Diagnostic Assessment & Prognosis

Predicting time to dementia using a quantitative template of disease progression

Murat Bilgel[a], Bruno M. Jedynak[b,*], for the Alzheimer's Disease Neuroimaging Initiative[1]

[a]*Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA*
[b]*Dept. of Mathematics and Statistics, Portland State University, Portland, OR, USA*

Abstract

Introduction: Characterization of longitudinal trajectories of biomarkers implicated in sporadic Alzheimer's disease (AD) in decades before clinical diagnosis is important for disease prevention and monitoring.

Methods: We used a multivariate Bayesian model to temporally align 1369 Alzheimer's disease Neuroimaging Initiative participants based on the similarity of their longitudinal biomarker measures and estimated a quantitative template of the temporal evolution of cerebrospinal fluid $A\beta_{1-42}$, p-tau$_{181p}$, and t-tau and hippocampal volume, brain glucose metabolism, and cognitive measurements. We computed biomarker trajectories as a function of time to AD dementia and predicted AD dementia onset age in a disjoint sample.

Results: Quantitative template showed early changes in verbal memory, cerebrospinal fluid $A\beta_{1-42}$ and p-tau$_{181p}$, and hippocampal volume. Mean error in predicted AD dementia onset age was <1.5 years.

Discussion: Our method provides a quantitative approach for characterizing the natural history of AD starting at preclinical stages despite the lack of individual-level longitudinal data spanning the entire disease timeline.

© 2019 The Authors. Published by Elsevier Inc. on behalf of the Alzheimer's Association. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Alzheimer; Dementia; Onset; Prediction; Longitudinal; Biomarkers; Cognition; Quantitative template; Kaplan-Meier

1. Background

Alzheimer's disease (AD)–related brain changes, including amyloid and phosphorylated tau deposition as well as neurodegeneration, begin years before the emer-

gence of clinical dementia [1]. There is great interest in understanding the progression of biological and cognitive markers (collectively referred to as biomarkers in this article) that are implicated in AD in the earliest stages, given that therapeutic intervention is hypothesized to be more effective if administered early in the disease before downstream brain damage occurs. Such an understanding will enable a better definition of preclinical AD and help identify individuals who are likely to benefit from therapy.

In studies of dominantly inherited AD, parents' ages at disease onset serve as strong predictors of an individual's onset age, allowing the use of time from expected onset as a surrogate of disease progression against which biomarker trajectories can be characterized [2]; however, in sporadic AD, such an estimate is not as good an indicator of expected onset, making it difficult to characterize biomarker evolution

in the earliest stages of neuropathology and neurodegeneration that mark the preclinical stages of AD. Furthermore, differences across individuals in the rate of disease progression make it difficult to characterize a quantitative template (QT) of biomarker changes as a function of disease state. Therefore, it is necessary to use novel statistical approaches that take into account such differences across individuals for combining short-term follow-up data per individual to reveal long-term biomarker trajectories.

Several studies have addressed this challenge of characterizing biomarker trajectories in preclinical stages of AD as a function of an underlying latent disease progression variable that reflects the natural history of AD neuropathology, neurodegeneration, and cognitive changes via generative models. These models emphasize the interpretability of model results rather than optimization of predictive performance, where discriminative methods might outperform but produce results that are not as easily interpretable. Existing generative models of AD progression can be divided into two major categories based on the granularity of their characterization of the latent disease progression variable, either as a sequence of discrete events [3–5] or as a continuous variable [6–14]. In addition to characterizing changes in biomarker trajectories as a function of latent disease stages, these statistical models provide individualized information that can be used for personalized disease staging and monitoring.

Recent work in this area has focused on Bayesian reformulations of statistical approaches to biomarker trajectory modeling [11,14,15] to enable probabilistic estimates of trajectories and better characterization of the individual-level uncertainty in disease progression variables. These improvements can lead to better disease monitoring and progression prediction at the individual level, thereby providing useful tools for clinical trial recruitment and assessment. We further this line of research into continuous latent disease progression models by reformulating our progression score models [6,7] as a Bayesian model and make substantial changes to improve the interpretability of our results. First, we impose weakly informative priors on model parameters. This allows us to compute credible intervals around estimated trajectories and individualized disease stage indicators. Second, based on our earlier observation that rate of progression is associated with disease stage, we revise the transformation between age and the latent disease progression variable, so that biomarker trajectories can be depicted as a function of time to diagnosis, which was not directly possible in our previous model.

In addition to these substantial improvements to our previously described model, our work addresses several limitations of previously described disease progression models. (1) Existing disease progression score models formulated in a Bayesian framework [11,14,15] specify the latent disease progression variable using a single unknown parameter, the time-shift, meaning that these models do not take into account that disease progression may accelerate over time, limiting their accuracy when working with individual-level data over long periods of time. Our proposed progression score (PS) takes this phenomenon into account through the use of an additional parameter in the construction of PSs. (2) Current models either assume independence across biomarkers [15], resulting in biased estimates of latent disease scores [7], or impose linear [11] or double exponential [14] parametric trajectories on biomarkers, limiting the possible functions that can be estimated, potentially resulting in mischaracterization of the time evolution of biomarkers. We model biomarker time courses using flexible monotonic nonlinear functions characterized by basis functions. This allows for trajectory models that are as versatile as those modeled via monotonic gaussian processes as proposed by Lorenzi et al. [15], but that do not require working with large covariance matrices involved in the monotonic gaussian process setting. This computational saving allows us to incorporate the correlations among biomarkers to reduce the bias in the estimation of PSs. (3) Although several classifiers have been developed to predict a diagnostic label (i.e., cognitively normal [NL], mild cognitive impairment [MCI], or AD; MCI nonconverter vs. MCI converter) at the end of a fixed and short time frame (typically <3 years) given baseline biomarkers [16–22], prior literature on continuous prediction of time to dementia over longer periods is quite limited. We describe a novel application of disease progression modeling results for predicting time to dementia onset and propose the use of Kaplan-Meier curves for assessing prediction accuracy.

Using our Bayesian Progression Score Model, we compute a QT of the temporal evolution of AD-related biomarkers by temporally aligning longitudinal data for individuals based on the similarity of a collection of biomarkers for individuals who are NL or have MCI or AD dementia. The QT reflects change in biomarkers as a function of a latent disease stage indicator that is simultaneously estimated per individual in the model. Our model enables the characterization of the trajectories in time domain, thereby facilitating the interpretation and applicability of our results to clinical settings. The estimated temporal QT shows that verbal memory decline is detectable early on the trajectory to MCI and AD. The temporal evolution is also marked by changes in hippocampal volume, cerebrospinal fluid (CSF) amyloid ($A\beta_{1-42}$), total tau (t-tau), and phosphorylated tau (p-tau$_{181p}$), as well as brain glucose metabolism and global measures of cognition and mental status. We demonstrate that the estimated latent disease progression variable is associated with known risk factors for late-onset AD and clinical diagnosis. Finally, we predict AD dementia onset age using baseline data in a disjoint testing set, achieving a root mean square error (RMSE) of <1.5 years. Our results provide insights into the natural history of late-onset AD starting with the preclinical stages. The estimated QT can be used to estimate individualized latent disease stages given a collection of biomarker measurements and predict future conversion to AD.

2. Method

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of the ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. Specifically, we used the ADNI data prepared for the Alzheimer's Disease Modelling Challenge and followed the recommendations prescribed in this document for easing the comparison with other QTs for the progression of AD.

2.1. Participants

We used data for 1706 participants with 6880 visits from ADNI 1/GO/2. Clinical diagnoses of MCI and AD dementia were determined by the ADNI Conversion Committee according to the criteria described in the ADNI protocol. We designated data for approximately 80% of the individuals selected at random as the training set (1369 individuals with 5533 visits). We designated the data for the remaining individuals, excluding 67 individuals with AD at baseline, as the testing set (270 individuals with 1165 visits) for evaluating the performance of age at AD onset prediction. Participant demographics are presented in Table 1. We compared demographic variables by diagnostic category using the Wilcoxon rank-sum test for continuous variables and Fisher's exact test for categorical variables.

2.2. Biomarkers and cognitive measures

We computed a QT of the temporal evolution of the following AD-related biomarkers: CSF $A\beta_{1-42}$, p-tau$_{181p}$, t-tau; intracranial volume-adjusted hippocampal volume; brain glucose metabolism measured by fluorodeoxyglucose PET; verbal memory measured by Rey Auditory Verbal Learning Test immediate recall (sum score across 5 learning trials) [23]; mental status measured by the Mini–Mental State Examination (MMSE) [24]; and AD and dementia indicators as measured by the Alzheimer's Disease Assessment Scale-Cognitive 13-item scale [25] and the Clinical Dementia Rating-Sum of Boxes [26]. These measures were selected based on their demonstrated involvement in the progression of AD and closely mirror those selected for analysis with an earlier version of our model [6]. We selected visits with at least 5 of these 9 measures for inclusion in our analysis.

2.3. Progression score model

We reformulated the previously described progression score models [6,7] using a Bayesian framework where biomarker trajectories are modeled using basis functions. Compared with linear or sigmoid functions used in our previous applications, basis functions allow for much richer models of biomarker trajectories. We make the following modeling assumptions:

1. Changes in biomarkers relative to one another in the progression from a NL state to AD dementia can be characterized uniquely.
2. Any deviation from this unique characterization of biomarker changes along the disease spectrum is assumed to be measurement noise.
3. Given enough time, all individuals will exhibit biomarker levels seen in AD dementia; however, an individual's progression might be slow enough that AD-level biomarkers will not be attained during the individual's life span.

2.4. Continuous-time model

We first describe the model in continuous time. We then discretize it for fitting the parameters. The model describes biomarker values collectively as a function of age using a two-level composition. First, there is a subject-dependent exponential mapping of age t to PS s:

$$\frac{\partial s}{\partial t} = \alpha s, \quad \text{with } s(t_0) = \gamma, \tag{1}$$

where $\alpha > 0$ is a global parameter and $\gamma > 0$ is a subject-specific variable. t_0 is a fixed age, here $t_0 = 70$, so that γ is the PS at 70 years. This model is partially motivated by Jedynak et al.[6, Fig. 4] and Bilgel et al. [7, Fig. 6 right]. In both applications, the PS is an affine function of age for each subject, yet Eq. 1 provides a reasonable approximation. Solution of Eq. 1 is given by

$$s = f(t, \gamma) = \gamma e^{\alpha(t-t_0)} \tag{2}$$

Second, the vector of biomarker values at time t is modeled with a monotonic function of the PS for each biomarker added to a correlated noise:

$$\mathbf{y}(t) = \mathbf{g}(f(t, \gamma); \omega) + \varepsilon(t) \tag{3}$$

Monotonicity of $\mathbf{g}$ is required for model identifiability.

2.5. Time from diagnosis

Diagnostic events include transitions from NL or MCI to AD. The aforementioned model allows for the computation of biomarker trajectories as a function of time from a diagnosis. Let $\tilde{s}$ be the PS corresponding to a diagnosis. For an individual with $s(t_0) = \gamma$, the age $\tilde{t}$ at diagnosis is, using Eq. 2,

$$\tilde{t}(\tilde{s}) = t_0 + \frac{1}{\alpha} \ln \frac{\tilde{s}}{\gamma} \tag{4}$$

Table 1
Participant demographics at baseline

Characteristic	Training set			Testing set	
	NL, $n = 405$	MCI, $n = 698$	AD, $n = 266$	NL, $n = 114$	MCI, $n = 156$
Age	74.5 (5.81)	72.8 (7.62)	74.9 (7.71)	73.3 (5.78)	73.5 (7.40)
Male	197 (48.6%)	413 (59.2%)	149 (56.0%)	56 (49.1%)	91 (58.3%)
White	372 (91.9%)	651 (93.3%)	246 (92.5%)	100 (87.7%)	146 (93.6%)
Education	16.3 (2.62)	15.9 (2.85)	15.3 (3.02)	16.5 (2.97)	16.0 (2.85)
Single $\varepsilon 4$	105 (26.0%)	280 (40.2%)	132 (50.0%)	32 (28.1%)	56 (35.9%)
$\varepsilon 4/\varepsilon 4$	7 (1.73%)	75 (10.8%)	46 (17.4%)	5 (4.39%)	18 (11.5%)
No. of visits	4.21 (2.28)	4.45 (1.97)	2.71 (1.14)	4.25 (2.27)	4.37 (2.04)
Follow-up	3.15 (2.38)	2.75 (1.84)	1.14 (0.84)	3.14 (2.44)	2.80 (1.99)

Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment; NL, cognitively normal.

NOTE. Continuous variables are reported as "mean (standard deviation)" and categorical variables are reported as "count (percentage)".

The time from diagnosis for this subject at age t is thus $u = t - \tilde{t}(\tilde{s})$. Biomarker trajectories as a function of time from diagnosis are characterized by the mapping

$$u \mapsto \mathbf{g}(f(u+\tilde{t}(\tilde{s}), \gamma); \omega) \qquad (5)$$

2.6. Discrete-time model

We now describe the discrete-time model and the priors. Let i indicate a subject and j a visit. The subject-specific model is

$$s_{ij} = \gamma_i e^{\alpha(t_{ij}-t_0)} \qquad (6)$$

$$\alpha \sim \text{half} - \mathcal{N}(0, \sigma_\alpha^2) \qquad (7)$$

$$\gamma_i \sim \text{half} - \mathcal{N}(0, \sigma_\gamma^2), \qquad (8)$$

where t_{ij} is the age of subject i at visit j, and half$-\mathcal{N}$ is the half-normal distribution. The choice of t_0 is important for the interpretation of γ. At $t = t_0$, the PS of individual i is equal to γ_i. Furthermore, the parameter σ_γ reflects the variability in the PSs at age t_0. In this study, we let $t_0 = 70$ so that it is reasonably close to the mean age of our sample at baseline. We use half-normal hyperpriors on σ_α and σ_γ, with their scale parameters fixed at 0.05 and 5, respectively:

$$\sigma_\alpha \sim \text{half} - \mathcal{N}(0, 0.05^2) \qquad (9)$$

$$\sigma_\gamma \sim \text{half} - \mathcal{N}(0, 5^2) \qquad (10)$$

Given K biomarkers, their trajectory models parameterized by ω are given by

$$\mathbf{y}_{ij} = \mathbf{g}(s_{ij}; \omega) + \varepsilon_{ij} \qquad (11)$$

$$\varepsilon_{ij} \sim \mathcal{N}_K(0, \Lambda C \Lambda) \qquad (12)$$

$$C \sim \text{LKJ}(\nu), \qquad (13)$$

where Λ is a diagonal matrix with $\Lambda_{kk} = \lambda_k$, C is a correlation matrix, and LKJ is the random correlation matrix distribution described by Lewandowski et al. [27]. We let $\nu = 1$ for a uniform distribution over correlation matrices.

We consider logistic basis functions to characterize monotonic trajectories for each biomarker k:

$$g_k(s_{ij}; \omega_k) = \pi_k \sum_{z=1}^{Z} a_{kz} \phi_{kz}(s_{ij}; \ell_z, \sigma_{\text{basis}}^2) + b_k \qquad (14)$$

$$\phi_{kz}(s_{ij}; \ell_z, \sigma_{\text{basis}}^2) = \frac{1}{1+\exp\left(-\frac{(s_{ij}-\ell_z)}{\sigma_{\text{basis}}^2}\right)}, \qquad (15)$$

where $\omega_k = \{\pi_k, a_{k1}, ..., a_{kZ}, b_k\}$, with the constraint that $a_{kz} > 0 \ \forall k, z$ to ensure monotonicity. π_k is a categorical random variable with equally likely observations $\{-1, +1\}$ to determine whether the trajectory is decreasing or increasing, $\ell \in \mathbb{R}^Z$ is a prespecified set of Z basis function locations, and σ_{basis}^2 determines the slant of each basis function.

Assuming that all biomarker observations y_{ijk} have been standardized to have zero mean and unit standard deviation at baseline, we can let

$$a_{kz} \sim \text{half} - \mathcal{N}(0) \qquad (16)$$

$$b_k \sim \mathcal{N}(0) \qquad (17)$$

$$\lambda_k \sim \text{half} - \mathcal{N}(0) \qquad (18)$$

The standard deviation or scale parameter of each is fixed at a large value (i.e., 10) such that the priors are essentially uninformative. In the absence of data, these priors slightly favor flat "trajectories" for the biomarkers.
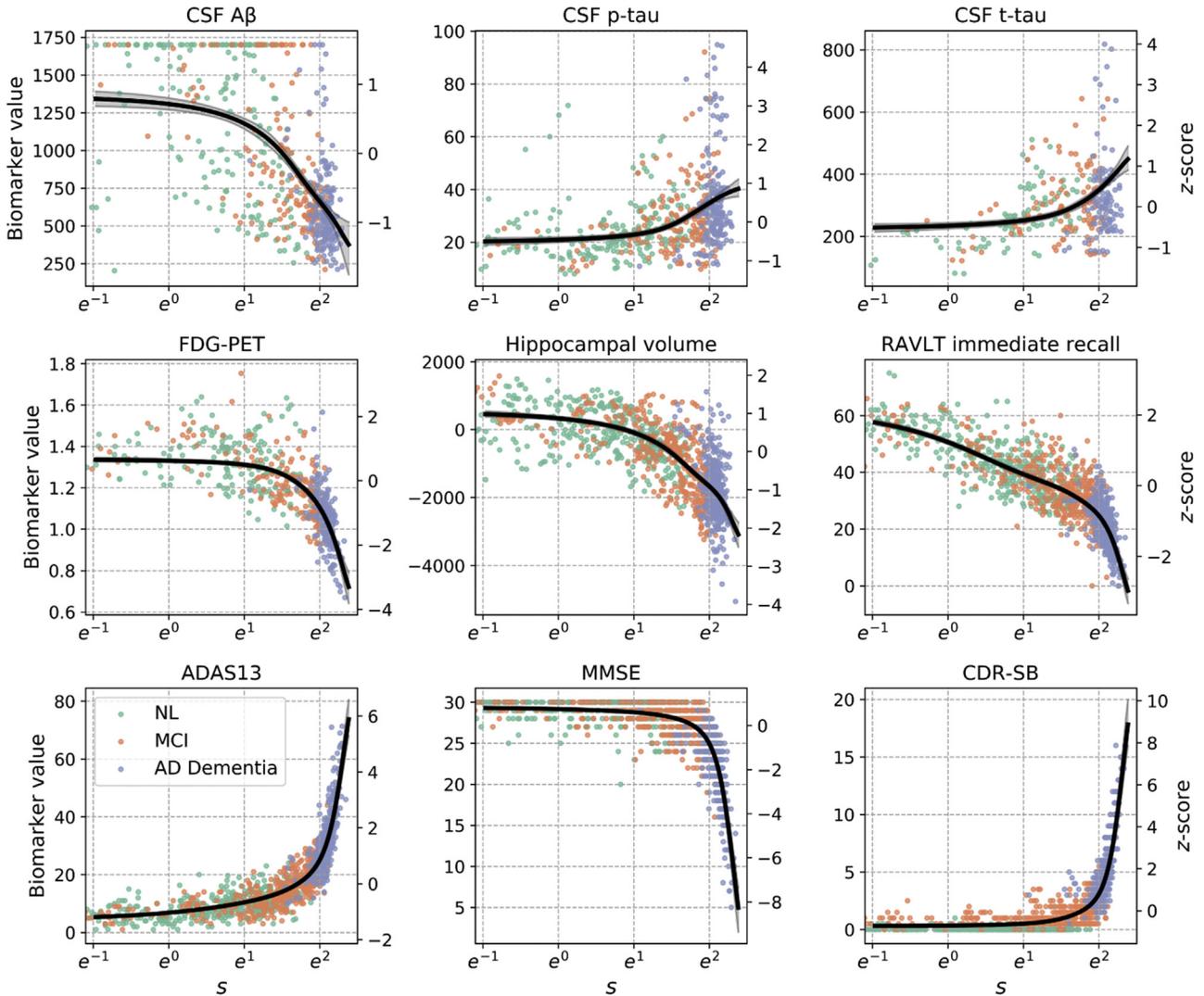
Fig. 1. Estimated population-level biomarker trajectories as a function of progression score *s*; *s* is plotted on a natural logarithm scale so that the *x*-axis is linear in time. Mean trajectories are plotted in black, along with their 95% credible intervals. Longitudinal data points for 100 randomly sampled individuals per diagnostic category are shown. Biomarker *z*-scores, shown on the right-hand-side *y*-axes, were computed using mean and standard deviations at baseline across 1369 in the training set. Abbreviations: Aβ, amyloid β; CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; PET, positron emission tomography; RAVLT, Rey Auditory Verbal Learning Test; ADAS13, Alzheimer's Disease Assessment Scale-Cognitive 13-item scale; MMSE, Mini–Mental State Examination; CDR-SB, Clinical Dementia Rating-Sum of Boxes.

2.7. Model fitting

We conducted Markov chain Monte Carlo sampling to generate posterior samples, using the No-U-Turn Sampler step method [28] for continuous variables and a Metropolis-within-Gibbs step method optimized for binary variables for π_k. Nonnegative variables were log-transformed to allow for unconstrained optimization. For visits with missing biomarker measurements, only the available measurements contributed to the model fitting procedure. We used the PyMC3 Python package for model specification and fitting [29]. Model fitting was performed using training data.

We used $Z=5$ logistic bases, with ℓ_z equally spaced between 0 and 10 and $\sigma^2_{\text{basis}} = 1$. We then fitted a series of models:

1. σ_α, σ_γ, and C were fixed at 0.05, 5.0, and the identity matrix $I_{K \times K}$, respectively. For tuning the sampling parameters, we obtained 300 samples, which were then discarded. The following 300 samples were used for model parameter estimation.
2. Next, we removed the constraint fixing the correlation matrix C. Samples obtained from the previous step were used to initialize ω_k, λ_k, α_i, β_i, π_k, and missing biomarker observations y_{ijk}. We continued to fix σ_α=0.05 and $\sigma_\gamma = 5.0$. This model was fitted using the longitudinal data set, with 200 tuning + 200 samples.
3. Finally, we removed the constraints fixing σ_α and σ_γ. Samples obtained from the previous step were used to initialize all previously estimated parameters, and we fitted this model using 2000 tuning + 2000 samples.
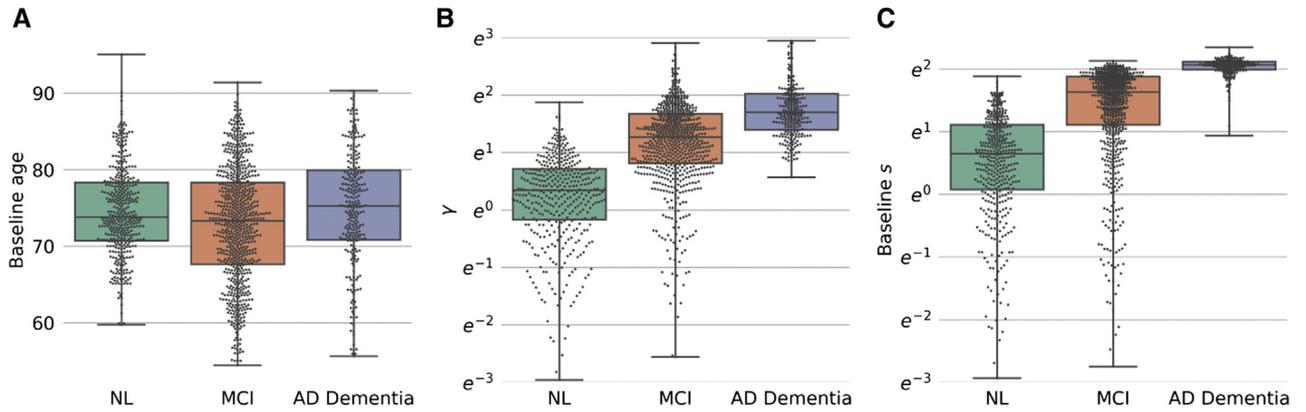
Fig. 2. Box and swarm plots of (A) baseline age, (B) estimated γ, and (C) estimated progression score s at baseline by baseline diagnosis for individuals in the training set. γ and s are plotted on a natural logarithm scale. All pairwise comparisons were statistically significant (all $P < 0.0013$), with the exception of baseline age comparison between NL and AD. Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment; NL, cognitively normal.

2.8. Correlates of progression scores

To understand correlates of the individualized PS estimated by our model, we performed a multiple linear regression analysis using the training data set investigating age, education, sex, apolipoprotein E (*APOE*) $\varepsilon 4$ status, and clinical diagnosis as predictors of PS at baseline.

2.9. Biomarker trajectories as a function of time from AD dementia

For each of the last $M = 200$ Monte Carlo iterations, we drew a sample $\tilde{s}_{AD}^{(m)}$ from the distribution of s at the first visit with an AD diagnosis among individuals who converted from NL or MCI to AD for the mth Monte Carlo iteration (excluding individuals who reverted) and computed the curve $u \mapsto \mathbf{g}(f(u + \tilde{t}(\tilde{s}_{AD}^{(m)}), \gamma_i^{(m)}); \omega^{(m)})$ for each individual i in the training set. The resulting MN biomarker curves, where N is the number of individuals in the training set, were used to characterize the mean curve and its 95% interval. To check if our estimated trajectories agreed with the time from onset values available in the data set, we plotted the computed curves, superimposed on a scattergram of the biomarker data versus observed time from onset.

2.10. Prediction of time to diagnosis

Age at the first occurrence of an AD dementia diagnosis for each individual who is NL or MCI at baseline was considered as the age at dementia onset. Using the known age at dementia onset data in the training set, we trained a linear regression model with time to dementia onset from baseline as the outcome and baseline PS as the independent variable.

To make predictions for time to dementia onset in the testing set, we first computed a PS for each individual at baseline given the trained PS model. The trained linear regression model was then applied to these baseline PSs to make predictions for time to dementia onset for each individual in the testing set.

We assessed the performance of our onset age estimation by computing the RMSE between the predicted and observed onset ages for individuals whose onset ages are known (i.e., for whom diagnostic conversion was observed in the longitudinal data set). This analysis is a biased reflection of the accuracy of onset age prediction given that it is restricted to those who convert to AD. To obtain measures based on all individuals regardless of their conversion status, we estimated Kaplan-Meier survival curves based on our predictions and observed onset ages. Survival curve estimation incorporates data for all individuals by assuming that AD conversion will occur after the last visit for individuals who did not convert during the study. For computing the Kaplan-Meier curves using observed data, event was defined as the first visit with an AD diagnosis following a visit with an NL or MCI diagnosis. We right-censored using age at last visit if the individual remained NL or MCI. For computing the Kaplan-Meier curves using predicted data, if onset was "predicted" to occur before baseline visit, age at onset was set to baseline age. We right-censored using age at last visit if the predicted onset age was greater than the age at last visit. We compared these two curves using the log-rank test χ^2 statistic.

We repeated the linear regression models using age, each biomarker, age + each biomarker, all biomarkers, and age + all biomarkers as independent variable(s). We also ran linear regression models using PS and its combinations with age and/or γ as independent variable(s). An intercept was included in all linear regression models. We then used permutation tests to compare the RMSE and χ^2 of the best model without PS to those of the best model with PS. The permutation test involved randomly swapping the onset ages predicted by the models being compared and computing the differences in RMSE and χ^2 between the two models. This was repeated 2000 times to obtain a distribution for RMSE difference and a distribution for χ^2 difference under the null hypothesis of equality between models. The observed difference was then quantified against this null distribution using a two-tailed test.

Table 2
Result of multiple linear regression analysis investigating the associations of age, sex, education, *APOE ε4* genotype, and clinical diagnosis with estimated progression score *s* at baseline

Independent variable	Estimate	SE	*t*-statistic	*P* value
Intercept	−3.72	0.545	−6.83	<0.001
Age	0.0869	0.006	14.1	<0.001
Male sex	0.303	0.089	3.40	0.001
Education	−0.0619	0.016	−3.98	<0.001
APOE ε4 heterozygous	0.719	0.094	7.63	<0.001
APOE ε4 homozygous	1.34	0.158	8.50	<0.001
MCI	2.45	0.103	23.9	<0.001
AD dementia	5.13	0.132	38.8	<0.001

Abbreviations: APOE, apolipoprotein E; AD, Alzheimer's disease; MCI, mild cognitive impairment; SE, standard error.

NOTE. Five individuals were excluded due to missing *APOE* information. 1364 observations were used to fit the model, yielding an R^2 = 0.635, an adjusted R^2 = 0.633, *F*-statistic = 336.7, and Prob(*F*-statistic) <0.0001.

Finally, we compared the prediction performance of PS computed using our model to the performance of "disease age" computed using two previously proposed models of AD progression (see Appendix C for details).

2.11. Reproducible research

We provide the code used for these analyses and all the software required to run it as a docker image accessible via https://hub.docker.com/r/bilgelm/bayesian-ps-adni/.

3. Results

Participant demographics are presented in (Table 1). Difference in the proportions of NL and MCI diagnoses at baseline between training and testing sets was not statistically significant (Fisher's exact test, *P* = 0.79). There were no statistically significant differences between training and testing sets at baseline by diagnostic category in age, sex, race, education, *APOE ε4* genotype, number of visits, or duration of follow-up.

3.1. Fitted progression score model

Assessment of the traceplots (Fig. 1) and the Geweke scores (Fig. 2) for the final model suggested convergence of the sampler. Estimated biomarker trajectories as a function of PS are presented in Fig. 1. The resulting temporal QT revealed early changes in verbal memory. The temporal evolution was also marked by changes in hippocampal volume and CSF Aβ_{1-42}, tau, and p-tau, as well as fluorodeoxyglucose PET and remaining cognitive measures.

3.2. Correlates of progression scores

Box plots of the estimated subject-specific variable γ and baseline *s* revealed differences among NL, mild cognitively impaired, and demented individuals. These group-wise differences were greater than those observed with age

(Fig. 2). Older age, male sex, fewer years of education, the number of *APOE ε4* alleles, and MCI or AD diagnoses were associated with higher PS at baseline (Table 2).

3.3. Time from diagnosis

Biomarker trajectories as a function of time from AD diagnosis are shown in Fig. 3.

Among prediction models that did not include PS as an independent variable, the model with the lowest χ^2 was the one with all 9 biomarkers ($\chi^2 = 20.5$, RMSE = 1.58), whereas additionally including age yielded the model with the lowest RMSE ($\chi^2 = 21.1$, RMSE = 1.57) (Table 3). Among models with PS, including age in addition to PS yielded the model with the lowest χ^2 ($\chi^2 = 1.68$, RMSE = 1.49), whereas the model with PS only had the lowest RMSE ($\chi^2 = 2.00$, RMSE = 1.48).

Results of age at AD onset prediction based on the linear regression model with PS and age are shown in Fig. 4. The Kaplan-Meier curve computed using predicted onset ages tracked the observed curve closely (Fig. 4B). We used the log-rank test to compare the curve based on predicted onset ages to the curve based on observed onset ages. The difference between the curves based on predicted and observed onset ages was not statistically significant based (log-rank test *P* = 0.2).

We compared the performance of the model with all 9 biomarkers to that of the model with PS + age because these models had the lowest χ^2 measures. Based on permutation testing, the difference in RMSE (*P* = 0.75) or χ^2 (*P* = 0.08) of the two models was not statistically significant. Unlike the predictions based on PS + age, predictions based on all 9 biomarkers yielded a survival curve that was different from the survival curve based on the observed ages (log-rank test, *P* < 0.0001).

Results of experiments investigating the predictive performance of disease age computed using two previously proposed models of AD progression are presented in the lower portion of Table 3. The performance of the disease age computed by Latent Time Joint Mixed effects Model [11] is very similar to that of age. The disease age computed using Gaussian Process Progression Model [15] results in a small RMSE and maximum absolute error, both of which are measures computed using only the individuals who convert; however, the survival curve–based measures, which take into account all individuals regardless of conversion, are poorer.

4. Discussion

We presented a model for aligning short-term longitudinal data across individuals to characterize long-term trajectories of a collection of biological and cognitive measurements that are implicated in AD. We applied our model to individuals who are NL or have MCI or AD dementia to estimate a QT of AD progression using the ADNI data set and recommendations presented in the Alzheimer's Disease Modelling Challenge. We
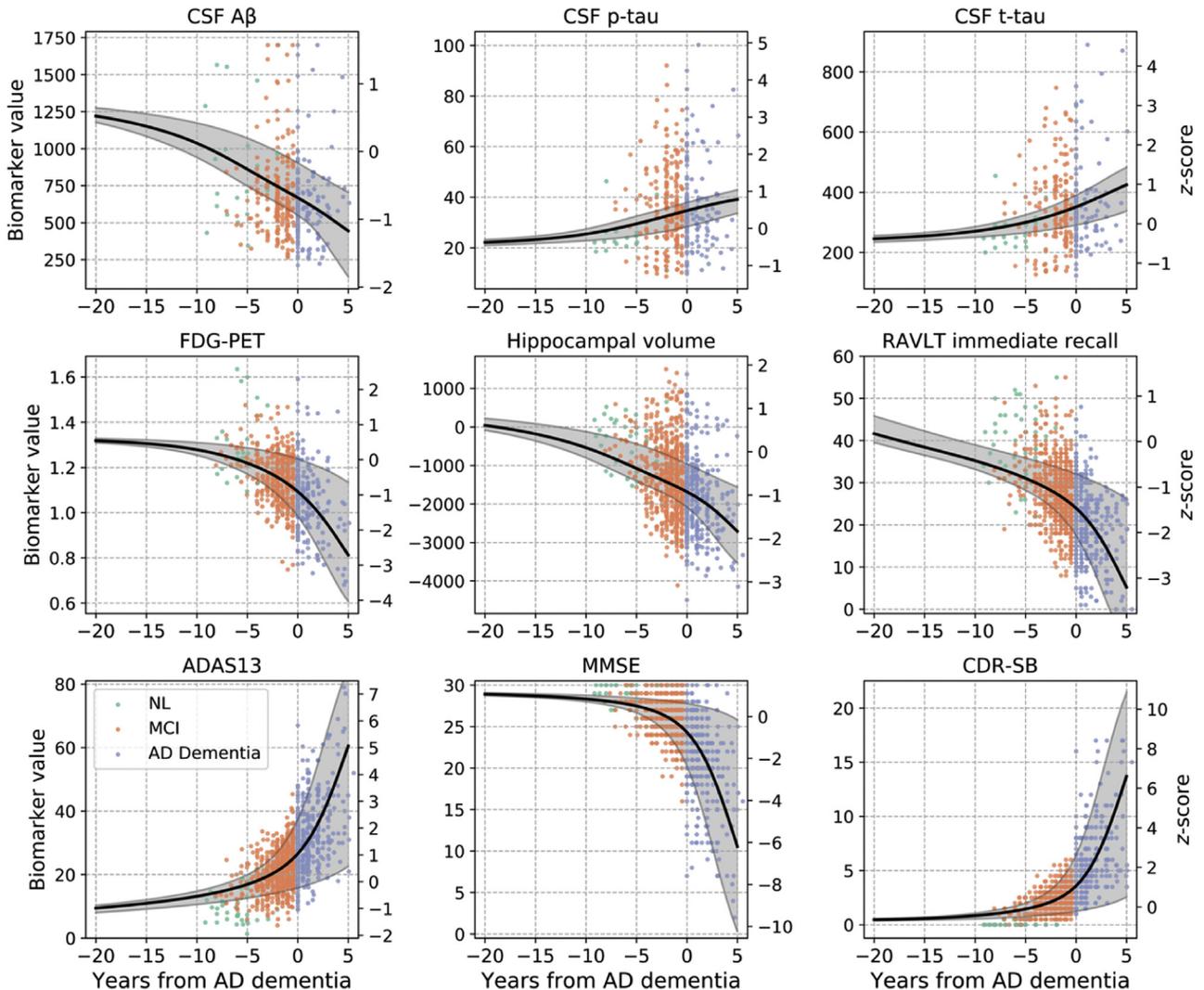
Fig. 3. Estimated biomarker trajectories as a function of time from initial AD dementia diagnosis are shown with the black curves, and the shaded areas depict 95% credible intervals. A value of 0 on the *x*-axis corresponds to the onset of AD dementia. Negative values are before the onset of AD dementia, and positive values are after the onset of AD dementia. Biomarker *z*-scores, shown on the right-hand-side *y*-axes, were computed using mean and standard deviations at baseline across 1369 individuals in the training set. Note that observed time from AD was not used in the estimation of the biomarker trajectories shown; trajectories were obtained using the model fit. Scattergrams of observations are shown as a function of observed time from AD, color-coded by diagnosis, to allow for a visual assessment of the agreement of the estimated trajectories with underlying data. Abbreviations: Aβ, amyloid β; CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; PET, positron emission tomography; RAVLT, Rey Auditory Verbal Learning Test; ADAS13, Alzheimer's Disease Assessment Scale-Cognitive 13-item scale; MMSE, Mini–Mental State Examination; CDR-SB, Clinical Dementia Rating-Sum of Boxes.

demonstrated that progression along our estimated temporal QT reflects AD processes by showing its association with clinical diagnosis and known AD risk factors, as well as by quantifying the accuracy of our model in predicting age at onset of AD dementia. The ability of our model to leverage short-term data to obtain long-term biomarker trajectories is highly relevant to the current need in clinical trial design for predicting who will develop cognitive impairment and AD dementia over time.

We found that the PS computed based on AD-related neuropathology, neurodegeneration, and cognitive measures were associated with age, *APOE* ε4 positivity, which is the most influential known genetic risk factor for sporadic AD [30], and clinical diagnoses of MCI

and AD. We also observed higher PS among men and a negative association between PS and years of education. Associations with *APOE* ε4 and clinical diagnoses provided evidence that our QT, which was constructed without these pieces of information, is reflective of AD progression. Although we recognize that there is heterogeneity in individual presentations across the biomarkers considered in our model and that not every individual included in our analyses is necessarily on an AD pathway, the strong associations we observed between PS and AD diagnosis as well as AD risk factors suggest that the estimated trajectories are mainly indicative of age- and disease-related changes that occur along the progression toward AD dementia.

Table 3
Prediction results based on linear regression models, presented for a select subsample of investigated models

Independent variable(s)	Measures based on those who convert		Survival curve–based measures	
	RMSE	Max. abs. err.	χ^2	Max. vert. dist.
Age	1.80	6.78	46.6	0.491
RAVLT imm.	1.64	5.52	35.7	0.438
RAVLT imm. + age	1.64	5.46	34.5	0.437
All biomarkers	1.58	4.67	20.5	**0.428**
All biomarkers + age	1.57	4.63	21.1	0.437
PS	**1.48**	3.79	2.00	0.453
PS + γ	**1.48**	3.75	1.72	0.453
PS + age	1.49	**3.70**	**1.68**	0.453
Comparison to disease age estimated from other methods				
LTJMM [11], $t_{i,j=0} + \delta_i$	1.80	6.77	47.7	0.492
GPPM [15], $\phi^j(\tau) = \tau + d^j$ (subset)	2.20	5.80	18.3	0.466
This article, PS (subset)	2.14	4.21	0.685	0.200

Abbreviations: RMSE, root mean square error; Max. abs. err., maximum absolute error; χ^2, log-rank test statistic; Max. vert. dist., maximum vertical distance between the survival curves based on predicted onset and observed onset; RAVLT imm., Rey Auditory Verbal Learning Test immediate recall (sum across learning trials); PS, progression score.

NOTE. The lower portion of the table presents the predictive performance achieved using disease ages estimated using two existing models of AD progression, Latent Time Joint Mixed effects Model (LTJMM) and Gaussian Process Progression Model (GPPM). Because the GPPM model had to be fitted on a subset of the data, for comparison purposes, we also present the predictive performance of PS computed using our model in this same subset.

Our QT was consistent with our previous findings in the ADNI [6]. By modifying our previously described progression score models [6,7] to include a global parameter governing the relationship between disease stage and rate of progression, we were able to depict biomarker trajectories as a function of time from AD diagnosis. This innovation enabled a more easily interpretable characterization of the natural history of AD starting in preclinical stages.

Our method achieved <1.5-year RMSE in predicting age at onset of AD over the course of 0.48 to 9.0 years. It is difficult to compare the performance of our prediction with results reported in the literature based on previous methods given differences in samples and model features used. Oxtoby et al. [5] reported an RMSE of 1.3 years for predicting age at onset over the course of 3.0 years in the context of dominantly inherited AD based on CSF, MRI, and PET biomarkers. Vogel et al. [22] reported an R^2 value of 0.15 using the ADNI data set for predicting years to conversion to MCI or AD based on functional and structural MRI measures. We fitted two previously proposed models and compared the performance of disease age computed by these models in predicting age at onset of AD in the same data set to the performance of PS computed from our model. PS computed using our model achieved the best predictive performance under every measure. It is necessary to continue model development in this area to enable more accurate onset age predictions over longer time intervals.

A limitation of our model is the simplifying assumption of a single pathway of biomarker changes from a NL state to AD dementia. It is known that there is heterogeneity in longitudinal progression in AD. Studying individuals who deviate from the estimated QT may be informative in understanding this heterogeneity, and models that relax the assumption of a single
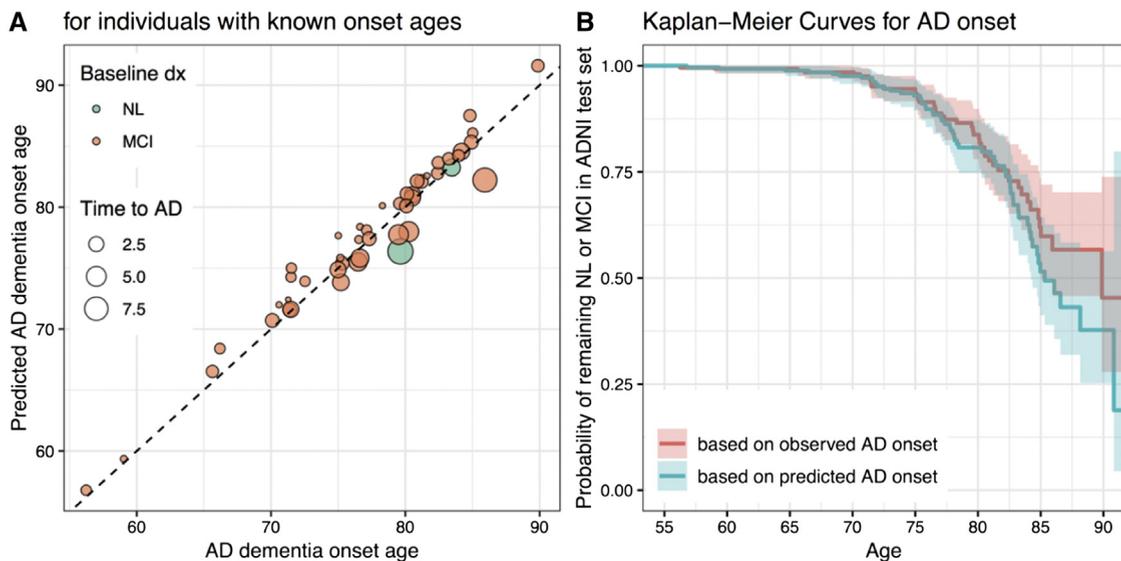

Fig. 4. (A) AD dementia onset ages predicted using the linear regression model with PS + age vs. observed AD dementia onset age (for individuals with known onset ages in the testing set). Time between baseline age and AD onset age (indicated by the size of the markers) varied between 0.48 and 9.0 years (median 1.6, IQR 1.0-3.0). There were 18, 15, 10, and 3 participants whose true time to diagnosis was in the interval (0.4,1.5), (1.5,2.5), (2.5,5.0), and (5, 9.1), respectively. The RMSEs corresponding to these ranges were 1.64, 1.06, and 0.84, 3.14. (B) Kaplan-Meier curves based on observed (red) and predicted (blue) AD dementia onset ages.

progression pathway will be necessary for a detailed analysis. We observed the greatest variability around trajectory estimates for CSF Aβ_{1-42}. This suggests that the timing of CSF Aβ_{1-42} can vary at the individual level in relation to changes in other biomarkers included in this study. Another limitation of our study is that higher PS indicates both age- and AD-related changes in the examined biomarkers. Understanding changes that occur with AD as distinct from age-related changes is clinically important. However, this is a challenging goal given that there is no scientific consensus regarding the definition of healthy aging. Several studies have approached this problem by including age and other demographic variables as covariates in addition to a disease stage variable [11,31,32] or by regressing out their effects before model fitting [5,33,34], whereas others have not included covariates to characterize trajectories that explain the natural history of AD dementia, including biological and cognitive changes that occur due to aging in addition to AD [4,6–9,13,15]. We formulated our goal with this study as the characterization of the natural history of AD dementia, including biological and cognitive changes that occur due to aging in addition to AD pathology, and therefore did not include an adjustment for age or any other demographic variables.

In conclusion, our method allows for the estimation of individualized latent disease progression indicators and population-level biomarker trajectories from longitudinal data. The estimated temporal QT of AD provides a mechanism for localizing individuals along biomarker trajectories based on multivariate data. Individualized PSs can be used as longitudinal composites to investigate associations with risk factors and outcomes. Furthermore, the ability to obtain individualized estimates of age at AD onset can allow for better participant recruitment in clinical trials aimed at preclinical AD.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.dadm.2019.01.005.

RESEARCH IN CONTEXT

1. Systematic review: The authors performed searches using PubMed and Google Scholar to cover relevant literature featuring the following keywords: Alzheimer, dementia, progression, staging, statistical model, Bayesian model, predict, onset age, and age at onset. Relevant scientific literature citing and cited by papers found through these searches were also included in the systematic review.

2. Interpretation: Our model provided a quantitative template of Alzheimer's disease (AD) progression, demonstrating early changes verbal memory, hippocampal volume loss, and cerebrospinal fluid measures of amyloid and tau. Our method enabled prediction of AD dementia onset age with a root mean square error of <1.5 years, and predicted onset ages yielded a survival curve that closely matched the survival curve based on observed onset ages.

3. Future directions: Future work should investigate how the proposed progression score model can be operationalized to enable decisions regarding participant recruitment, monitoring, and treatment in research studies and clinical trials.

References

[1] Villemagne VL, Burnham S, Bourgeat P, Brown B, Ellis KA, Salvado O, et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. Lancet Neurol 2013;12:357–67.

[2] Bateman RJ, Xiong C, Benzinger TLS, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. N Engl J Med 2012;367:795–804.

[3] Fonteijn HM, Modat M, Clarkson MJ, Barnes J, Lehmann M, Hobbs NZ, et al. An event-based model for disease progression and its application in familial Alzheimer's disease and Huntington's disease. NeuroImage 2012;60:1880–9.

[4] Young AL, Oxtoby NP, Daga P, Cash DM, Fox NC, Ourselin S, et al. A data-driven model of biomarker changes in sporadic Alzheimer's disease. Brain 2014;137:2564–77.

[5] Oxtoby NP, Young AL, Cash DM, Benzinger TLS, Fagan AM, Morris JC, et al. Data-driven models of dominantly-inherited Alzheimer's disease progression. Brain 2018;141:1529–44.

[6] Jedynak BM, Lang A, Liu B, Katz E, Zhang Y, Wyman BT, et al. A computational neurodegenerative disease progression score: Method and results with the Alzheimer's Disease Neuroimaging Initiative cohort. NeuroImage 2012;63:1478–86.

[7] Bilgel M, Prince JL, Wong DF, Resnick SM, Jedynak BM. A multivariate nonlinear mixed effects model for longitudinal image analysis: Application to amyloid imaging. NeuroImage 2016;134:658–70.

[8] Marinescu RV Eshaghi A, Lorenzi M, Young AL, Oxtoby NP, Garbarino S, Shakespeare TJ, et al. A vertex clustering model for disease progression: Application to cortical thickness images. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap PT, et al., eds. Lecture Notes in Computer Science, Information Processing in Medical Imaging;10265. Springer; 2017. p. 134–45. ISBN 9783319590493.

[9] Schiratti JB, Allassonniere S, Routier A, Colliot O, Durrleman S. A mixed-effects model with time reparametrization for longitudinal univariate manifold-valued data. In: Ourselin S, Alexander DC, Westin CF, Cardoso MJ, eds. Lecture Notes in Computer Science, Information Processing in Medical Imaging;9123. Springer; 2015. p. 564–75. ISBN 3-540-54246-9.

[10] Donohue MC, Jacqmin-Gadda H, Le Goff M, Thomas RG, Raman R, Gamst AC, et al. Estimating long-term multivariate progression from short-term data. Alzheimers Dement 2014;10:S400–10.

[11] Li D, Iddi S, Thompson WK, Donohue MC. Bayesian latent time joint mixed effect models for multicohort longitudinal data. Stat Methods Med Res 2017:1-11; https://doi.org/10.1177/0962280217737566. arXiv:1703.10266.

[12] Guerrero R, Schmidt-Richberg A, Ledig C, Tong T, Wolz R, Rueckert D. Instantiated mixed effects modeling of Alzheimer's disease markers. NeuroImage 2016;142:113–25.

[13] Koval I, Schiratti JB, Routier A, Bacci M, Colliot O, Allassonnière S, et al. Spatiotemporal propagation of the cortical atrophy: Population and individual patterns. Front Neurol 2018;9:235.

[14] Ishida T, Tokuda K, Hisaka A, Honma M, Kijima S, Takatoku H, et al. A novel method to estimate long-term chronological changes from fragmented observations in disease progression. Clin Pharmacol Ther 2019;105:436–47.

[15] Lorenzi M, Filippone M, Frisoni GB, Alexander DC, Ourselin S. Probabilistic disease progression modeling to characterize diagnostic uncertainty: Application to staging and prediction in Alzheimer's disease. NeuroImage 2017; https://doi.org/10.1016/j.neuroimage.2017.08.059 [Epub ahead of print].

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

[17] Moradi E, Pepe A, Gaser C, Huttunen H, Tohka J. Machine learning framework for early MRI-based Alzheimer's conversion prediction in MCI subjects. NeuroImage 2015;104:398–412.

[18] Long X, Chen L, Jiang C, Zhang L. Prediction and classification of Alzheimer disease based on quantification of MRI deformation. PLoS One 2017;12:1–19.

[19] Lebedev AV, Westman E, Van Westen GJ, Kramberger MG, Lundervold A, Aarsland D, et al. Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. NeuroImage: Clin 2014;6:115–25.

[20] Wang T, Qiu RG, Yu M. Predictive Modeling of the Progression of Alzheimer's Disease with Recurrent Neural Networks. Scientific Rep 2018;8:1–12.

[21] Mathotaarachchi S, Pascoal TA, Shin M, Benedet AL, Kang MS, Beaudry T, et al. Identifying incipient dementia individuals using machine learning and amyloid imaging. Neurobiol Aging 2017; 59:80–90.

[22] Vogel JW, Vachon-Presseau E, Binette AP, Tam A, Orban P, Joie RL, et al. Brain properties predict proximity to symptom onset in sporadic Alzheimer's disease. Brain 2018;141:1871–83.

[23] Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.). [The psychological examination in cases of traumatic encephalopathy. Problems.]. Arch de Psychol 1941; 28:215–85.

[24] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98.

[25] Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's disease assessment scale that broaden its scope. Alzheimer Dis Associated Disord 1997;11:S13–21.

[26] Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules; 1993.

[27] Lewandowski D, Kurowicka D, Joe H. Generating random correlation matrices based on vines and extended onion method. J Multivariate Anal 2009;100:1989–2001.

[28] Hoffman MD, Gelman A. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. J Machine Learn Res 2014; 15:1593–623.

[29] Salvatier J, Wiecki T, Fonnesbeck C. Probabilistic programming in Python using PyMC. Peerj Comput Sci 2016;2.

[30] Corder EH, Lannfelt L, Bogdanovic N, Fratiglioni L, Mori H. The role of APOE polymorphisms in late-onset dementias. Cell Mol Life Sci 1998;54:928–34.

[31] Younes L, Albert M, Miller MI. Inferring changepoint times of medial temporal lobe morphometric change in preclinical Alzheimer's disease. NeuroImage: Clin 2014;5:178–87.

[32] Iturria-Medina Y, Sotero RC, Toussaint PJ, Mateos-Perez JM, Evans AC, Initiative TADN. Early role of vascular dysregulation on late-onset Alzheimer's disease based on multifactorial data-driven analysis. Nat Commun 2016;7:11934.

[33] Bilgel M, Koscik R, An Y, Prince J, Resnick S, Johnson S, et al. Temporal order of Alzheimer's disease-related cognitive marker changes in BLSA and WRAP longitudinal studies. J Alzheimer's Dis 2017; 59:1335–47.

[34] Oxtoby NP, Garbarino S, Firth NC, Warren JD, Schott JM, Alexander DC. Data-driven sequence of changes to anatomical brain connectivity in sporadic Alzheimer's disease. Front Neurol 2017; 8:580.