

Functioning

A computational method for computing an Alzheimer's disease progression score; experiments and validation with the ADNI data set



Bruno M. Jedynek^{a,b,c,*}, Bo Liu^a, Andrew Lang^d, Yulia Gel^{a,e}, Jerry L. Prince^d, for the Alzheimer's Disease Neuroimaging Initiative¹

^a Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, USA

^b Center for Imaging Science, Johns Hopkins University, Baltimore, MD, USA

^c Laboratoire de Mathématiques Paul Painlevé, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France

^d Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD, USA

^e Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario, Canada

ARTICLE INFO

Article history:

Received 27 May 2013

Received in revised form 8 March 2014

Accepted 9 March 2014

Available online 17 October 2014

Keywords:

Alzheimer's disease

Biomarkers

Progression score

Sampling from the residuals

ABSTRACT

Understanding the time-dependent changes of biomarkers related to Alzheimer's disease (AD) is a key to assessing disease progression and measuring the outcomes of disease-modifying therapies. In this article, we validate an AD progression score model which uses multiple biomarkers to quantify the AD progression of subjects following 3 assumptions: (1) there is a unique disease progression for all subjects; (2) each subject has a different age of onset and rate of progression; and (3) each biomarker is sigmoidal as a function of disease progression. Fitting the parameters of this model is a challenging problem which we approach using an alternating least squares optimization algorithm. To validate this optimization scheme under realistic conditions, we use the Alzheimer's Disease Neuroimaging Initiative cohort. With the help of Monte Carlo simulations, we show that most of the global parameters of the model are tightly estimated, thus enabling an ordering of the biomarkers that fit the model well, ordered as: the Rey auditory verbal learning test with 30 minutes delay, the sum of the 2 lateral hippocampal volumes divided by the intracranial volume, followed (by the clinical dementia rating sum of boxes score and the mini-mental state examination score) in no particular order and at last the AD assessment scale-cognitive subscale.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The ability to precisely identify the stage of disease, predict the rate of disease progression, and accurately measure the outcomes of potential therapies (Morris and Price, 2001; Perrin et al., 2011) is critical to the successful management of Alzheimer's disease (AD). The classical characterization of late-onset AD progression is a

time-ordered succession of 3 stages: normal (N), mild cognitive impairment (MCI), and AD. Physical measurements of disease progression, that is, biomarkers, are used to classify patients into these 3 stages, but it has been challenging to reliably define finer stages of the disease. In Jedynek et al. (2012), we have proposed a method for computing an Alzheimer's disease progression score (ADPS) by computationally combining 7 biomarkers of AD. A byproduct of this study was a temporal ordering of the biomarkers. Experiments conducted on the combined AD Neuroimaging Initiative (ADNI) I, GO, and II data sets provided an ordering of the biomarkers which was consistent with Jack et al. (2010) except for a cognitive test, the Rey Auditory Verbal Learning Test, 30 minutes recall (RAVLT30), which was found to become dynamic very early in the development of the disease. The statistical validation of this ordering was performed by resampling from the collection of subjects. Because the statistical model used is a regression, it can be argued that an additional validation would be obtained by sampling from the residuals.

* Corresponding author at: Department of Applied Mathematics and Statistics, Johns Hopkins University, Whitehead, 3400 North Charles Street, Baltimore, MD, 21218, USA. Tel.: +1 410 516 7341; fax: +1 410 516 7459.

E-mail address: bruno.jedynek@jhu.edu (B.M. Jedynek).

¹ Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

In this article, we validate the methodology presented in Jedynak et al. (2012) using the technique of sampling from the residuals. We adopt the model-based bootstrap that is widely used in regression analysis (see M. C, 2007; Shao and Tu, 1995, and references therein) and time series (Bühlman, 2002; Politis, 2003) and has become increasingly popular for inference of longitudinal processes (Newsom et al., 2013; Skrondal and Rabe-Hesketh, 2008). The benefit of the proposed approach is that it enables us to quantify estimation uncertainty and impute missing values (Graham, 2012). The rest of this article is organized as follows: in Section 2.1, we present the statistical model. The optimization algorithm used for fitting the parameters is presented in Section 2.2. The ADNI data set and the selection of the biomarkers is presented in Sections 2.3 and 2.4, respectively. Section 2.5 motivates and then describes the sampling method. The results are presented in Section 3 followed by a discussion of our findings in Section 4. Concluding remarks are presented in Section 5.

2. Method

Our research first describes then evaluates an algorithm for computing an ADPS, which assigns a time-dependent score to each subject.

2.1. Statistical model

The method we use is based on 3 assumptions:

1. All subjects follow a common disease progression but differ in their age of onset and rate of progression.
2. As the disease progresses, each biomarker changes continuously and monotonically following a sigmoid shaped curve.
3. In the longitudinal period over which biomarkers are observed, the rate of progression for a given subject is constant.

The proposed computation positions the longitudinal measurements of each subject on a common disease progression scale. Because it is considered to be a common scale, all subjects are expected to undergo the same biological and cognitive changes when they reach the same value (or score) on this scale. Thus, individuals are generally mapped to different positions on this scale and they progress at different rates regardless of their age of disease onset.

The age t of subject i is to be transformed into the ADPS s_i as:

$$s_i(t) = \alpha_i t + \beta_i \quad (1)$$

after estimation of the subject-dependent parameters α_i and β_i , which indicate rate and onset of disease, respectively. A linear transformation is justified because the interval over which longitudinal observations of the ADNI subjects occur is short relative to the overall disease duration. Our objective is to compute a score for all I subjects in the ADNI database by estimating $\alpha = (\alpha_1, \dots, \alpha_I)$ and $\beta = (\beta_1, \dots, \beta_I)$. The subject-dependent parameters α and β are deliberately modeled as fixed effects, not random effects, as the ADPS may ultimately be used as a covariate.

The longitudinal dynamic of each biomarker is assumed to be the same across the population and can be represented as a sigmoidal function f of ADPS s . Using $\theta_k = (a_k, b_k, c_k, d_k)$ to represent the vector of sigmoid function parameters for the k th biomarker, we can write the form of the k th biomarker as

$$f(s; \theta_k) = a_k \left(1 + e^{-b_k(s-c_k)} \right)^{-1} + d_k \quad (2)$$

The minimum and maximum values of the sigmoid function are d_k and $d_k + a_k$, and the value of s for which the biomarker is

the most dynamic, having maximum slope $a_k b_k / 4$ corresponding to its inflection point, is c_k . Sigmoids offer a parsimonious parametric model which is often a better fit than linear models for biomarkers (Caroli and Frisoni, 2010; Sabuncu et al., 2011). They are also similar in form to the conceptual evolution of biomarkers envisioned by Jack et al. (2010) (Fig. 1). A comparison of different shapes (linear, sigmoid, quadratic, and splines) of various biomarkers as function of ADAS-COG is presented in Mouiha and Duchesne (2012).

The ADNI database contains measurements y_{ijk} of biomarker k for subject i at visit j . Because there are irregularities in data collection, we use I to denote the set of triples (i, j, k) for which measurements are available. Each biomarker observation can be written as

$$y_{ijk} = f(\alpha_i t_{ij} + \beta_i; \theta_k) + \sigma_k \varepsilon_{ijk}, \quad (i, j, k) \in I, \quad (3)$$

where t_{ij} is the age of subject i at visit j . Observation noise in each biomarker is modeled for simplicity by the product of ε_{ijk} , which are independent random variables with zero mean and unit variance, and σ_k , which is the standard deviation of biomarker k . The collection of standard deviations $\sigma = (\sigma_1, \dots, \sigma_K)$ comprise another unknown that must be estimated.

The unknowns in this problem are α , β , θ , and σ and the least squares problem associated with the observation model in Equation 3 is

$$l(\alpha, \beta, \theta, \sigma) = \sum_{(i,j,k) \in I} \log \sigma_k + \frac{1}{2\sigma_k^2} (y_{ijk} - f(\alpha_i t_{ij} + \beta_i; \theta_k))^2 \quad (4)$$

Necessary conditions on the available data I for guaranteeing the identifiability of the parameters are as follows:

1. For each biomarker, there is at least 1 subject i with $\alpha_i \neq 0$ and with at least 4 distinct time points in I .
2. For each subject, there is at least 1 biomarker which is available at 2 time points in I .

In practice, a sufficient number of data points per parameter is needed to obtain tight estimators. Examining first the case with no missing data, the number of equations in Equation (3) is ijk where i is the number of subjects, j is the number of time points, and K is the number of biomarkers. The number of parameters is $2i + 5k$, counting 2 parameters per subject and 5 per biomarker (4 for the sigmoid and 1 for the standard deviation). In applications where i is large compared with k , the number of data points per parameter is close to $jk/2$. Note that longitudinal data ($j > 1$) is critical for such modeling. However, a small number j of time points together with a small number k of biomarkers is in principle acceptable. The subset of ADNI presented in Section 2.4 has numerous missing data points. We will use simulations to study the quality of the estimation of the parameter c for each biomarker. This parameter is critical for ordering the biomarkers.

2.2. Parameter fitting

Parameter fitting is performed using alternating least squares wherein the parameters θ , α , β , and σ are optimized iteratively starting from the values computed in the previous step. The details of the fitting algorithm are shown in Algorithm 1. The initial values (line 1) are $\alpha^{(0)} \equiv 1$ and $\beta^{(0)} \equiv 0$. Because of the additive form of Equation (4), optimization over θ is done serially over each of the k biomarkers while keeping $(\alpha$ and β fixed). Similarly, optimization over (α, β) is performed serially over each of the i subjects while keeping θ fixed. Fitting of θ , α , and β requires

optimization of continuously differentiable nonconvex functions, which is carried out using the Levenberg-Marquardt algorithm (Levenberg, 1944) (lines 4 and 8). I_k (line 4) is the number of subjects and visits available for biomarker k . The denominator in the equation of line 5 is the number of degrees of freedom. A canonical way to parameterize sigmoid functions is to constrain the parameter b_k in Equation (2) to be nonnegative. This is enforced with the loop over biomarkers (lines 12–16), which does not modify the objective function in Equation (4). Our experiments confirm that successful fitting is accomplished in 30 iterations, that is, $L = 30$ on line 2.

The units of ADPS are arbitrarily defined, which implies that we must choose 2 specific numerical values to fully specify the ADPS. This situation is analogous to the selection of a scale for temperature, where the numerical values of the freezing and boiling points of water determine the scale. In our experiments, we chose to fix the ADPS such that after computation over the entire population, the computed ADPS for all visits of subjects with normal clinical assessment had a trimmed mean value (m_N) and a trimmed standard deviation (σ_N), which are set respectively to 0 and 1. This is accomplished in lines 17–19.

Algorithm 1. Algorithm for the fitting of the parameters

1. Initialize $\alpha^{(0)}, \beta^{(0)}$
2. for $l = 1$ to L do
3. for $k = 1$ to K do
4. $\theta_k^{(1)} = \operatorname{argmin}_{\theta_k} \sum_{(i,j) \in I_k} (y_{ijk} - f(\alpha_i^{(0)} t_{ij} + \beta_i^{(0)}; \theta_k))^2$
5. $\sigma_k^{(1)^2} = \frac{1}{|I_k - 2| - 4|} \sum_{(i,j) \in I_k} (y_{ijk} - f(\alpha_i^{(0)} t_{ij} + \beta_i^{(0)}; \theta_k^{(1)}))^2$
6. end for
7. for $i = 1$ to I do
8. $(\alpha_i^{(1)}, \beta_i^{(1)}) = \operatorname{argmin}_{\alpha_i, \beta_i} \sum_{(j,k) \in \mathcal{I}_i} \frac{1}{\sigma_k^{(1)^2} (y_{ijk} - f(\alpha_i t_{ij} + \beta_i; \theta_k^{(1)}))^2$
9. end for
10. $\alpha^{(0)} = \alpha^{(1)}, \beta^{(0)} = \beta^{(1)}$
11. end for
12. for $k = 1$ to K do
13. if $b_k < 0$ then
14. $a_k^{(1)} = -a_k^{(1)}, b_k^{(1)} = -b_k^{(1)}, d_k^{(1)} = d_k^{(1)} + a_k^{(1)}$
15. end if
16. end for
17. for $i = 1$ to I do
18. $\alpha_i^{(1)} = \frac{\alpha_i^{(1)} - m_N}{\sigma_N}, \beta_i^{(1)} = \frac{\beta_i^{(1)} - m_N}{\sigma_N}$
19. end for

2.3. ADNI cohort

Data used in the preparation of this article were obtained from the ADNI database (<http://adni.loni.ucla.edu/>). The ADNI was launched in 2003 by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug Administration, private pharmaceutical companies, and nonprofit organizations, as a \$60 million, 5-year public private partnership. The primary goal of ADNI has been to test whether serial magnetic resonance imaging, positron emission tomography, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment and early AD. Determination of sensitive and specific markers of very early AD progression is intended to aid researchers and clinicians to develop new treatments and monitor their effectiveness, as well as lessen the time and cost of clinical

trials. The Principal Investigator of this initiative is Michael W. Weiner, MD, VA Medical Center and University of California—San Francisco. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the United States and Canada. The initial goal of ADNI was to recruit 800 adults, aged 55–90 years, to participate in the research, approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years, and 200 people with early AD to be followed for 2 years. For up-to-date information, see www.adni-info.org.

2.4. Biomarker selection

The available data in ADNI are measurements in elderly humans of multiple biomarkers associated with AD. Hundreds of subjects, categorized into N, MCI, and AD, were examined at baseline and with repeat visits every 6–12 months for a period of up to 60 months.

The following 7 biomarkers were selected for use based on their relevance in assessing the progression of AD.

1. The sum of the 2 lateral hippocampal volumes (Freesurfer version 4.4.0 for longitudinal data <http://surfer.nmr.mgh.harvard.edu>) normalized by dividing the intracranial volume (HIPPO);
2. The Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS);
3. The Mini-Mental State Examination score (MMSE);
4. The A β 42 protein level measured from the cerebrospinal fluid (ABETA);
5. The tau protein level measured from the cerebrospinal fluid (TAU);
6. The Clinical Dementia Rating Sum of Boxes score (CDRSB);
7. The Rey Auditory Verbal Learning Test, 30 minutes recall (RAVLT30)

A detailed description of the ADNI population, protocols, and biomarkers is provided at <http://adni.loni.ucla.edu/>.

The ADNI, ADNI GO, and ADNI 2 biomarker data sets were downloaded from the ADNI server (<http://adni.loni.ucla.edu/>) on November 24, 2011. All visits without date information were removed. Subjects not having at least 2 measurements for at least 1 of the 7 biomarkers were also removed. Subjects not having at least 2 measurements of the HIPPO biomarker were removed. The total number of subjects remaining was 687, where 389 were male, 275 were female, and 23 had unknown sex. The total number of visits was 3658, and the clinical diagnoses at these visits were 1103 normal, 1513 MCI, and 1010 AD (note that we used ABETA and TAU data from files UPENNBIOMK, UPENNBIOMK2, UPENNBIOMK3, and UPENNBIOMK4. For each subject, we use the latest available file [batch], i.e., if UPENNBIOMK4 data are available for this subject, we use it. Otherwise, we use UPENNBIOMK3 and so on. For each subject, all his/her ABETA and TAU data always come from the same file [batch]). Of the 7 biomarkers considered, only ADAS and RAVLT30 were available at the time of download from the ADNI 2/GO data set. The protocol for these biomarkers is the same in ADNI, ADNI 2, and ADNI GO.

2.5. Sampling from the residuals

The analysis of a longitudinal simultaneously acquired collection of biomarkers of ADNI data set is a complex task for several reasons. First, for each biomarker, the sequences of measurements obtained across timepoint are correlated. Table 1 shows the correlation

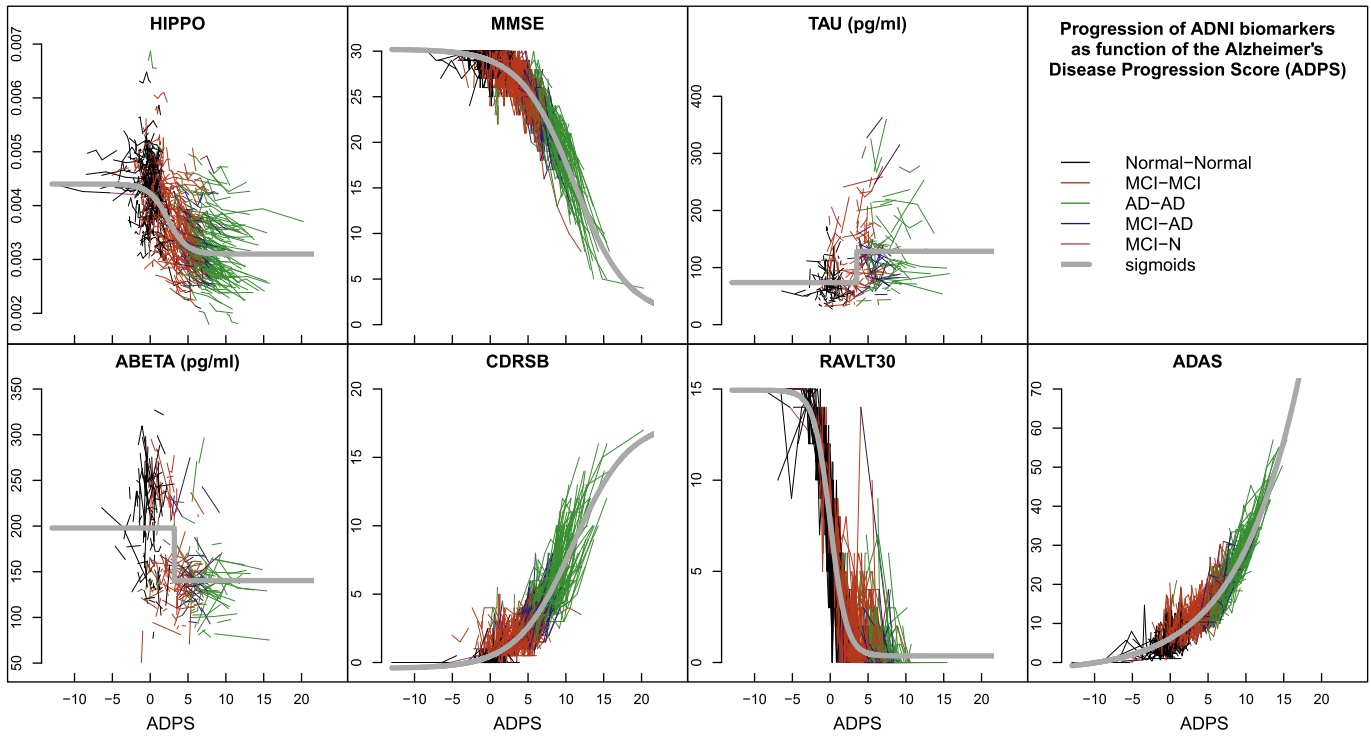


Fig. 1. The values of 7 biomarkers, measured at all visits of all ADNI subjects, are plotted on the normalized ADPS. Each connected polyline represents the consecutive visits of a single subject and each line segment is colored according to the subject's clinical diagnoses between visits (see legend). The gray curves are the sigmoid functions representing the fitted behavior of each biomarker in the normalized space. Reproduced from Jedynak et al. (2012). Abbreviations: ADNI, Alzheimer's Disease Neuroimaging Initiative; ADPS, Alzheimer's disease progression score. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

between measurements taken at baseline and 1 year after baseline. The correlation is larger than 0.7 for all the biomarkers and is as high as 0.98 in case of the volume of the hippocampus over all subjects.

Second, the structure of the missing data in ADNI is complex. The schedule of visits depends on the status (N, MCI, and AD) at baseline and is different for different biomarkers (see ADNI General Procedures Manual, pages 6, 7, and 8). Moreover, the subjects are recruited over a relatively long period; hence, the number of available measurements at an earlier visit (e.g., baseline) is significantly larger than at a later one (e.g., baseline plus 36 months). Table 2 shows the proportion of available measurements relative to the total number of subjects for each of the 7 selected biomarkers and each timepoint.

We now describe the sampling from the residual algorithm which we use. The key idea is to construct a bootstrap distribution of y_{ijk} while taking into account its dependence structure, which then

can be used for inference and imputation purposes. In particular, the residuals ε_{ijk} are ideally independent and identically distributed. However, because measurements are taken consecutively over time, this is not the case, as Table 1 indicates. Thus, we can attempt to filter out temporal dependence among ε_{ijk} by applying an appropriate filter (e.g., an autoregressive model) and the resulting new filtered residuals η_{ijk} should be close to the independent and identically distributed assumption. Hence, given conditional independence of η_{ijk} , we can use the classical bootstrap and sample new realizations of η_{ijk}^* from the estimated empirical probability distribution of η_{ijk} . Substituting in the sampled η_{ijk}^* into the filter yields new bootstrap residuals ε_{ijk}^* , which in turn can be plugged-in into the estimated model Equation (3), and the new proxy bootstrap values y_{ijk}^* are thus obtained. The resulting bootstrap distribution of y_{ijk}^* serves as a proxy to the unknown distribution of y_{ijk} and can be used to assess errors in parameter estimation and imputation of missing values as described in the Algorithm 2 in the following.

Table 1
Correlation coefficient for each biomarker between the measurements at baseline and 1 year after baseline

Biomarker	Correlation coefficient
HIPPO	0.98
MMSE	0.74
ABETA	0.94
RAVLT30	0.8
ADAS	0.83
TAU	0.93
CDRSB	0.85

Key: ADAS, Alzheimer's disease assessment scale; CDRSB, clinical dementia rating sum of boxes; MMSE, Mini-Mental State Examination; RAVLT30, Rey Auditory Verbal Learning Test, 30 minutes recall.

Table 2
Proportion in percentage of available data for each biomarker and each time point

Time point	HIPPO	ADAS	MMSE	TAU	ABETA	CDRSB	RAVLT30
Baseline	99	100	100	53	53	100	99
6 mo	97	99	99	0	0	99	99
12 mo	96	99	99	45	45	98	98
18 mo	40	45	45	0	0	45	45
24 mo	73	88	88	13	13	87	86
36 mo	20	62	64	3	3	62	63

The denominator is the number of subjects with ADAS measurement at baseline. Note that only MCI subjects have a visit at 18 months after baseline which explains why the numbers are lower for all the biomarkers at the 18-month visit. Key: ADAS, Alzheimer's disease assessment scale; MCI, mild cognitive impairment.

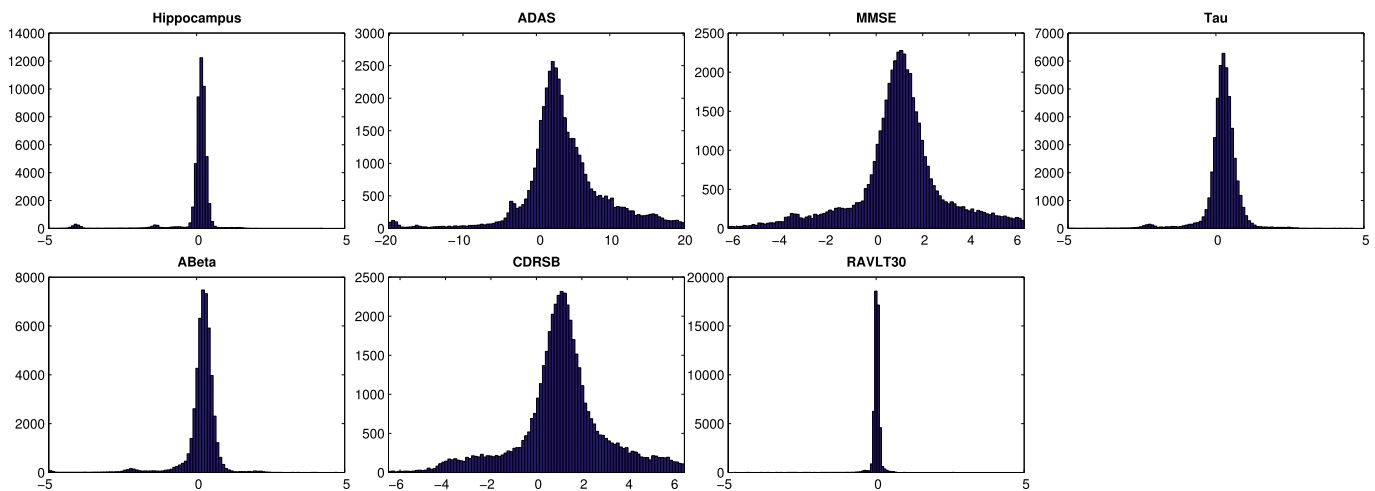


Fig. 2. Distribution of the errors in estimating the parameter c of the sigmoid (see Equation 2) as function of the progression time line of the disease. (For interpretation of the references to color in this Figure, the reader is referred to the web version of this article.)

To assess the quality of the parameter fitting algorithm presented in Algorithm 1, we generated a large number ($M = 40,000$) of simulated data sets by sampling from the residuals as follows. First, we estimated all the parameters in the model using Algorithm 1 and the ADNI data set. Second, we computed the residuals ε_{ijk} from Equation (3). Because these residuals showed correlation over time for each biomarker, we fit an autoregressive AR(1) model for the residuals of each biomarker. We then repeated the procedure described in Algorithm 2 $M = 40,000$ times.

Algorithm 2. Algorithm for sampling from the residuals

1. Sample η_{ijk} independently from a standard normal distribution, assuming no missing data.
2. Compute ε_{ijk} using the AR(1) model, imputing the values η_{ijk} as residuals.
3. Compute y_{ijk} using the model Equation (3), imputing ε_{ijk} as residuals.
4. Determine the missing data for each biomarker with the following sampling procedure: sample from a 2 state (not available [NA]) nonhomogeneous Markov chain indexed by the successive visits. The transition matrix was estimated separately for each status at baseline (N, MCI, and AD) and before hand.
5. Estimate all the parameters from the simulated data set with missing data obtained at the last step using Algorithm 1. Notate $c_k^{(m)}$ the estimated parameter c for biomarker k at iteration m .

3. Results

3.1. ADPS computed for ADNI subjects

The ADPS was computed for all subject visits in the combined ADNI, ADNI 2, and ADNI GO data sets (with minimal exclusions as described in Section 2.4). Results are presented in Fig. 1. Overall, normal subjects (black) have the smallest ADPS, MCI subjects (red) have moderate ADPS, and AD subjects (green) have the largest ADPS. Lower ADPS are therefore consistent with the normal population and higher ADPS are indicative of increased presence of dementia. Those subjects whose clinical status changes from MCI to AD (blue) are found mostly between the red and green colors. The estimated sigmoidal behaviors of each biomarker are shown in gray.

3.2. Assessment of the quality of the estimation of the inflection point of each biomarker curve

The estimator of the parameter c_k for biomarker k using the ADNI data set is denoted as c_k^T , where “T” stands for “target.” Each sampling from the residuals, indexed by m , produces a simulated data set from which an estimator of c_k is computed using Algorithm 1 and is notated $c_k^{(m)}$. The histograms for the data $\{c_k^{(m)} - c_k^T\}_{m \in \{1, \dots, 40,000\}}$ for each biomarker k are presented in Fig. 2. In comparing these histograms, be advised that the scale of the horizontal axis is not the same for each biomarker. Recall that the parameter c_k , the inflection point of each sigmoid, records the moment in the disease progression where the biomarker k changes the most (is the most dynamic). First, consider the biomarkers HIPPO, TAU, ABETA, and RAVLT30. The mean square error for these estimators is small. Indeed, the histograms obtained by sampling from the residuals are centered close to the origin with a small standard deviation. These results give validity to the choices and the settings of the optimization algorithm. Second, in the case of MMSE and CDRSB, some bias is observed, and the standard deviation is moderated. Finally, in the case of ADAS, the bias as well as the standard deviation is larger. Note that the fitted sigmoids for these 3 biomarkers do not level off in the later stages of the disease (see Fig. 1), which might explain why the fit is less stable (large bootstrap standard deviation). One remedy to stabilize the estimation could be to constrain the extremal values, d_k and $d_k + a_k$. For example, the maximum value for ADAS is 70 which could be enforced during the optimization process.

3.3. Ordering of the biomarkers

For each simulation, the parameter c_k , that is, the inflection point of the sigmoid fitted to biomarker k , was obtained, and the summary of the ordering of the c_k values for each of the 40,000 simulations is presented in Table 3. Clearly, the ordering of the most prevalent is as follows: RAVLT30, HIPPO, ABETA, TAU, then (CDRSB, MMSE) in no particular order, and then ADAS. This result is consistent with our results obtained in Jedynak et al. (2012) where the statistical method used to assess this ordering, that is, resampling from the subjects, was different than the method used here, that is, sampling from the residuals. Care should be taken in interpreting this result in term of the ordering of the biomarkers. In particular, ABETA and TAU biomarkers are not well explained by the

Table 3
Ordering of the biomarkers according to the location of the inflection point

	#1	#2	#3	#4	#5	#6	#7
RAVLT30	99.7	0.3	0	0	0	0	0
HIPPO	0	99.4	0.4	0.1	0	0.1	0
ABETA	0.1	0.2	95.9	3.7	0	0	0
TAU	0.1	0.1	3.8	96	0	0	0
CDRSB	0	0	0	0.1	65.4	33.5	1
MMSE	0	0	0	0	32.3	66	1.7
ADAS	0	0	0	0	2.2	0.4	97.3

For each biomarker (line), the value recorded in column *j* is the number of times, in percentage, this biomarker has an inflection point which is the *j*th smallest of the 7 biomarkers. This was computed using 40,000 independent samples. The results are truncated to the nearest 10th of a percentage.

Key: ADAS, Alzheimer's disease assessment scale; CDRSB, clinical dementia rating sum of boxes; MMSE, Mini-Mental State Examination; RAVLT30, Rey Auditory Verbal Learning Test, 30 minutes recall.

model, see Fig. 1. One can visually see that there is a large residual variance in the case of these 2 biomarkers by comparing the spread of the data points around the sigmoid curves in gray. As a consequence, we prefer to eliminate these biomarkers from our analysis and propose the following ordering of the biomarkers: RAVLT30, HIPPO, then (CDRSB and MMSE) in no particular order, and then ADAS.

4. Discussion

Amyloid concentration in the brain is known to change very early in the disease process (Vemuri et al., 2010; Villemagne et al., 2013). Why is it then that we do not detect an early change of ABETA? There are several nonexclusive possible explanations. One is that ABETA is a noisy measurement of the amyloid burden, eventually contaminated by 1 or several physiological covariates. A complementary explanation is that there are several paths to disease within the ADNI subjects. For example, there might be MCI or even AD subjects who are impaired for reasons unrelated to AD (depression, tau only pathologies, and vascular dementia). Recall that the first assumption in our model is that there is a unique disease progression for all the subjects. A violation of this hypothesis would result in a lack of fit for some of the biomarkers and could also modify our conclusions about the ordering of the biomarkers.

The fact that the RAVLT30 biomarker is dynamic early in the disease process is an interesting result which deserves further investigations in ADNI. Using data from the Canadian Study of Health and Aging, it was found in Tierney et al. (2005) that the RAVLT was predictive of neurodegenerative changes up to 10 years before diagnosis. Also, early changes in the hippocampus volume might occur very early in the disease process while being too subtle to be detected with the current protocol and that progress in the acquisition and/or image processing technology might reveal these subtleties. Finally, the reader is reminded that the results were obtained for the ADNI data set as of November 4, 2011 and do not necessarily extrapolate to a larger or different population.

5. Conclusion

We have presented a validation of a multi-biomarker, data-driven approach to assess time-dependent changes of biomarkers in AD and to localize subjects on a common scale of disease progression over the entire range of progression represented in the ADNI cohort. The sampling from the residuals analysis shows that the inflection point of the biomarker sigmoid curves is well estimated for most biomarkers. Our presented model and subsequent validation argue that the following ordering of the biomarkers

should be considered: RAVLT30, HIPPO, then (CDRSB and MMSE) in no particular order, and at last ADAS.

Disclosure statement

The authors have no conflicts of interest to disclose.

Acknowledgements

Personnel costs for this research were partially supported by a grant from Pfizer as well as from an Ossoff scholar award. Other support came from grants numbered P41EB015909 and R01EB012547 from the National Institute of Biomedical Imaging and Bioengineering. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer Drug Discovery Foundation; Amorphix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Foundation; Eisai; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd; and its affiliated company "Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514.

References

- Bühlman, P., 2002. Bootstraps for time series. *Stat. Sci.* 17, 52–72.
- Caroli, A., Frisoni, G., the Alzheimer's Disease Neuroimaging Initiative, 2010. The dynamics of Alzheimer's disease biomarkers in the Alzheimer's disease neuroimaging initiative cohort. *Neurobiol. Aging* 31, 1263–1274.
- Graham, J., 2012. *Missing Data: Analysis and Design*. Springer Verlag, New York.
- Jack Jr., C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Trojanowski, J.Q., 2010. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* 9, 119–128.
- Jedynak, B.M., Lang, A., Liu, B., Katz, E., Zhang, Y., Wyman, B.T., Raunig, D., Jedynak, C.P., Caffo, B., Prince, J.L., 2012. Alzheimer's Disease Neuroimaging Initiative. A computational neurodegenerative disease progression score: method and results with the Alzheimer's disease Neuroimaging Initiative cohort. *Neuroimage* 63, 1478–1486.
- Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* 2, 164–168.
- Chernick, M.R., 2007. *Bootstrap Methods: A Guide for Practitioners and Researchers*. Wiley, New Jersey.
- Morris, J., Price, J., 2001. Pathologic correlates of nondemented aging, mild cognitive impairment, and early-stage Alzheimer's disease. *J. Mol. Neurosci.* 17, 101–118.
- Mouha, A., Duchesne, S., 2012. Toward a dynamic biomarker model in Alzheimer's disease. *J. Alzheimers Dis.* 30, 91–100.
- Newsom, J., Jones, R., Hofer, S., 2013. *Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health, and Social Sciences*. Routledge, New York.
- Perrin, R.J., Craig-Schapiro, R., Malone, J.P., Shah, A.R., Gilmore, P., Davis, A.E., Roe, C.M., Peskind, E.R., Li, G., Galasko, D.R., Clark, C.M., Quinn, J.F., Kaye, J.A., Morris, J.C., Holtzman, D.M., Townsend, R.R., Fagan, A.M., 2011. Identification and validation of novel cerebrospinal fluid biomarkers for staging early Alzheimer's disease. *PLoS One* 6, e16032.
- Politis, D., 2003. The impact of bootstrap methods on time series analysis. *Stat. Sci.* 18, 219–230.

- Sabuncu, M.R., Desikan, R.S., Sepulcre, J., Yeo, B.T.T., Liu, H., Schmansky, N.J., Reuter, M., Weiner, M.W., Buckner, R.L., Sperling, R.A., Fischl, B., ADNI, 2011. The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch. Neurol.* 68, 1040–1048.
- Shao, J., Tu, D., 1995. *The Jackknife and Bootstrap*. Springer, New York.
- Skrdal, A., Rabe-Hesketh, S., 2008. *Multilevel and Related Models for Longitudinal Data*. Springer Verlag, London.
- Tierney, M.C., Yao, C., Kiss, A., McDowell, I., 2005. Neuropsychological tests accurately predict incident Alzheimer disease after 5 and 10 years. *Neurology* 64, 1853–1859.
- Vemuri, P., Wiste, H.J., Weigand, S.D., Knopman, D.S., Trojanowski, J.Q., Shaw, L.M., Bernstein, M.A., Aisen, P.S., Weiner, M., Petersen, R.C., Jack, C.R., Alzheimer's Disease Neuroimaging Initiative, 2010. Serial MRI and CSF biomarkers in normal aging, MCI, and AD. *Neurology* 75, 143–151.
- Villemagne, V., Burnham, S., Bourgeat, P., Brown, B., Ellis, K.A., Salvado, O., Szoek, C., Macaulay, L., Martins, R., Maruff, P., Ames, D., Rowe, C., Masters, C., 2013. Amyloid-beta deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* 12, 357–367.