Multi-stage biomarker models for progression estimation in Alzheimer's disease

Alexander Schmidt-Richberg[1], Ricardo Guerrero[1], Christian Ledig[1], Helena Molina-Abril[2], Alejandro F. Frangi[2], Daniel Rueckert[1], and on behalf of the Alzheimers Disease Neuroimaging Initiative

Biomedical Image Analysis Group, Imperial College London, London
Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), University of Sheffield, UK
a.schmidt-richberg@imperial.ac.uk,

Abstract. The estimation of disease progression in Alzheimer's disease (AD) based on a vector of quantitative biomarkers is of high interest to clinicians, patients, and biomedical researchers alike. In this work, quantile regression is employed to learn statistical models describing the evolution of such biomarkers. Two separate models are constructed using 1) subjects that progress from a cognitively normal (CN) stage to mild cognitive impairment (MCI) and 2) subjects that progress from MCI to AD during the observation window of a longitudinal study. These models are then automatically combined to develop a multi-stage disease progression model for the whole disease course. A probabilistic approach is derived to estimate the current disease progress (DP) and the disease progression rate (DPR) of a given individual by fitting any acquired biomarkers to these models. A particular strength of this method is that it is applicable even if individual biomarker measurements are missing for the subject. Employing cognitive scores and image-based biomarkers, the presented method is used to estimate DP and DPR for subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Further, the potential use of these values as features for different classification tasks is demonstrated. For example, accuracy of 64% is reached for CN vs. MCI vs. AD classification.

1 Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative condition and the most common form of dementia. Patients that show first symptoms like general memory loss are usually diagnosed as suffering from Mild Cognitive Impairment (MCI), an early stage of dementia. Later, throughout the disease, these symptoms are followed by behavioural changes and further cognitive and functional decline. Patients become less able to perform simple tasks and increasingly depend on carers' support. However, an objective staging of the disease is non-trivial due to the considerable variability in the patient's age at disease onset and the individual rate of progression [3]. Estimating the current disease severity

and the future rate of progression is of interest for patients and caregivers but also has potential to improve clinical trials as more homogeneous study groups can be recruited.
Most existing approaches with this aim employ the results of cognitive tests like the Clinical Dementia Rating – Sum of Boxes (CDR-SB) as biomarkers to characterise disease progression. Many of those methods (see e.g. [5] for a detailed overview) use Markov transition models and Cox proportional hazard models to estimate disease progression. However, as detailed in [5], the main limitations of these methods are "the use of a limited number of health states to capture events related to disease progression over time", and the fact that "a single symptom, such as cognition, is not able to characterise AD progression". Therefore, additional meaningful biomarkers that better describe anatomical changes can be quantified from imaging data. For example, [4, 12] use an event-based model to determine the order in which CSF-, image- and cognition-based biomarkers become abnormal and then employ this information to assign a subject to one of several discrete disease stages.
As modelling disease progression by a number of discrete stages is a strong simplification, some approaches have been developed that acknowledge the course of disease as a continuous process. For example, Yang et al. [10] assume an exponential-shaped trajectory of the ADAS score. The authors then estimate a *time shift* γ indicating the disease progress of a subject by fitting its ADAS scores to this curve. Similarly, Delor et al. [2] compute a *disease onset time* by adjusting subjects according to their CDR-SB score.
The approach presented in this work builds upon these methods. Here, quantile regression is used to estimate typical trajectories of clinical biomarkers (see Sec. 2). In detail, two models are trained, one for the transition from CN to MCI and one for the MCI-to-AD conversion. These models are then combined to a multi-stage model for the whole course of the disease. Thereafter, a probabilistic model is derived that allows the estimation of a subjects current disease progress and rate of progression based on measured biomarker values. The approach is flexible with regard to the considered biomarkers, which can be based, for example, on cognitive scores, neuroimaging, or both. Moreover, missing measurements are handled in a natural way, this means, the approach can be employed even if the set of observed biomarkers is incomplete. The proposed disease progress estimation is evaluated in Sec. 3 using clinical data. Its applicability for different classification tasks is demonstrated at the end of this section.

2 Methods

To model disease progression, the existence of a set of biomarker values y_{sv}^b acquired from multiple subjects $s \in S = \{1, \dots, n_S\}$ during multiple visits $v \in V_s$ is assumed. Here, $b \in B_{sv}$ denotes the index of the biomarker. Each biomarker vector is associated with the time $t_{sv} \in T$ of acquisition, measured in days after the first (baseline) visit, as well as the diagnosis d_{sv} that was given during each visit. The number of visits can vary for each subject, $V_s \subseteq V = \{1, \dots, n_V\}$. Also, the

biomarkers acquired at each visit might differ, such that $B_{sv} \subseteq B = \{1, \ldots, n_B\}$. In the training phase of the presented method, characteristic trajectories of biomarkers in the course of disease progression are learned based on a number of training subjects (Sec. 2.1). These models are then employed in the test phase to estimate how far and how fast test subjects have progressed along the disease trajectory (Sec. 2.2).

2.1 Model learning

Aim of the model training phase is to learn the temporal trajectory of biomarker evolution throughout the disease by determining the probability that a certain biomarker b has a value y^b at a specified time point. More technically, each measured biomarker value y^b_{sv} is understood as an observation of a response variable Y^b at a *disease progress (DP)* $p_{sv} \in \mathbb{R}$ (the explanatory variable or covariate). The conditional distribution of Y^b given p is then denoted by $f_{Y^b}(y|p)$.
A *disease progression model* $\mathcal{M}(p)$ comprises the distributions of all biomarkers in B on a domain $P \subset \mathbb{R}$, such that $\mathcal{M}(p) = \{\mathcal{M}^1(p), \ldots, \mathcal{M}^{n_B}(p)\}$ with $\mathcal{M}^b(p) := f_{Y^b}(y|p)$ for $p \in P$. Another way of representing the model is by its q-quantile functions $y^b_q(p)$, which can be derived directly from $f_{Y^b}(y|p)$ (for example, the median trajectory is denoted by $y^b_{0.5}(p)$).
The learning of a model consists of three main steps. First, the training subjects have to be temporally aligned to establish correspondences between the time points of observation. Progression models are then estimated using quantile regression to learn the probability distributions $f_{Y^b}(y|p)$. Since temporal alignment is based on the point of conversion from either CN to MCI or MCI to AD, two separate models are learned. These two models are then combined to a multi-stage progression model.

Temporal alignment of the training data Temporal alignment aims at associating the time points t_{sv} of biomarker acquisition to the corresponding DP p_{sv}. In detail, the goal is to find a strictly monotonically increasing *time warp* function $\tau(t)$ that maps the subject-specific acquisition time $t_{sv} \in T$ to the population-based disease progress $p_{sv} \in P$, such that $p_{sv} = \tau(t_{sv})$. During model training, the time point t^0_s at which the clinical diagnosis changes and thus indicates transition to a more severe disease state is set to $p = 0$, that means

$$p_{sv} = \tau(t^0_s; t_{sv}) := t_{sv} - t^0_s \ . \tag{1}$$

For this reason, a specific model is trained for each transition phase (here, CN-to-MCI and MCI-to-AD). To identify t^0_s, the visits v^* and v^{**} with the last CN (MCI) and the first MCI (AD) diagnosis are determined. The time point of conversion is then assumed to be the average of these two visits, i.e. $t^0_s := 0.5 \cdot (t_{sv^*} + t_{sv^{**}})$.

Learning disease progression model The conditional distributions f_{Y^b} are learned independently for each biomarker using quantile regression via *vector*

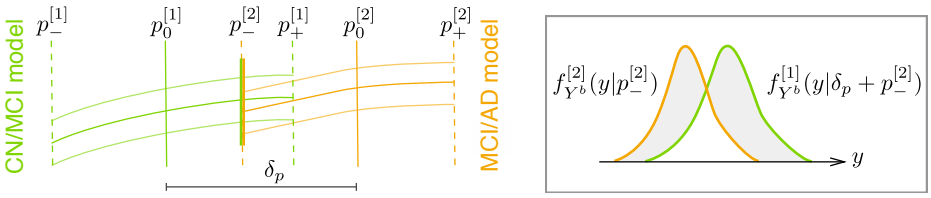

Fig. 1. Approach for automatically determining the optimal model offset δ_p.

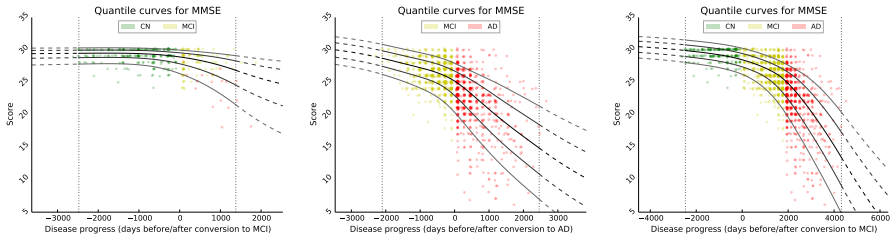

Fig. 2. Example of the model composition. First, separate models are trained for CN-to-MCI and MCI-to-AD converters. An optimal offset between these models is then automatically determined and model training is repeated with the whole set of samples.

generalised additive models (VGAMs) [11]. In contrast to logistic or exponential regression [10], VGAMs do not depend on prior assumptions on the functional form for each predictor variable other than their smoothness. However, the domain P of $\mathcal{M}^b(p)$ is limited to the progress interval contained in the sample set. This means P is given by $P = [p_-, p_+]$, with $p_- := \min_{s,v}(p_{sv})$ and $p_+ := \max_{s,v}(p_{sv})$ being the earliest and latest observed DP, respectively. Therefore, the models are extrapolated by a linear extension of the underlying predictor functions (see [11] for details) and $P = \mathbb{R}$ is assumed in the following.

Model composition To combine the CN/MCI and MCI/AD models, an optimal offset δ_p is determined by optimising the similarity of the models in the overlapping region. Given δ_p, the end point $p_+^{[1]}$ of the CN/MCI model $\mathcal{M}^{[1]}$ corresponds to $p_+^{[1]} - \delta_p$ in the MCI/AD model $\mathcal{M}^{[2]}$. Similarly, $\mathcal{M}^{[2]}(p_-^{[2]})$ corresponds to $\mathcal{M}^{[1]}(\delta_p + p_-^{[2]})$ (cf. Fig. 1). The quality of the fit is quantified by

$$\hat{\delta}_p := \underset{\delta_p}{\operatorname{argmax}} \frac{1}{2} \left[\left(\mathcal{M}^{[1]}(p_+^{[1]}) - \mathcal{M}^{[2]}(p_+^{[1]} - \delta_p) \right) + \left(\mathcal{M}^{[1]}(\delta_p + p_-^{[2]}) - \mathcal{M}^{[2]}(p_-^{[2]}) \right) \right]$$

with $\mathcal{M}^{[1]}(p) - \mathcal{M}^{[2]}(q) = \frac{1}{|B|} \sum_{b \in B} \int f_{Y^b}^{[1]}(y|p) - f_{Y^b}^{[2]}(y|q)\, dy$ being the area between the corresponding density functions (averaged over all biomarkers). After determining $\hat{\delta}_p$, the multi-stage model is retrained using the full set of samples with $p = 0$ defined as the point of conversion from CN to MCI (see Fig. 2).

2.2 Progress estimation

Once the disease progression model is built, the aim is to estimate the progress of any given subject. However, the point of conversion t_s^0 is usually unknown and thus Eq. (1) cannot be employed. Progress estimation is accomplished by finding the most likely time warp $\tau(t)$ that optimally fits the evolution of the biomarkers, as measured from the patient, into the progression model $\mathcal{M}$.
Let $\boldsymbol{t}_s = (t_{s1}, \dots, t_{sn_V})^T$ be the vector containing the time points of all visits of subject s and $\tau(\boldsymbol{t}_s) = (\tau(t_{s1}), \dots, \tau(t_{sn_V}))^T$. Let further $\boldsymbol{y}_s = (\boldsymbol{y}_{sv})_{v \in V_s}$ be the biomarker vector measured for s, with $\boldsymbol{y}_{sv} = (y_{sv}^b)_{b \in B_{sv}}$ denoting the values acquired at visit v. Based on $\boldsymbol{t}_s$, the most probable time warp $\hat{\tau}_s$ given $\boldsymbol{y}_s$ is determined by maximising the logarithm of the likelihood function $\mathcal{L}(\tau(\boldsymbol{t}_s) \,|\, \boldsymbol{y}_s)$. This means

$$\hat{\tau}_s := \underset{\tau}{\operatorname{argmax}} \ \log \mathcal{L}(\tau(\boldsymbol{t}_s) \,|\, \boldsymbol{y}_s) = \underset{\tau}{\operatorname{argmax}} \ \log f_{\boldsymbol{Y}}(\boldsymbol{y}_s \,|\, \tau(\boldsymbol{t}_s)) \tag{2}$$

with $\boldsymbol{Y} = (Y^1, \dots, Y^{n_B})$. The joint probability of all observations y_{sv}^b is then

$$f_{\boldsymbol{Y}}(\boldsymbol{y}_s \,|\, \tau(\boldsymbol{t}_s)) = \prod_{v \in V_s} f_{\boldsymbol{Y}}(\boldsymbol{y}_{sv} \,|\, \tau(t_{sv})) = \prod_{v \in V_s} \prod_{b \in B_{sv}} f_{Y^b}(y_{sv}^b \,|\, \tau(t_{sv})) \ .$$

whereat all biomarker observations are assumed to be independent of each other. A simple *translational time warp* parameterisation is given by

$$\tau(p^0; t) := p^0 + t \ . \tag{3}$$

Here, the *disease progress* (DP) $p^0 \in \mathbb{R}$ is an offset that indicates how far the subject has progressed in the course of disease at the time point of the first visit. However, this simple model cannot accommodate for different rates of progression, which are known to exist between subjects [3]. If $|V_s| > 1$, the extended *affine time warp* definition

$$\tau(p^0, r; t) := p^0 + rt \tag{4}$$

can be employed, where $r \in \mathbb{R}^+$ is a scaling factor indicating the *disease progression rate* (DPR). The optimal values $\hat{p}_s^0$ and $\hat{r}_s$ for DP and DPR are determined by maximising Eq. (2) over all p^0 and r, i.e. $\hat{\tau}_s(\cdot) \mathrel{\hat{=}} \tau(\hat{p}_s^0, \hat{r}; \cdot)$. In general, more complex time warps definitions are possible.

3 Experiments and results

This section evaluates the presented approach using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI, Sec. 3.1). First, in Sec. 3.2, progression models are trained for all available biomarkers. These models are then employed to estimate DP and DPR. The applicability of these values for different classification tasks is shown in Sec. 3.3.

3.1 Materials

For this study, all subjects enrolled in either ADNI1, ADNIGO or ADNI2 (up to an acquisition date of 30/01/2014) were considered. For model training, all subjects were selected that converted either from CN to MCI (60 subjects) or from MCI to AD (248 subjects). The number of available biomarkers can vary between the training subjects (see below). To obtain consistent results in all experiments, the test set consists of the subjects with a stable diagnosis for which all biomarkers are present at baseline (bl), month 12 (m12) and month 24 (m24) visit. In total, these are 158 (88 male, 70 female) subjects classified as cognitive normal (CN), 90 (50 male, 40 female) patients with early MCI (EMCI), 94 (63 male, 31 female) with late MCI (LMCI), and 88 (45 male, 43 female) patients diagnosed with AD.
The set B of biomarkers considered in this work consists of cognitive scores and image-based features as detailed in the following. Biomarker values are corrected for age using a linear regression on the baseline samples of all control subjects.

Cognitive scores $B^{\mathbf{cog}}$: A number of cognitive tests are performed by subjects participating in the ADNI study at each visit. The achieved scores are used as biomarkers. These tests include the Minimental State Examination (MMSE), the Alzheimer's Disease Assessment Scale (ADAS 11 and ADAS 13), the Functional Activities Questionnaire (FAQ), the Clinical Dementia Rating – Sum of Boxes (CDR-SB) and the Rey Auditory Verbal Learning Test (RAVLT). Not every test score is available for each subject and visit, such that the absolute number of available training samples varies between 393 to 399 samples from 60 subjects for $\mathcal{M}^{[1]}$, and 1452 to 1480 samples from 248 subjects for $\mathcal{M}^{[2]}$.

Volumes of brain structures $B^{\mathbf{vol}}$: Further, the volumes of 35 distinct brain structures are used as biomarkers. For this, MR scans are first automatically segmented into 134 regions using the whole brain segmentation proposed by [8], which is based on multi-atlas label propagation with expectation-maximisation based refinement (MALPEM). Here, brain atlases from the MICCAI 2012 Grand Challenge on Multi-Atlas Labeling[1] are employed. Corresponding manual expert segmentations are provided by Neuromorphometrics, Inc.[2] under academic subscription. The 30 atlas segmentations are transformed to an unsegmented scan and fused into a consensus probabilistic segmentation estimate using a local weighting approach. Subsequently, all 134 probabilistic label estimates are further refined using image intensity information.
To reduce the total number of models, left and right cortex are fused to single structures, resulting in 35 distinct anatomical regions. For procedural reasons, only segmentations for images acquired before 20/11/2013 were available, such that the total number of training samples for each structure is 219 (59 subjects) and 955 (247 subjects) for the two models.

[1] `https://masi.vuse.vanderbilt.edu/workshop2012`
[2] `http://Neuromorphometrics.com`

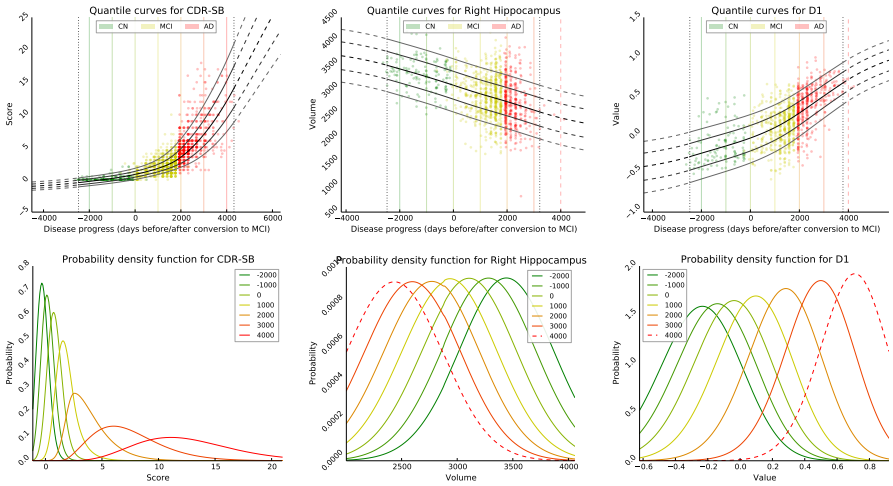

Fig. 3. Examples for the learned multi-stage disease progression models.

Biomarkers derived from manifold learning $B^{\mathbf{ml}}$: Features obtained from MR images using manifold learning (ML) have been shown to contain valuable information about disease severity and progression [6]. The main idea of ML is to find a meaningful, low-dimensional representation of a high-dimensional feature space, such that similar scans also have similar coordinates in the low-dimensional manifold. This is achieved in three steps. First, the image regions that are most relevant with regard to information about disease state are automatically learned using sparse regression as in [6]. To compensate for varying intensity values in the images caused by different scanners and acquisition protocols, local binary patterns (LBP) are computed in a 26-connected neighbourhood for voxels within these regions and used as features in the high-dimensional space. The manifold is then learned using Laplacian eigenmaps. The local geometry is determined via a sparse similarity graph, built using the sum of squared differences (SSD) as similarity measure. Connections in the graph are made between the k nearest neighbours, with the additional constraint that an instance can only be connected to one instance per subject.
The manifold coordinates are computed for all training subjects with at least 5 visits, which results in 185 samples from 41 subjects available to train the CN/MCI model and 859 samples (155 subjects) for the MCI/AD model. The manifold dimension is chosen to be $d = 20$, that means 20 features are obtained per subject per visit and denoted as D1 to D20.

3.2 Model learning and progress estimation

Progression models are trained for all 60 biomarkers. CDRSB and FAQ models could not be built for the CN-to-MCI transition because the majority of samples

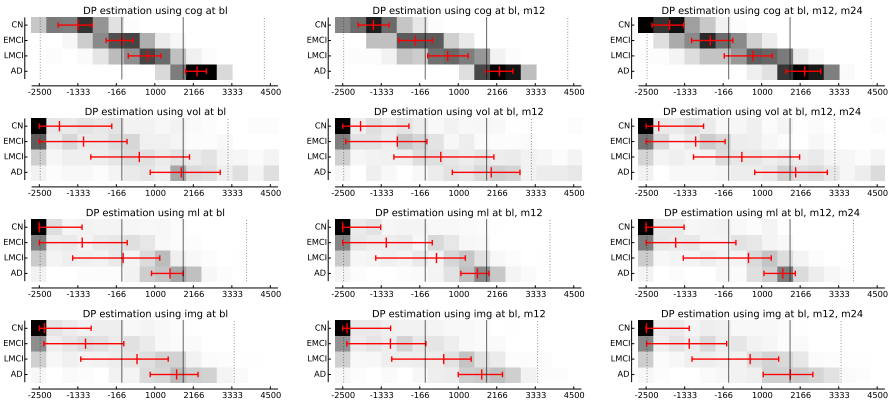

Fig. 4. Visualisation of the disease progress (DP) estimated with different biomarker settings (columns: number of visits; rows: biomarker sets). The distribution of DPs is indicated by the grey values. The x-axis gives the DP with solid and dashed lines at the points of conversion at $p = 0$ and $p = \delta_p$, and the mean model range p_- and p_+. The red bars show the median and 25th/75th percentile of the estimated DPs.

is clustered at $y = 0$. These biomarkers were therefore omitted for determining the model offset. Based on the remaining biomarkers, an $\hat{\delta}_p = 1860$ days is determined. Composed models are then retrained based on all samples and example results visualised in Fig. 3. Using the composed models, disease progression of all test subjects is estimated as proposed in Sec. 2.2. Plausibility of the estimated DPs is evaluated with regard to their ability do differentiate between the different diagnoses, that means to what extent an ordering of CN $<$ EMCI $<$ LMCI $<$ AD is achieved. To this end, the disease progress $\hat{p}_s^0$ is estimated on the search space $[-2500, 4500]$ for all CN, EMCI, LMCI and AD subjects in the test set using several biomarker configurations. On the one hand, different sets of biomarkers B^{est} are considered for estimation: B^{cog}, B^{vol}, B^{ml}, as well as all imaging-based biomarkers $B^{\text{img}} := B^{\text{vol}} \cup B^{\text{ml}}$. On the other hand, biomarkers from one (baseline), two (baseline and m12) and three (baseline, m12 and m24) visits V^{est} are employed. The distribution of the estimated DPs depending on the diagnosis is visualised in Fig. 4.

3.3 Application: Classification

Image-based classification of disease stage One of the main research topics of image-based analysis of Alzheimer's disease is the classification of subjects according to their diagnosis based on structural MR images. The high interest is highlighted by a classification challenge held in the course of MICCAI 2014[3]. The estimated DPs are therefore employed as single features to distinguish between CN and AD, CN and MCI, MCI and AD, and all three classes simultaneously.

[3] `http://caddementia.grand-challenge.org`

Table 1. Results for the image-based classification of subjects according to their diagnosis. Different imaging biomarkers acquired at a different number of visits are compared. For each test, the accuracies (ACC) of a 10-fold cross validation are given.

Biomarkers B^{est}	Visits V^{est}	CN/AD ACC	CN/MCI ACC	MCI/AD ACC	CN/MCI/AD ACC
B^{ml}	{bl}	0.89	0.71	0.82	0.63
	{bl, m12}	0.91	0.71	0.82	0.63
	{bl, m12, m24}	0.91	0.69	0.82	0.61
B^{vol}	{bl}	0.83	0.68	0.80	0.59
	{bl, m12}	0.86	0.68	0.78	0.58
	{bl, m12, m24}	0.85	0.69	0.78	0.58
B^{img}	{bl}	0.88	0.72	0.81	0.63
	{bl, m12}	0.90	0.71	0.83	0.64
	{bl, m12, m24}	0.90	0.71	0.83	0.64

Support Vector Machines (SVMs) are used as classifiers. Corresponding to Sec. 3.2, DPs are estimated using the translational time warp definition (3) based on different biomarker settings and compared to each other. The cognitive scores are excluded because the diagnosis is made largely based on the CDR, such that including them biases the classification (e.g., classification accuracy reaches 1.0 for CN vs. AD and 0.96 for MCI vs. AD using B^{cog} at baseline). The results of a 10-fold cross validation are shown in Tab. 1.

Classification of stable and progressive MCI Of high clinical interest is the task of identifying subjects with memory complaints that will develop Alzheimer's disease within a given period of time [9]. For example, this information is valuable to select subjects for clinical trials. Therefore, classification between MCI/AD converters and non-converters is performed using different features: A) DPs estimated with the time warp definition Eq. (3) using m12 measurements, B) DPs using baseline and m12 measurements (translational time warp Eq. (3)) and C) DPs and DPRs computed with the affine time warp Eq. (4) based on baseline and m12 visits. The set of converters (progressive MCI, pMCI) consists of all subjects that convert from MCI to AD between m12 and m36 visit. It is to be noted that these subjects are part of the training set used for learning the progression models, which introduces a bias in the test. However, adding one subject to a set of more than 1500 samples for the quantile regression barely changes the models, such that this effect can be neglected. Non-converters are test subjects with a stable MCI diagnosis (sMCI). In total, the data consists of 231 sMCI and 106 pMCI subjects. Weighted SVMs are used for classification to compensate for the unbalanced sets. Classification results are given in Tab. 2.

Classification of subjects with rapid cognitive decline (RCD) The rate of cognitive decline is known to vary considerably between subjects. It is therefore

Table 2. Results for the classification between subjects with stable and progressive MCI (sMCI vs. pMCI), and the identification of subjects with rapid cognitive decline (RCD). For all test, accuracy (ACC), sensitivity (SENS) and specificity (SPEC) of a 10-fold cross validation are given.

Biomarkers B^{est}	Visits V^{est}	Time warp $\tau(t)$	sMCI vs. pMCI			RCD vs. non-RCD		
			ACC	SENS	SPEC	ACC	SENS	SPEC
B^{cog}	{m12}	DP	0.78	0.82	0.76	0.89	0.90	0.89
	{bl, m12}	DP	0.77	0.82	0.75	0.84	0.83	0.84
	{bl, m12}	DP/DPR	0.79	0.82	0.78	0.92	0.97	0.92
B^{ml}	{m12}	DP	0.73	0.82	0.68	0.65	1.00	0.63
	{bl, m12}	DP	0.72	0.77	0.69	0.64	1.00	0.63
	{bl, m12}	DP/DPR	0.72	0.88	0.64	0.69	0.87	0.68
B^{vol}	{m12}	DP	0.68	0.59	0.71	0.73	0.83	0.73
	{bl, m12}	DP	0.68	0.54	0.74	0.74	0.87	0.73
	{bl, m12}	DP/DPR	0.66	0.68	0.66	0.75	0.83	0.74
B^{img}	{m12}	DP	0.68	0.71	0.67	0.70	0.90	0.69
	{bl, m12}	DP	0.70	0.65	0.72	0.72	0.87	0.71
	{bl, m12}	DP/DPR	0.71	0.84	0.64	0.73	0.83	0.73

of interest for family and researchers alike to predict if a subject will suffer a rapid cognitive decline (RCD). Following [1], RCD is defined as a decrease of 8 or more MMSE points in the course of 2 years (28 subjects in the test set of 633 subjects). All remaining 605 test subjects are labeled as non-RCD. Classification results are given in Tab. 2.

4 Discussion

Visually assessed, the trained multi-stage models appear plausible for all biomarkers (cf. Fig. 2 and Fig. 3). In particular, the automatically determined offset of $\hat{\delta}_p = 1860$ days entails a smooth transition between CN/MCI and MCI/AD models and is in the same range as a manual fit would be. This indicates the validity of the presented approach for model composition.

While classic machine learning methods would be restricted to the set of training subjects for which all considered biomarkers are present, all available samples are used for model learning in the presented approach. For example, a subject's CDR-SB score is considered even though no MMSE was acquired.

The estimated DPs visualised in Fig. 4 show a good class separation. Naturally, the cognitive scores perform best for distinguishing between the four classes, while image-based measurements suffer from a larger inter-subject variability. Neither volumetric nor manifold features seem to be clearly superior. Adding measurements from multiple visits slightly reduces noise and increases class separability. A particular advantage of the presented approach for progress estimation is that all available data can be used without retraining the model. It is,

for example, possible to estimate the DP if only cognitive scores are available for one visit and only image-based biomarkers for another visit. In this way, all available information can be employed in an optimal way.
The experiments show that DP and DPR are powerful features for different classification tasks. The results of the three-class classification of the disease stage are, for example, on par with [7] (even though not on exactly the same data and therefore not directly comparable). Interestingly, CN/MCI classification is worse and MCI/AD classification considerably better than in [7]. A possible explanation for this observation is the fact that fewer training samples were available from CN/MCI converters and the model is therefore less precise in this region. DP and DPR are also successfully employed to distinguish between stable and progressive MCI subjects and to identify patients with rapid cognitive decline. While not directly comparable (due to a different set of subjects), the results for sMCI vs. pMCI classification are on par with the literature (e.g., ACC=0.67 is reported for MRI-based biomarkers in [9]). No results for an automatic RCD prediction have been published so far (to the best of our knowledge), however, classification accuracy is in the same range as for predicting conversion to AD and therefore appears reasonable. In all tests, cognitive scores excel as biomarkers. Interestingly, manifold coordinates perform better then volumetric features for separating stable from progressive MCI, but worse for RCD detection. The combination of the image-based features enhances robustness and performs best on average. Further, joint DP and DPR estimation considerably enhances classification of subjects with RCD.
In summary, classification results highlight the validity of the estimated DP and DPR values. However, the method still relies on the definition of meaningful biomarkers, as the superiority of B^{cog} in all experiments shows. In future work, it would therefore be interesting to employ biomarkers based on brain atrophy, tensor-based morphometry or PET imaging. Given enough training data, it would further be highly interesting to generate personalised models for certain groups, e.g. male and female patients or APoE $\epsilon 4$ positive and negative subjects.

5 Conclusion

In this work, a biomarker-based method for modelling disease progression from a cognitively normal stage to Alzheimer's disease was proposed. This was achieved by learning two separate models for the CN/MCI and MCI/AD transition phases using quantile regression and then combining these to a multi-stage model. Further, a probabilistic approach for estimating the disease progress and the rate of progression for any given subject was presented.
Model training and progression estimation were then evaluated on the ADNI database. The estimated DPs showed good class separability on the whole domain from CN to AD. DP and DPR were further successfully employed as features to classify subjects according to their disease stage, to differentiate between stable and progressive MCI and to identify subjects with rapid cognitive decline, highlighting the versatile applicability of the presented approach.

Acknowledgements

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007 2013) under grant agreement no. 601055, VPH-DARE@IT.

References

1. Atchison, T.B., Bradshaw, M., Massman, P.J.: Investigation of profile difference between Alzheimer's disease patients declining at different rates: examination of baseline neuropsychological data. Arch Clin Neuropsychol 19(8), 1007–1015 (2004)
2. Delor, I., Charoin, J.E., Gieschke, R., Retout, S., Jacqmin, P.: Modeling Alzheimer's Disease Progression Using Disease Onset Time and Disease Trajectory Concepts Applied to CDR-SOB Scores From ADNI. CPT Pharmacometrics Syst Pharmacol 2(10), e78 (2013)
3. Doody, R.S., Pavlik, V., Massman, P., Rountree, S., Darby, E., Chan, W.: Predicting progression of Alzheimer's disease. Alzheimers Res Ther 2(1), 2 (2010)
4. Fonteijn, H.M., Modat, M., Clarkson, M.J., Barnes, J., Lehmann, M., Hobbs, N.Z., Scahill, R.I., Tabrizi, S.J., Ourselin, S., Fox, N.C., Alexander, D.C.: An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. NeuroImage 60(3), 1880–1889 (2012)
5. Green, C., Shearer, J., Ritchie, C.W., Zajicek, J.P.: Model-based economic evaluation in Alzheimer's disease: a review of the methods available to model Alzheimer's disease progression. Value Health 14(5), 621–630 (2011)
6. Guerrero, R., Ledig, C., Rueckert, D.: Manifold Alignment and Transfer Learning for Classification of Alzheimer's Disease. In: Machine Learning in Medical Imaging, 5th International Workshop, MLMI 2014. pp. 77–84 (2014)
7. Ledig, C., Guerrero, R., Tong, T., Gray, K., Schmidt-Richberg, A., Makropoulos, A., Heckemann, R., Rueckert, D.: Alzheimer s disease state classification using structural volumetry, cortical thickness and intensity features. In: MICCAI workshop Challenge on Computer-Aided Diagnosis of Dementia Based on Structural MRI Data. pp. 55–64 (2014)
8. Ledig, C., Heckemann, R.A., Makropoulos, A., Hammers, A., Lötjönen, J., Menon, D., Rueckert, D.: Robust whole-brain segmentation: Application to traumatic brain injury. Medical Image Analysis 21(1), 40–58 (2015)
9. Trzepacz, P.T., Yu, P., Sun, J., Schuh, K., Case, M., Witte, M.M., Hochstetler, H., Hake, A., Alzheimer's Disease Neuroimaging Initiative: Comparison of neuroimaging modalities for the prediction of conversion from mild cognitive impairment to Alzheimer's dementia. Neurobiol Aging 35(1), 143–151 (2014)
10. Yang, E., Farnum, M., Lobanov, V., Schultz, T., Verbeeck, R., Raghavan, N., Samtani, M.N., Novak, G., Narayan, V., DiBernardo, A., Alzheimer's Disease Neuroimaging Initiative: Quantifying the pathophysiological timeline of Alzheimer's disease. J Alzheimers Dis 26(4), 745–753 (2011)
11. Yee, T.W.: Quantile regression via vector generalized additive models. Stat Med 23(14), 2295–2315 (2004)
12. Young, A.L., Oxtoby, N.P., Daga, P., Cash, D.M., Fox, N.C., Ourselin, S., Schott, J.M., Alexander, D.C., Alzheimer's Disease Neuroimaging Initiative: A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 137(Pt 9), 2564–2577 (2014)