_i$ , to the centers of mass of e.g. the CTRL group ( $CM_{CTRL}$ ) and the AD group ( $CM_{AD}$ ), is expressed as:

$$(4)$$

with

$$(3)$$

being the formulation for the centers of mass calculations, where  $M$  is the total mass for all subjects in the group,  $m_i$  their individual masses, and  $p_i$  their individual positions in feature space as derived in the previous section. The distance metric that can be used can be anyone of the aforementioned distances; for the purposes of the current study, the Euclidean distance as formulated in Eq. 3 was employed.

We chose to retain the concept of “mass” even though it has no real bearing within the present context of an image-based feature space. It could be replaced with different information regarding individuals in the groups, for example Braak histopathological staging [23]. Alternatively one can vary the specificity and sensitivity of the attraction field by increasing the “mass” of subjects in one of the groups (e.g. CTRL or probable AD). For our purposes however we set the mass of each subject to unity, and, further, for equal considerations of simplicity, we set the gravitational constant  $G$  also to unity.

Statistics and measurements were computed using the MATLAB Statistics Toolbox (The MathWorks, Natick, MA).

## Experiments

Once the reference space was created, all of the experiments that we conducted were performed in a leave-one-out fashion whereby one subject from the study groups was temporarily removed, allowing for an independent estimate of the low-dimensional discriminant function and the calculation of the eigendistribution means and centers of mass. Only then was the left-out subject entered in the system and its morphological factor computed. The final results consist in the comparison of the independently acquired morphological factors for each subject. We ran two distinct experiments: (a) determination of the relative *accuracies* of each distance formulation (Manhattan, Euclidean, Weighted Distance (Wilk’s  $\lambda$ ), Weighted Distance (Koikkalainen), Gravitational model) at the discrimination of CTRL vs. probable AD; (b) determination of the *accuracy* of the best distance formulation at the discrimination of MCI-S vs. MCI-P; and (c) determination of the *resolution* of the best distance formulation based on the CTRL vs. probable AD discriminant function using the E-ADNI scan-rescan dataset.

## Results

### Demographics

There were no statistically significant differences for age between the 75 probable AD (mean = 73,3 yrs; SD = 8,4 yrs) and 75 NC individuals (mean = 73,3 yrs; SD = 4,6 yrs) (Student’s T test, DF = 148,  $P > 0.05$ ). There was a statistical difference for age (Student’s T test, DF = 47,  $p = 0.001$ ) between the MCI-S (mean = 74.2 yrs; SD = 6.4 yrs) and MCI-P groups (mean = 63.6 yrs; SD = 14.2 yrs).

### Data processing and feature selection

We set the variance ratio  $r$  (see eq. (1)) to 0.997, resulting in a reference PCA model composed of 256 intensity and local volume change eigenvectors. We proceeded with forward stepwise regression analysis using Wilk’s  $\lambda$  method ( $P$ -to-enter = 0.005) to select the discriminating variables forming the hyperplane separating each group (e.g. CTRL from probable AD; MCI-S from MCI-P). This was performed in a leave-one-out fashion to eliminate over-learning of the dataset. The median number of eigenvectors  $\lambda_F$  retained in the discriminating function for either CTRL vs. probable AD or MCI-S vs. MCI-P was four.

### Morphological factor calculation

**Table I** displays the different accuracies obtained for the five different formulations for the morphological factor at the task of discriminating CTRL vs. probable AD (leave-one-out). The Gravitational model’s accuracy was 90%, superior to the Weighted Distance models.

	Gravity model	Weighted distance		Euclidean	Manhattan
		Koikkalainen	Wilk’s $\lambda$		
<b>Accuracy</b>	0.90	0.86	0.85	0.73	0.78

**Table I – Model results**

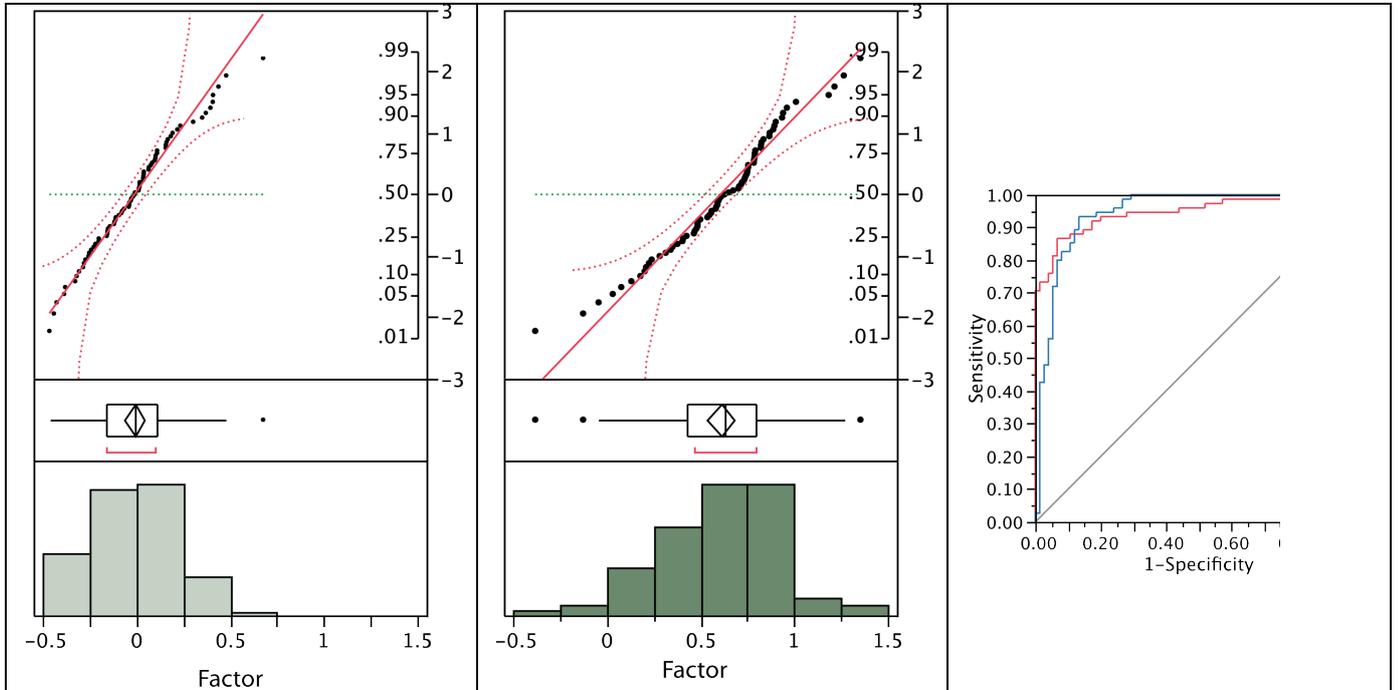
Using the Gravitational model, we report the results for the morphological factor for the CTRL vs. probable AD experiment and the MCI-S vs. MCI-P experiment in Table II. The distributions of morphological factors for all groups, alongside quantile plots to assess normality (CTRL and probable AD groups) are shown in Figures 1 and 2.

The Receiver Operating Characteristic (ROC) curve for the task of discriminating CTRL from probable AD shows the trade-offs possible in sensitivity and specificity (Figure 1; Right). The Area under the ROC curve was 0.9444. At the 90% accuracy point (135/150), specificity was 87.5% and sensitivity 92.9%.

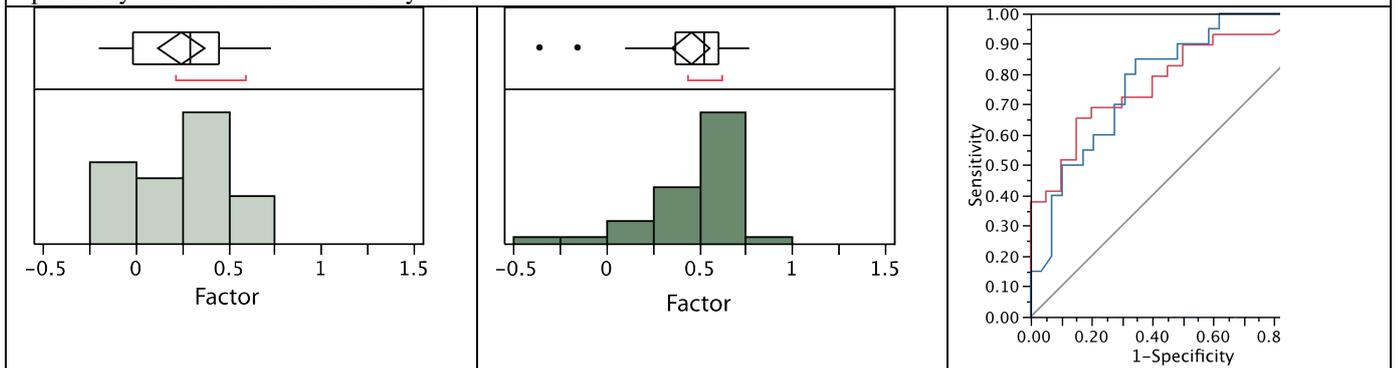
CTRL	AD	MCI Stable	MCI Progressed
------	----	------------	----------------

<b>N</b>	75	75	29	20
<b>Mean</b>	0.61	-0.01	0.45	0.24
<b>Std Dev</b>	0.32	0.23	0.26	0.27
<b>Std Err Mean</b>	0.04	0.03	0.05	0.06
<b>Upper 95% Mean</b>	0.68	0.04	0.55	0.37
<b>Lower 95% Mean</b>	0.53	-0.06	0.35	0.12

**Table II - Morphological factor results**



**FIGURE 1 – (LEFT AND MIDDLE)** Distributions of morphological factors for the CTRL (left) and probable AD groups (middle) alongside quantile plots based on the Gravitational model (see Eq. 4). **(RIGHT)** Receiver operating characteristic curve (ROC) for the morphological factor displaying the trade-offs between sensitivity and specificity at the task of discriminating CTRL vs. probable AD. The Area under the ROC curve was 0.9444. At the 90% accuracy point (135/150), specificity was 87.5% and sensitivity 92.9%.



**FIGURE 2 – (LEFT AND MIDDLE)** Similar distributions of morphological factors for the MCI-S (left) and MCI-P groups (middle). **(RIGHT)** ROC for the discrimination of MCI-S and MCI-P. The Area under the ROC curve was 0.7940. At 72.3% accuracy, specificity was 62%, and sensitivity 75%.

With the Gravitational model we computed the ROC curve for the discrimination of MCI-S from MCI-P (Figure 2; Right). The Area under the ROC curve was 0.7940. At 72.3% accuracy, specificity was 62%, and sensitivity 75%.

We then computed the morphological factor for the E-ADNI human phantom volunteer, using the CTRL and probable AD cohorts as a training group (for the determination of the discriminating function). Using the Gravitational model, the average factor value was -0.4, or 4 standard deviations away from the mean of the CTRL

distribution, with an average difference in scan-rescan factor of 4%. Notably, the morphological index obtained via a weighted distance method (Koikkalainen factor) had an average difference in scan-rescan factor of less than 1%.

## Discussion

The gravitational model approach constitutes a novel development in the strategies towards obtaining a single quantitative factor from data reduction and machine learning of very high-dimensional MRI input data towards discrimination of individual subjects. Its inherent flexibility makes multi-group comparisons trivial, alongside the introduction of other sources of data. Its performance compares favorably to other results in the MRI literature within the context of discriminating CTRL vs. probable AD [2]. As a single dimensional scalar, the morphological factor metric achieves strong accuracy (90%), especially when compared to a multidimensional discrimination function on the same data set (e.g. 92% as reported in [3]). It is also a strong result when put within the *clinical* context of discriminating CTRL vs. probable AD, where inclusion evaluations are reportedly 78% accurate (against longitudinal evaluation and final histopathological diagnostic). While lower, accuracy figures for the prediction of progression to probable AD in the MCI cohort (on average, 1.5 years before clinical diagnostic) are also strong, and compare favorably to published results on MRI data [4, 5]. A study comparing these approaches (e.g. within a mono-centric setting, such as the Open Access Series of Imaging Studies [24] or multi-centric setting such as the Alzheimer’s Disease Neuroimaging Initiative [25]) would be worthwhile.

The paper uses the leave-one-out approach to feature selection (stepwise regression analysis), which allows a correct generalization of the morphological factor as it is not tested on the same data.

Clinical interpretation of changes in image features associated with changes in the morphological factor should provide insight into the development of AD and would need to be compared to existing results from voxel-based morphometry studies, structural studies (e.g. hippocampal and entorhinal atrophy) and histopathological confirmation studies. Overall, we speculate that the specific patterns of intensity and local volume change differences result from different levels of advanced extra-cellular plaque formation, neurofibrillary tangles accumulation and other pathological processes between CTRL and probable AD, and between stable and progressing MCI. With regards to the features employed in this method, the differences in local volume changes should mirror the changes noticed in other reports, such as visual assessment [26], while differences in grey-level might reflect the intensity of neuronal loss induced by the neuropathological changes [27], which precede volume loss as visualized on MRI. Such an evaluation however is beyond the scope of this manuscript.

There are a number of limitations in this study. One pertains to the fact that the MRI images for the probable AD subjects were acquired at the time of diagnosis; therefore, some of the patients have had AD for a number of years. In turn, this implies that extensive neurodegeneration has taken place at this point, and should artificially facilitate the discrimination with CTRL. However, the fact that the latter were age-matched, and the fact that the results in the MCI cohort remain significant, alleviate part of this concern. It would be useful to assess if the morphological factor correlates with different indices of disease severity, cognitive deficits or other biomarkers. Neuropathological confirmation is also required to replace the clinical evaluation as a gold standard. Finally, the patterns of abnormalities that can be found by the method are restricted to a space which is built from healthy, young controls. It is not the optimal space to describe normal aging and/or AD-related variability. However, it does tend to maximize the distance between both groups, as we noticed from building a few reference spaces in a N-fold validation of the CTRL/probable AD groups that achieved lower accuracies.

We estimate that the proposed formulation of the morphological factor is relevant within the context of aid to diagnostic and prediction of future clinical status in probable AD.

## Acknowledgments

We thank the International Consortium for Brain Mapping and the pilot European ADNI study for data. This project was supported by funding from the Ministère du Développement Économique, de l’Innovation et de l’Exportation du Québec.

## References

1. Lao, Z., et al., Morphological classification of brains via high-dimensional shape transformations and machine learning methods. *Neuroimage*, 2004. **21**(1): p. 46-57.
2. Davatzikos, C., et al., Detection of prodromal Alzheimer’s disease via pattern classification of magnetic resonance imaging. *Neurobiol Aging*, 2008. **29**(4): p. 514-23.
3. Duchesne, S., et al., MRI-Based Automated Computer Classification of Probable AD Versus Normal Controls. *IEEE Trans Med Imaging*, 2008. **27**(4): p. 509-20.

4. Duchesne, S., et al., Amnesic MCI future clinical status prediction using baseline MRI features. *Neurobiol Aging*, 2008. **Accepted for publication.**
5. Fan, Y., et al., Spatial patterns of brain atrophy in MCI patients, identified via high-dimensional pattern classification, predict subsequent cognitive decline. *Neuroimage*, 2008. **39**(4): p. 1731-1743.
6. Kloppel, S., et al., Automatic classification of MR scans in Alzheimer's disease. *Brain*, 2008.
7. Vemuri, P., et al., Alzheimer's disease diagnosis in individual subjects using structural MR images: validation studies. *Neuroimage*, 2008. **39**(3): p. 1186-97.
8. Fan, Y., et al., Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. *Neuroimage*, 2008. **41**(2): p. 277-85.
9. Duchesne, S. Quantitative evaluation of Alzheimer's disease. in *SPIE - Medical Imaging*. 2009. Orlando, FL: SPIE Society.
10. Mazziotta, J.C., et al., A probabilistic atlas of the human brain: theory and rationale for its development. The International Consortium for Brain Mapping (ICBM). *Neuroimage*, 1995. **2**(2): p. 89-101.
11. McKhann, G., et al., Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 1984. **34**(7): p. 939-44.
12. Rossi, R., et al., Clinical and neuropsychological features associated with structural imaging patterns in patients with mild cognitive impairment. *Dement Geriatr Cogn Disord*, 2007. **23**(3): p. 175-83.
13. Erkinjuntti, T., et al., Research criteria for subcortical vascular dementia in clinical trials. *J Neural Transm Suppl*, 2000. **59**: p. 23-30.
14. McKeith, I.G., et al., Prospective validation of consensus criteria for the diagnosis of dementia with Lewy bodies. *Neurology*, 2000. **54**(5): p. 1050-8.
15. Knopman, D.S., et al., Antemortem diagnosis of frontotemporal lobar degeneration. *Ann Neurol*, 2005. **57**(4): p. 480-8.
16. Frisoni, G.B., et al., The pilot European Alzheimer's Disease Neuroimaging Initiative of the European Alzheimer's Disease Consortium. *Alzheimers Dement*, 2008. **4**(4): p. 255-64.
17. Sled, J.G., A.P. Zijdenbos, and A.C. Evans, A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data. *IEEE Transactions on Medical Imaging*, 1998. **17**: p. 87-97.
18. Collins, D.L., et al., Automatic 3D Intersubject Registration of MR Volumetric Data in Standardized Talairach Space. *Journal of Computer Assisted Tomography*, 1994. **18**: p. 192--205.
19. Collins, D.L. and A.C. Evans, ANIMAL: Validation and Applications of Non-linear Registration Based Segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 1997. **11**(8): p. 1271--1294.
20. Chung, M.K., et al., A unified statistical approach to deformation-based morphometry. *NeuroImage*, 2001. **14**(3): p. 595-606.
21. Coupe, P., et al., An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. *IEEE Trans Med Imaging*, 2008. **27**(4): p. 425-41.
22. Koikkalainen, J., et al. Estimation of disease state using statistical information from medical imaging data. in *Medical Image Computing and Computer Assisted Intervention - From statistical atlases to personalized models workshop*. 2006. Copenhagen, Denmark: MICCAI Society.
23. Braak, H. and E. Braak, Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol*, 1991. **82**(4): p. 239-59.
24. Marcus, D.S., et al., Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young, Middle Aged, Nondemented, and Demented Older Adults. *J Cogn Neurosci*, 2007. **19**(9): p. 1498-507.
25. Mueller, S.G., et al., Ways toward an early diagnosis in Alzheimer's disease: The Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement*, 2005. **1**(1): p. 55-66.
26. Wahlund, L.O., et al., Visual rating and volumetry of the medial temporal lobe on magnetic resonance imaging in dementia: a comparative study. *J Neurol Neurosurg Psychiatry*, 2000. **69**(5): p. 630-5.
27. Wahlund, L.O. and K. Blennow, Cerebrospinal fluid biomarkers for disease stage and intensity in cognitively impaired patients. *Neurosci Lett*, 2003. **339**(2): p. 99-102.