



Published in final edited form as:

Stat Methods Med Res. 2021 January ; 30(1): 316–330. doi:10.1177/0962280220948097.

A Bayesian hierarchical change point model with parameter constraints

Hong Li¹, Andreana Benitez², Brian Neelon¹

¹Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, USA

²Department of Neurology and Radiology, Medical University of South Carolina, Charleston, SC, USA

Abstract

Alzheimer's disease is the leading cause of dementia among adults aged 65 or above. Alzheimer's disease is characterized by a change point signaling a sudden and prolonged acceleration in cognitive decline. The timing of this change point is of clinical interest because it can be used to establish optimal treatment regimens and schedules. Here, we present a Bayesian hierarchical change point model with a parameter constraint to characterize the rate and timing of cognitive decline among Alzheimer's disease patients. We allow each patient to have a unique random intercept, random slope before the change point, random change point time, and random slope after the change point. The difference in slope before and after a change point is constrained to be nonpositive, and its parameter space is partitioned into a null region (representing normal aging) and a rejection region (representing accelerated decline). Using the change point time, the estimated slope difference, and the threshold of the null region, we are able to (1) distinguish normal aging patients from those with accelerated cognitive decline, (2) characterize the rate and timing for patients experiencing cognitive decline, and (3) predict personalized risk of progression to dementia due to Alzheimer's disease. We apply the approach to data from the Religious Orders Study, a national cohort study of aging Catholic nuns, priests, and lay brothers.

Keywords

Alzheimer's disease; Bayesian inference; change point model; parameter constraints; block Metropolis-Hastings; personalized risk prediction

1 Introduction

As adults age, they undergo subtle cognitive changes that may not initially manifest as clinical symptoms. These cognitive changes are often gradual, making them difficult

Corresponding author: Hong Li, Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC 29425, USA. liho@musc.edu.

Supplemental material

Supplemental material for this article is available online.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

to distinguish at first from more debilitating cognitive decline. However, as individuals approach dementia, cognitive impairment becomes more obvious, and the decline in cognitive function begins to accelerate. Once dementia sets in, the cognitive decline pattern differs markedly from the pattern in normal aging adults.^{1–3} Alzheimer’s disease (AD) is the leading cause of dementia among adults aged 65 or above, resulting in progressive memory loss, impaired thinking and disorientation, as well as changes in personality and mood. AD is marked histologically by the degeneration of brain neurons, primarily in the cerebral cortex, and by the presence of neurofibrillary tangles and plaques containing betaamyloid.⁴

Physiologically, AD presents as a continuum of syndromes ranging from no cognitive impairment (NCI) to mild cognitive impairment (MCI), and finally to dementia.⁵ There is heterogeneity in the onset of the disease: some individuals may experience cognitive decline early, while for others, the onset of disease occurs later in life.¹ There are a number of other defining features associated with the disease. First, for those who develop AD-related dementia, it is believed that an accelerated decline in cognitive function will occur at a certain time point (or “change point”) during the course of cognitive decline, leading ultimately to dementia.^{5,6} However, the timing of the change point is highly variable and is influenced by patient-specific factors, such as age, family history, and lifestyle factors.^{7,8} It is therefore critical to monitor the occurrence of this acceleration so that clinicians and care providers can pursue timely treatments.⁹ Second, the slope before the change point varies: some adults may show a linear cognitive decline prior to the point of accelerated decline, while others may show relatively stable performance prior to the change point.¹⁰ Third, the slope after the change point is nonincreasing, as cognitive decline progresses more rapidly following the change point.⁵ An appropriate statistical model should therefore take into account the above features. In particular, the model should restrict the change in slope following the change point to be nonpositive, with no change in slope occurring only for adults who do not experience or only temporarily experience cognitive decline.

Numerous random effect change point models have been formulated to examine the acceleration of cognitive change.^{5,6,9,11–13} However, this prior research has focused primarily on methods to characterize the decline in cognitive function for diseased patients only, without attempting to distinguish normal aging from diseased cognitive decline, or to predict the likelihood of cognitive decline in advance of clinical diagnosis. Ji et al.¹⁴ recognized that not every older adult experiences diseased cognitive decline and developed a hypothesis testing approach to assess whether a biomarker change point occurred during long-term follow-up. If the hypothesis test confirmed the presence of a change point, the authors then fit a general bilinear model to estimate a global change point time for all individuals. However, this approach did not allow for subject-specific change points. Slate and Turnbull¹⁵ used a subject-specific change point model to model the growth of prostate-specific antigen. The authors restricted the slope after the change point to be larger than a fixed, positive lower bound. The purpose of this restriction was to aid model identifiability by ensuring that the slopes before and after the change point could be distinguished from one another during estimation. However, the authors assumed a separate normal distribution, independent of other model parameters, for the slope after the change point, which may be restrictive in some settings. In our application, for example, patients with low cognitive scores at baseline will likely have more precipitous declines following

AD onset compared to patients with higher baseline scores. In addition, Slate and Turnbull¹⁵ assumed a prespecified value for the lower bound of the slope following the change point rather than allowing this to be estimated from the data. Through simulations, we show that fixing the threshold a priori can lead to poor inference if the threshold parameter is misspecified, which is likely in practice.

In this article, we develop a flexible model to (1) distinguish normal aging patients from patients with diseased cognitive decline who are not yet formally diagnosed as dementia, which will help to identify the profile of healthy cognitive aging and diseased cognitive aging, (2) estimate subject-specific change points and rates of cognitive decline, which will help to characterize cognitive decline, and (3) develop a prediction model to assist with subject-specific disease prognoses to identify patients at high risk of advancing to dementia.

The gold standard diagnosis of AD requires brain autopsy, whereby a neuropathologist diagnoses AD if two pathognomonic signs of AD (amyloid plaques and neuritic tangles) are found in brain tissue. In living patients, the diagnosis of AD is typically performed through clinical evaluation.¹⁶ However, as recent research shows, the sensitivity of the clinical diagnosis ranges from 70.9% to 93%, and the specificity ranges from 44.3% to 91%.^{17,18} In addition, physicians often feel uncomfortable in providing a timely diagnosis of AD due to the limitation of the current diagnostic approach.¹⁹ Hence, there is a need to develop a prediction tool to predict the risk of disease progression to dementia in the next several years without increasing healthcare costs. While the clinical practicality of such a tool may be limited at present to research studies that collect large amounts of longitudinal data, such a tool might eventually assist clinicians (e.g. geriatricians, geriatric psychiatrists, neurologists, neuropsychologists) in deciding whether to implement treatment and care management plans in the hopes of delaying dementia onset and improving quality of life. Furthermore, this tool would aid patients in planning ahead while they are still able to make important decisions regarding their care, support, and financial or legal matters. In this study, we focus on using data from a neuropsychological test battery since these are the primary tools for evaluating cognitive decline. These measurements are comprehensive, less invasive, and provide the information needed to determine the syndromal cognitive stage. Although there are other biomarkers and approaches (e.g. neuroimaging) available, the majority are not available for routine use due to expense, risks, and the need for specialized equipment or laboratories.^{20,21} Hence, it is critical to develop statistical methods that can be easily implemented to improve the prognosis of cognitive decline and to help triage patients when deciding whether further diagnostic biomarker testing is necessary.

To address these goals, we propose a Bayesian change point model for the analysis of cognitive decline for older adults. Our proposed model has several attractive features. First, it incorporates correlated subject-specific intercepts, slopes before change points, and slopes after change points. Second, it constrains the difference in slope before and after change points to be nonpositive, since it is well documented that cognitive decline accelerates following the onset of AD.^{10,14} Third, it defines a clinically meaningful “null region” in which the difference in slope before and after the change point is effectively 0, representing NCI, where the threshold for this null region is estimated from the data rather than fixed at a prespecified value. While this threshold is similar in spirit to Slate and Turnbull,¹⁵

we use it in a fundamentally different way—namely, as a prognostic tool to determine the likelihood of dementia due to AD rather than as a device to ensure identifiable model parameters. Through simulation studies, we demonstrate that a data-driven approach to estimating the threshold leads to more accurate and precise parameter estimates. Fourth, the model allows for a random change point time for each individual. Finally, the model can make personalized prediction of the risk of progression to dementia due to AD as individuals age. For implementation, we propose an efficient Markov chain Monte Carlo (MCMC) algorithm that relies in large part on easily sampled Gibbs steps. Our approach can be useful for clinicians or researchers involved in large cohort AD research studies (such as those in AD research centers) as well as memory clinics where extensive longitudinal cognitive data are available.

In the next section, we describe the Religious Orders Study (ROS) dataset. Then, we present a Bayesian hierarchical random change point model with a parameter constraint and discuss the estimation approach. This is followed by the Simulation study section where we evaluate the model's performance using a simulation study. In the penultimate section, we apply the approach to the data of ROS and discuss the results. We summarize the approach in the final section.

2 Motivating dataset: The ROS

Our analysis was based on the ROS,¹⁶ a longitudinal clinical cohort study of aging Catholic nuns, priests, and lay brothers from more than 40 religious orders across the United States. Participants without known dementia were enrolled into the study, and data were collected annually for more than 20 years. Clinical evaluation of cognitive function was performed annually, and patients were categorized as NCI, MCI, or dementia due to AD (hereafter referred to as “AD”) at each annual visit. Clinical diagnoses were performed by assessment of cognitive impairment by neuropsychologists and determination of disease stages by clinical consensus using standard criteria.⁵ In general, patients' disease stage could switch between NCI and MCI over the course of the study. However, because AD is currently an irreversible condition,²² it was unlikely for disease stage to return to either MCI or NCI once a patient was diagnosed as AD unless there was a misdiagnosis. The investigators collected measures of cognitive function (e.g. episodic memory, semantic memory, and working memory), motor function (e.g. manual strength, manual dexterity, and gait), disabilities and blood tests, genetic risk factors, and demographic and psychological traits. For our analysis, we used a well-validated, global measure of cognitive function (“global cog”) as the outcome, which is a composite score of cognitive measures.⁴ This composite score is derived from 19 tests that cover five cognitive domains, namely episodic memory, semantic memory, working memory, perceptual speed, and visuo-spatial ability. The composite scores are created by converting each test to a *z*-score and averaging the *z*-scores from the five domains. Higher scores indicate better cognitive function. Detailed information about the individual tests and the derivation of the composite measure has been previously reported.^{16,23}

Figure 1 shows 3 panels of global cog functional trajectories for a 20% random sample of ROS patients, resulting in an analytic sample of 189 patients. We chose subsample of

patients for illustrative purpose and to expedite the case study presented in the Application to ROS data section. This random sample was more or less equally split between NCI, MCI, and AD patients, where disease stage was defined according to the diagnosis at the final clinical visit. Of these 189 patients, 13.5% of the patients reversed their disease stage from AD to MCI or NCI, 32.8% patients switched between MCI and NCI, 30.7% patients stayed at NCI, 11.6% progressed from NCI to MCI to AD, 3.7% transitioned from NCI to MCI, and 6.3% transitioned directly from NCI to AD based on clinicians' assessments over the course of the study. The global cog scores ranged from -4 to 2 in the entire available ROS data. As expected, participants with NCI tended to have higher scores compared to patients with MCI and AD, with the mean score at the last clinical visit of 0.44 , while MCI patients generally had higher scores than AD patients with the mean score of -0.24 . For AD patients, the mean score at the last clinical visit is -1.75 . Although no participants had known AD-type dementia at enrollment, over the course of the study, some participants appeared to develop disease more rapidly than others. The left panel shows global cog scores for the NCI participants. The scores fluctuated between -1 and 2 , but overall, there was no downward trend in the trajectories. For MCI patients (the middle panel), the global cog measurements fluctuated between -2 and 1 , suggesting increased cognitive impairment, with a slight downward trend over time. For AD patients (the right panel), the global cog measurements fluctuated between -4 and 1 and generally showed a sharp decline at some point during the study, potentially heralding the onset of AD.

3 Model and estimation

3.1 Model

In this section, we describe a Bayesian constrained random change point model for cognitive functional trajectories. For each individual, we estimate if and when a change point occurs, as well as the slope before and after the change point. Our model allows for individuals to have unique random intercepts, random slopes before the change points, random change point times, and random slopes after the change points. The slope difference before and after the change point is constrained to be nonpositive, as we expect a more precipitous decline following the change point. However, we do not constrain the slope before the change point since its pattern varies.¹⁰ Thus, our model enables us to achieve two goals simultaneously: (1) to identify whether an individual's cognitive trajectory experiences a change point during the follow-up and (2) to estimate the rate and timing of rapid cognitive decline if a change point occurs.

We consider a piecewise-linear random change point model characterized by potentially different slopes before and after a random change point

$$Y_{ij} = \beta_{1i} + \beta_{2i}t_{ij} + \beta_{3i}(t_{ij} - \kappa_i)_+ + z'_{ij}\beta_{4i} + \epsilon_{ij} \quad (1)$$

where for individual $i = 1, \dots, n$ at follow-up visit $j = 1, \dots, n_i$, Y_{ij} is the continuous outcome measure (e.g. global cog), t_{ij} is the measurement time, z_{ij} is a $m \times 1$ vector of individual level predictors (i.e. age and education), $w_+ = w$ if $w \geq 0$ and 0 otherwise, κ_i denotes an unknown random change point, and ϵ_{ij} is an error term, which follows a normal $N(0, \sigma_\epsilon^2)$ distribution;

alternative distributions, such as the skew normal or finite mixture models, could be used to accommodate additional features of the data, such as skewness or multimodality.

Because AD onset has a high degree of heterogeneity in its progression, we use subject-specific parameters for all terms in the model. For individual i , $\beta_1 i$ denotes the random intercept, $\beta_2 i$ is the random slope before the change point k_i , and $\beta_3 i$ is the difference in random slopes before and after the change point k_i . For brevity, we refer to $\beta_3 i$ below as simply the “slope decrement” after the change point k_i . As previous research has shown, cognitive decline accelerates after the change point for AD patients, but cognitive trajectories are relatively stable for normal aging patients. Hence, we assume that the slope decrement $\beta_3 i$ is less than or equal to zero, which ensures clinically and statistically meaningful estimation for the slope after the change point. The parameter β_{4j} is the $m \times 1$ vector of coefficients corresponding to the predictors z_{ij} . The conditional likelihood function of $Y = (y_{ij})$ given all parameters and data $\mathbf{x}_{ij} = (1, t_{ij}, (t_{ij} - \kappa_i)_+, z_{ij})'$ and $\mathbf{X} = (\mathbf{x}_{ij})$, which is an $N \times q$ matrix, where $N = \sum_{i=1}^n n_i$, is

$$L(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\kappa}, \sigma_\epsilon^2, \mathbf{X}) = \prod_{i=1}^n \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp\left\{-\frac{1}{2\sigma_\epsilon^2}(y_{ij} - \mathbf{x}_{ij}\boldsymbol{\beta}_i)^2\right\}$$

where $\boldsymbol{\kappa} = (\kappa_1, \kappa_2, \dots, \kappa_n)'$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2, \dots, \boldsymbol{\beta}'_n)'$ in which $\boldsymbol{\beta}_i = (\beta_{i1}, \beta_{i2}, \beta_{i3}, \boldsymbol{\beta}'_{i4})'$ is a $q \times 1$ vector of regression parameter. The conditional likelihood can be maximized subject to the constraint $\beta_{3i} \leq 0$ for $i = 1, 2, \dots, n$. However, calculation of the maximum likelihood estimator subject to constraints is often difficult,²⁴ since estimates of β_{3i} for individuals without cognitive impairment fall on the boundary of the parameter space (i.e. $\hat{\beta}_{3i}$ is equal to or close to 0). Hence, we use Bayesian methods for parameter estimation, since order constraints can be readily incorporated into the prior distributions for the model parameters.

3.2 Prior specification

For normal aging individuals, the slopes after the change points may be minimal, in which case we can ignore their effects. However, it is not practical to assume the slope decrement is exactly equal to zero. Instead, we introduce a “null region” within which the effect is negligible, representing stable cognitive trajectories. We follow Neelon and Dunson²⁵ to select a prior distribution for $\boldsymbol{\beta}$ which achieves the following two goals simultaneously: (1) constrains the trajectory to be nonincreasing after the change point and (2) allows for a null region. Hence, we express the constrained slope decrement for patient i , β_{3i} , in terms of a latent, unconstrained slope decrement β_{3i}^* , namely, we set $\beta_{3i} = 11\{\beta_{3i}^* < \delta\}\beta_{3i}^*$, $i = 1, 2, \dots, n$, where δ is a threshold parameter representing clinically relevant cognitive decline, and $11\{\cdot\}$ is the indicator function. This representation facilitates posterior computation because it is often easier to sample from an unconstrained conditional distribution and apply a suitable constraint as part of the MCMC algorithm than it is to draw directly from a constrained posterior itself.²⁶ The parameter β_{3i}^* is an unconstrained latent slope decrement,

and d is a small negative threshold above which the slope decrement β_{3i}^* is effectively zero. Specifically, for individual i , if $\beta_{3i}^* < \delta$ then $\beta_i = \beta_i^* = (\beta_{1i}^*, \beta_{2i}^*, \beta_{3i}^*, \beta_{4i}^*)'$, which ensures feature (1); otherwise if $\delta \leq \beta_{3i}^* \leq 0$, then $\beta_i = \beta_{i0}^* = (\beta_{1i}^*, \beta_{2i}^*, 0, \beta_{4i}^*)'$ here, we set $\beta_{3i} = 0$ to ensure feature (2). We choose a conditionally conjugate MVN (β_0, Σ_β) prior for the unconstrained parameter β_i^* , where β_0 is the population mean, and Σ_β is an unstructured covariance matrix that accounts for dependence in the model parameters. This is relevant in our application, as we expect AD patients have much steeper slopes before and after change points compared to the MCI and NCI patients.⁵ More formally, given the threshold δ , the joint prior for (β, β^*) is

$$\begin{aligned} \pi(\beta, \beta^* \mid \delta, \Sigma_\beta, \beta_0) &= \prod_{i=1}^n \pi(\beta_i, \beta_i^* \mid \delta, \Sigma_\beta, \beta_0) \\ &= \prod_{i=1}^n \left(11\{\beta_{3i}^* < \delta\} 11\{\beta_i = \beta_i^*\} + 11\{\delta \leq \beta_{3i}^* \leq 0\} 11\{\beta_i = \beta_{i0}^*\} \right) \text{MVN}(\beta_0, \Sigma_\beta) \end{aligned} \quad (2)$$

where $\beta^* = (\beta_1^{*'}, \beta_2^{*'}, \dots, \beta_n^{*'})'$. Equation (2) represents a mixture of a point mass at 0 when $\delta \leq \beta_{3i}^* \leq 0$ and a truncated multivariate normal distribution (MVN) when $\beta_{3i}^* < \delta$. This mixture prior approach is similar to spike- and-slab variable selection and offers a computationally efficient alternative to dimension-switching approaches such as reversible jump.²⁷ Essentially, we fit a fully parameterized model that includes the decrement slope, β_{3i} , but allow β_{3i} to equal zero if the unconstrained slope, β_{3i}^* falls within the null region. Thus, the dimension of the model stays fixed, but the prior structure for β_{3i} is a mixture distribution whose sample space encompasses a point mass at zero. In this way, we partition the constrained parameter space for the slope decrement, β_{3i} , into two regions: (1) $[\delta, 0]$, corresponding to individuals whose β_{3i} s fall into this null region and who do not experience cognitive impairment and (2) $(-\infty, \delta)$, corresponding to individuals whose β_{3i} s fall into this rejection region and who experience AD-type cognitive decline. By using this joint prior in equation (2), we simultaneously estimate both the slope decrement β_{3i} , which guarantees the stable or downward trend in cognitive trajectories, and the threshold δ of the null region, which encourages flatness in the trajectories for patients without clinically relevant cognitive decline.

To complete the prior specification and ensure an identifiable model, we assign an MVN $(\mathbf{0}, \Omega)$ prior for the population mean β_0 . For Σ_β , we assign an inverse-Wishart $\text{IW}(v, \Gamma)$ prior where v is the degree of freedom. Throughout, we assume that Ω and Γ are known scale matrices (e.g. the $q \times q$ identity matrix, I_q , where q represents the dimension of β_0). We assign an inverse-gamma $\text{IG}(a_e, b_e)$ prior for σ_ϵ^2 . To be consistent with prior research,^{5,28} we assume $\kappa_i \sim N(\mu_\kappa, \sigma_\kappa^2)$ truncated between 3 and 8 for $i = 1, 2, \dots, n$ likewise, we assume $\delta \sim N(\mu_\delta, \sigma_\delta^2)$ truncated between -0.1 and 0 , which provides a clinically meaningful null region. Thus, unlike previous work involving change point models, we allow the threshold parameter to be estimated from the data rather than fixed a priori. Using simulation studies,

we demonstrate below that estimating d rather than fixing it a prespecified value generally improves inferences.

3.3 Posterior

For posterior computation, we adopt an efficient MCMC algorithm that combines Gibbs steps and Metropolis-Hastings steps. Conditional on the data $(y_{ij}, \mathbf{x}_{ij})$ $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, n_i$, the joint posterior density of parameters $\beta, \Sigma_\beta, \kappa, \sigma_\epsilon^2, \beta_0$, the latent variable β^* , and δ is proportional to

$$\begin{aligned} & \prod_{i=1}^n \prod_{j=1}^{n_i} N(y_{ij}; \beta_i, \sigma_\epsilon^2, \mathbf{x}_{ij}, \kappa_i) \pi(\beta_i, \beta_i^* | \delta, \Sigma_\beta, \beta_0) \pi(\beta_0) \pi(\delta) \pi(\kappa_i) \pi(\Sigma_\beta) \pi(\sigma_\epsilon^2) \\ &= \prod_{i=1}^n \prod_{j=1}^{n_i} N(y_{ij}; \beta_i, \sigma_\epsilon^2, \mathbf{x}_{ij}, \kappa_i) \pi(\beta_0) \pi(\delta) \pi(\kappa_i) \pi(\Sigma_\beta) \pi(\sigma_\epsilon^2) \\ & \quad \left(11\{\beta_{3i}^* < \delta\} 11\{\beta_i = \beta_i^*\} + 11\{\delta \leq \beta_{3i}^* \leq 0\} 11\{\beta_i = \beta_{i0}^*\} \right) \text{MVN}(\beta_0, \Sigma_\beta) \end{aligned}$$

where $\pi(\beta_i, \beta_i^* | \delta, \Sigma_\beta, \beta_0)$ is defined in equation (2), $\pi(\beta_0)$ is a prior for β_0 , $\pi(\delta)$ is a prior of δ , $\pi(\kappa_i)$ is a prior for the change point, $\pi(\Sigma_\beta)$ is a prior for Σ_β and $\pi(\sigma_\epsilon^2)$ is a prior for σ_ϵ^2 . The full conditional distributions for β_0, Σ_β , and σ_ϵ^2 have straightforward conjugate forms. We estimate $\kappa_i (i = 1, 2, \dots, n)$ using a Metropolis-Hastings step. Averaging k_i over the course of the MCMC algorithm results in a smooth posterior estimate of the mean egression function, reflecting our intuition that the change point is not instantaneous but instead represents a smooth transition (or “bent cable”) occurring over a short period of time.²⁹ Following Neelon and Dunson,²⁵ we update the threshold parameter d , the latent slope decrements $\beta_{3i}^* (i = 1, \dots, n)$, and the observed slope decrements $\beta_{3i} (i = 1, \dots, n)$ jointly using a Metropolis-Hastings step. Thus, after setting initial values, the sampler iterates through the following steps:

1. Sample σ_ϵ^2 from its inverse gamma full conditional
2. Sample Σ_β from its inverse-Wishart full conditional
3. Sample $\kappa_i^{new} (i = 1, 2, \dots, n)$ from an $N(\kappa_i^{old}, \sigma_\kappa^{*2})$ truncated between 3 and 8. Accept or reject the candidate κ_i^{new} using a random walk Metropolis-Hastings step
4. Sample β_0 from its multivariate normal full conditional
5. Update δ, β, β^* in one block as follows:
 - a. Sample a candidate value of δ^{new} from an $N(\delta^{old}, \sigma_\delta^{*2})$ truncated between -0.1 and 0

- b. For $i = 1, 2, \dots, n$, sample β_i^{*new} from its unconstrained MVN full conditional, and if $\beta_{3i}^{*new} < \delta^{new}$ then set $\beta_i^{new} = \beta_i^{*new}$; otherwise, if $\delta^{new} \leq \beta_{3i}^{*new} \leq 0$, then set $\beta_i^{new} = \beta_{i0}^{*new}$
 - c. Accept or reject the candidate δ^{new} , $\beta^{new} = (\beta_1^{new'}, \dots, \beta_n^{new'})'$ and $\beta^{*new} = (\beta_1^{*new'}, \dots, \beta_n^{*new'})'$ in one block with a Metropolis-Hastings step
6. Repeat steps 1 to 5 until convergence and calculate posterior summaries based on a large number of iterations.

MCMC convergence is monitored by trace plots, and the model fitting is assessed by standardized residual plots. The detailed MCMC algorithm (including priors and posteriors) is described in Online Appendix A. R code (<https://www.r-project.org>) for fitting the model is available from the first author.

3.4 Personalized prediction of disease progression to AD

In the following, we illustrate how to use the proposed approach to identify patients at high risk to progress to AD at various stages during the follow-up period. Our approach can be applied when extensive longitudinally measured cognitive data are available (such as AD research centers and some memory clinics). Additionally, clinicians with access to longitudinal data can use the model to predict the risk of progressing to AD and to adjust patients' treatment plans accordingly at each follow-up visit.

Suppose we wish to predict the probability of progression to AD for a new patient i at visit j . We first fit model 1) using the longitudinal cognitive measurements for patient i up to and including visit j (i.e. we observe only the first j measurements for patient i), together with the available data for the remaining patients. Let β_{3ij}^* denote the latent slope decrement following the change point for patient i at visit j , and let δ_j be the corresponding threshold indicating the null region boundary. At each MCMC iteration, we compare β_{3ij}^* with δ_j to decide whether β_{3ij}^* falls into the null region or the rejection region. We then estimate the probability, p_{ij} , of β_{3ij}^* falling inside the rejection region (i.e. $11\{\beta_{3ij}^* < \delta_j\} = 1$ at visit j over the course of the MCMC iterations. If individual i does not experience a change point at visit j , the majority of the β_{3ij}^* s will fall into the null region (i.e. the majority of $\beta_{3ij} = 0$) and the corresponding posterior mean probability \hat{p}_{ij} will be small; otherwise, if individual i does experience a change point at visit j , the majority of the β_{3ij}^* s will fall inside the rejection region (i.e. the majority of $\beta_{3ij} = \beta_{3ij}^*$), and the corresponding posterior mean probability \hat{p}_{ij} will be large. Hence, \hat{p}_{ij} provides an estimate of AD risk at visit j . By refitting the model for various choices of j , we can monitor the changes in the probability of progression to AD over the course of the follow-up.

4 Simulation study

4.1 Simulation model

To evaluate the proposed model, we conducted a simulation study, applying the method to 100 simulated patients in which 50 observations whose cognitive functional trajectories encounter change points (i.e. labeled “AD” patients) and the other 50 observations whose cognitive functional trajectories do not encounter change points (i.e. labeled “non-AD” patients). To illustrate how to fit the model with predictors, we added two risk factors: age at enrollment centered at 76 years old and years of education as a continuous variable. We generated data using the following model

$$Y_{ij} = \beta_{1i} + \beta_{2i}t_{ij} + \beta_{3i}(t_{ij} - \kappa_i)_+ + \beta_{4i}age_i + \beta_{5i}education_i + \epsilon_{ij} \quad (3)$$

where $i = 1, 2, \dots, 100$, $j = 1, 2, \dots, n_i$. We first made assumptions similar to the preliminary results from the ROS data,⁵ which we refer to Scenario 1. For each individual i , we generated the number of annual visits, n_i , from a discrete uniform $U[7, 15]$ distribution, the change point κ_i followed an $N(5, 1)$ truncated between years 3 and 8, the error term ϵ_{ij} followed an $N(0, 0.05)$, and the years of education followed a uniform distribution $U[15, 25]$. For observations with change points, we assumed that age followed an $N(85, 64)$, and the parameter $\beta_i (i = 1, 2, \dots, 50)$ followed an $MVN(\beta_0, \Sigma_\beta)$ in which the population mean $\beta_0 = (-0.18, -0.05, -0.3, -0.04, 0.04)'$. For observations without change points, we assumed that age followed an $N(72, 64)$, and the parameter $\beta_i (i = 51, \dots, 100)$ followed an $MVN(\beta_0, \Sigma_\beta)$ in which the population mean $\beta_0 = (-0.18, -0.05, 0, -0.04, 0.04)'$. The specification for Σ_β is provided in Online Appendix B, which resulted in heterogeneity similar to the ROS data. We assumed the threshold for the null region $\delta = -0.05$ and $\beta_3 = -0.3$ for AD patients to be consistent with the results in Yu et al.⁵ To assess the robustness of the model, we simulated two additional scenarios, denoted as Scenarios 2 and 3, in which we increased the variability in the data and changed the value of δ to accommodate different null regions (see Online Appendix B).

When fitting model (3), we assumed independent $MVN(\beta_0, \Sigma_\beta)$ priors for the random effects β_i^* , an $MVN(\mathbf{0}, \mathbf{I}_5)$ prior for β_0 , an $IG(0.001, 0.001)$ prior for σ_ϵ^2 , and an $IW(5, \mathbf{I}_5)$ prior for Σ_β . We ran the model for 20000 MCMC iterations with a burn-in of 5000, which was sufficient to ensure convergence based on trace plots and to verify that the normality assumption held based on the standardized residual plot.

4.2 Results

Table 1 presents the posterior means and 95% credible intervals (CIs) for the model parameters in Scenario 1, where $\delta = -0.05$, $\rho = 0.1$ and $\sigma_\epsilon^2 = 0.05$. The estimated population mean, $\hat{\beta}_0$, and the estimated threshold of the null region, $\hat{\delta} = -0.05$ (with acceptance rate of 60%), are close to the true parameter values, and the 95% CIs overlap with the true values used in the simulation.

Figure 2 shows the simulated and estimated trajectories. This figure clearly demonstrates the adequate model fit, with the estimated average trajectories overlapping with the simulated average trajectories very well for both AD and non-AD patients. Figure D1 in Online Appendix D presents the simulated trajectories (blue dot curves) and the estimated trajectories (red long-dash curves) for six individuals. The figure shows that the estimated trajectories closely mimic the simulated trajectories. Specifically, AD patients 5, 25, and 42 experience change points in which the $\hat{\beta}_{3s}$ are -0.17 , -0.25 , and -0.65 , respectively, which are outside the null region. The follow-up times for these patients are 14, 15, and 15 years. Non-AD patients 69, 70, and 93 do not experience change points during their follow-up periods, which ranged from 14 to 15 years. As expected, the $\hat{\beta}_{3s}$ for these individuals are -0.001 , -0.0003 , and -0.02 , which are inside the null region. Finally, the standardized residual plot in Figure D2 of Online Appendix D follows a standard normal distribution, indicating that the normality assumption holds.

To assess the robustness of the model performance given the nature of the disease and covariates in the model, we simulated additional scenarios, denoted as Scenarios 2 and 3 (details are provided in Online Appendix B). Scenario 2 corresponds to $\delta = -0.1$, Σ_{β} taking the same values as in Scenario 1, and $\sigma_{\epsilon}^2 = 0.5$, which implies greater within-subject variability relative to Scenario 1. Scenario 3 corresponds to $\delta = -0.05$, $\sigma_{\epsilon}^2 = 0.05$, $\rho = 0.2$ and the diagonal elements in Σ_{β} taking the values given in Scenario 1, leading to higher random effect correlation compared to Scenario 1. As shown in Table 1, for both scenarios, the estimated model parameters closely mimic the true values, with 95% CIs overlapping with the true values in all cases (Scenarios 2 and 3 in Table 1). Figure D3(a) and (b) in Online Appendix D, which is similar to Figure 2, presents estimated trajectories for Scenario 2 and 3, respectively. Both panels clearly demonstrate satisfactory model fit in which the estimated average trajectories overlap with the simulated average trajectories very well for all cases. Figure D4 in Online Appendix D presents estimated global cog trajectories for six patients under Scenario 2 (Panel (a)) and Scenario 3 (Panel (b)). Again, the simulated trajectories and the estimated trajectories mirror one another. Figure D5 in Online Appendix D presents the standardized residual plots for Scenario 2 (Panel (a)) and Scenario 3 (Panel (b)). Both residual plots follow a standard normal distribution indicating that the normality assumption holds.

Finally, we conducted simulations to compare our approach to that of Slate and Turnbull,¹⁵ which prespecifies the threshold parameter, δ , and assumes that the random effect β_{3i} is uncorrelated with other random effects in the model. We selected four settings for the correlation between $\beta_{3i}(i = 1, \dots, n)$ and other random effects in the model, denoted by ρ^* , and for the prespecified value of d that reflects an “informed guess” as to the true, unknown δ value:

- Setting 1a: $\rho^* = 0$, true $\delta =$ prespecified $\delta = -0.05$. This scenario most closely conforms to the approach of Slate and Turnbull.¹⁵
- Setting 1b: Same settings as in (1a) but with a misspecified δ of -0.03 under the approach of Slate and Turnbull.¹⁵

- Setting 2a: $\rho^* = 0.2$, true $\delta =$ prespecified $\delta = -0.1$. This scenario corresponds to a modest correlation between the random effects.
- Setting 2b: Same settings as in (2a) but with a misspecified δ of -0.15 under the approach of Slate and Turnbull.¹⁵

For each setting, we assumed $\beta_{01} = -0.18$, $\beta_{02} = -0.05$, $\beta_{03} = -0.3$, $\beta_{04} = -0.04$ and $\beta_{05} = 0.04$. For Settings a and 1b, where $\rho^* = 0$, we assumed that $(\beta_{1i}, \beta_{2i}, \beta_{4i}, \beta_{5i})'$ followed an MVN distribution with mean $(\beta_{01}, \beta_{02}, \beta_{04}, \beta_{05})'$ and covariance matrix taking the values of Scenario 1 in Online Appendix B. We further assumed an $N(\beta_{03}, \sigma_{33}^2)$ distribution for β_{3i} , where σ_{33}^2 is given in Scenario 1 of Online Appendix B. For Settings 2a and 2b, where $\rho^* = 0.2$, we assumed that $(\beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_{4i}, \beta_{5i})'$ followed an MVN distribution with mean $(\beta_{01}, \beta_{02}, \beta_{03}, \beta_{04}, \beta_{05})'$ and covariance matrix taking values given in Scenario 1 of Online Appendix B with the common correlation of ρ^* . These two settings have the assumed true $\delta = -0.1$, which is different from the above Scenario 3. We then compared the model of Slate and Turnbull¹⁵ to our approach in which δ was estimated as part of the MCMC algorithm.

Table C1 in Online Appendix C presents the posterior means and 95% CIs for the model parameters using both approaches. As expected, when d is incorrectly specified under the approach of Slate and Turnbull,¹⁵ the parameter estimates deviate from their true values, and the 95% CIs often fail to cover the true parameter values. These trends increase as the random effect correlation increases. In contrast, under the proposed approach, the parameter estimates are generally more accurate, and the 95% CIs are more precise. For example, under Setting 1b, where $\rho^* = 0$, the true $\delta = -0.05$, and the prespecified $\delta = -0.03$, which is only a small deviate from the true δ , under Slate and Turnbull¹⁵ approach, $\hat{\beta}_{03} = -0.21$ (95%CI = $[-0.24, -0.18]$), which fails to cover the true value of -0.30 . Although $\hat{\beta}_{02}$ and $\hat{\sigma}_e^2$ are not far away from the true values, their 95% CIs overlap with the true values at one of the boundaries. Conversely, under the proposed approach, $\hat{\beta}_{03} = -0.23$ (95%CI = $[-0.32, -0.14]$), and all other parameter estimates are quite accurate with correct posterior intervals around the true values. Even when we correctly specified the δ value, the approach of Slate and Turnbull¹⁵ did not always yield accurate estimates and correct posterior intervals, since it fails to consider the correlation between the random slope after the change point and other random effects in the model. For example, under Setting 2a, where $\rho^* = 0.2$ and δ is correctly specified at -0.1 , $\hat{\beta}_{03} = -0.36$ (95%CI = $[-0.41, -0.32]$), which fails to cover the true value of -0.30 . In addition, $\hat{\beta}_{02}$ and $\hat{\sigma}_e^2$ are away from the true values and the corresponding 95% CIs fail to cover the true values. In contrast, under the proposed method, $\hat{\beta}_{03} = -0.34$ (95%CI = $[-0.41, -0.20]$), which achieves the proper estimate and posterior interval encompassing the true value. All other parameter estimates are quite accurate with proper posterior intervals. Taken together, these results highlight the importance of estimating the threshold parameter during model fitting and taking into account the potential correlation among all model parameters.

5 Application to ROS data

Next, we applied model (3) to the random sample from the ROS dataset. The follow-up time for each individual ranged from 7 to 15 years. We also included age at enrollment (centered) and years of education in our analyses, as these are known predictors of cognitive decline.⁵ Among those 189 patients, the percentage of NCI, MCI, and AD patients diagnosed at the last clinical visit was 32%, 37%, and 31%, respectively. The average year of education was 15.5 years ranging from 5 to 26 years, and the average age at enrollment was 76.9 years old ranging from 56.7 to 91.3 years old.

5.1 Model fitting

We fit the model based on the estimation approach presented in the Model and estimation section. We assumed an $MVN(0; 1000 \times I_5)$ prior for β_0 , an $IG(0.001, 0.001)$ prior for σ_ϵ^2 , an $MVN(\beta_0, \Sigma_\beta)$ prior for β_i^* , and an $IW(5, 0.01 \times I_5)$ prior for Σ_β .

We ran the MCMC algorithm for 20,000 iterations with a burn-in of 5000. The acceptance rate for estimating κ_i ranged from 47% to 49%, and the acceptance rate for δ was 46%. The standardized residual plot (Panel (a) of Figure D6 in Online Appendix D) follows a standard normal distribution, and the trace plot (Panel (b) of Figure D6 in Online Appendix D) shows adequate mixing for all parameters. The posterior population mean and 95% CI for each component of β_0 are presented in Table C2 of Online Appendix C. Age at enrollment and years of education are significant predictors of AD. The posterior mean of the threshold for the null region $\hat{\delta} = -0.05$, which is consistent with previous research.⁵ Figure 3 shows individual trajectory plots for six patients from the ROS data (blue dotdash curves), estimated trajectories (red long-dash curves), and 95% credible bands (shaded area), which follow the trend of and overlap with the observed sample trajectories; \hat{p} represents the estimated probability of progression to AD at the 11th visit using the prediction algorithm described in the Personalized prediction of disease progression to AD section.

5.2 Potential clinical usage: Personalized risk prediction

Next, we used the model to predict the risk of disease progression to AD on another randomly selected sample. For brevity, we illustrate the approach by randomly selecting 14 new patients from different stages of disease progression. For these patients, seven were diagnosed as AD (the top seven patients in Table 2), three patients' disease stage switched between NCI and MCI (the middle three patients in Table 2), and four were NCI during the course of the study (the bottom four patients in Table 2). We present the prediction results for additional 18 patients (6 patients in each disease stage) in Table C3 of Online Appendix C, which has the same structure as Table 2. We repeatedly fit model (3) combining these new patients with the 189 patients from the previous analysis. We first fit the model using their global cog measurements from the first six visits (measurements at baseline and the first five follow-up visits), which allowed sufficient time for patients to experience potential change points. We then fit the model again using their global cog measurements from the first seven visits (global cog measurement at baseline and the first six follow-up visits), and repeatedly fitting model (3) with additional observations up to follow-up visit eight, nine,

and so on until first AD diagnosis or the eleventh visit for the non-AD patients, where the eleventh visit is the maximal follow-up time for seven AD patients.

In the following, we suppress the time index j and the individual index i for $\hat{\beta}_3$ and \hat{p} . Table 2 presents the posterior mean of the slope decrement, along with \hat{p} (as a %), the estimated personalized risk of progression to AD, from the sixth to the eleventh visit. By repeatedly fitting model (3) with additional global cog measurements at each follow-up visit, the slope decrement, $\hat{\beta}_3$, and the probability of progression to AD, \hat{p} , were updated.

Using both the posterior means of $\hat{\beta}_3$ and \hat{p} at each follow-up visit, we can obtain information about patients' likelihood of progression to AD. For example, for patient 50108200, at the sixth visit, $\hat{\beta}_3$ is just outside the null region (where $\hat{\delta} = -0.044$ at visit 6), and \hat{p} is 53%. At the seventh visit, $\hat{\beta}_3$ decreases, and \hat{p} increases slightly. At the eighth visit, \hat{p} has a relatively larger jump compared to the previous visits, but $\hat{\beta}_3$ is not far from the null region. These findings may indicate the onset of cognitive decline for this patient. However, at the ninth visit compared to the estimates at the previous visits, $\hat{\beta}_3$ jumps to -0.11 , which is a steeper decline and is farther away from the null region. Additionally, \hat{p} is 79%, which represents a relatively large increase compared to the sixth to the eighth visits. These results indicate that this patient's cognitive decline continues the trend we observed at the eighth visit. Hence, we can view this as an early warning sign of cognitive impairment. At the tenth visit, \hat{p} is 95%, which is a stronger warning of cognitive decline. Clinicians may consider more frequent visits or enhanced treatments to mitigate the cognitive decline for this patient. At the tenth and eleventh visits, $\hat{\beta}_3$ s bound well away from the null region, and \hat{p} s are nearly 100%. These estimates are in concordance with the clinical diagnosis of AD at the eleventh visit. However, our model is able to provide an early warning of serious cognitive decline at least three years before the actual diagnosis of AD for this patient. For patients 50102790, 50103679, 81874628, 64336939, and 20195344, the model provides a strong warning of cognitive decline ($\hat{p} > 80\%$ and $\hat{\beta}_3$ far away from the null region) at the seventh, sixth, sixth, ninth, and eighth visit, respectively, well in advance of their actual AD diagnoses.

In some cases, there is a discrepancy between the model prediction and clinical diagnosis. For example, patient 59796318's first AD diagnosis was at the eighth visit. However, the model's suggestion is different. The estimated posterior mean of β_3 and p at the eighth visit are -0.09 and 72%, which is a suggestive but not conclusive indication of progression to AD. In fact, the AD diagnosis at the eighth visit was a misdiagnosis: this patient was subsequently followed for two more visits and the disease stage was NCI at the ninth visit and MCI at the tenth visit. Thus, the proposed method could be used to identify potential misdiagnoses when the model predictions disagree with clinical assessments.

For the three patients whose disease stage switched between NCI and MCI, the posterior means of $\hat{\beta}_3$ s and \hat{p} s fluctuate along with their global cog trajectories, which indicates that our model is able to capture subtle cognitive changes and the switch of the disease stage from time to time. Consider, for example, patient 5632732. At the sixth and seventh

visits, $\hat{\beta}_3$ s for this patient are just outside the boundary of the null region, and \hat{p} s are less than 50%. However, at the eighth visit, \hat{p} jumps to 74% compared to that of 44% at the seventh visit, and $\hat{\beta}_3$ is farther away from the null region, which can be viewed as an early warning of cognitive decline. At the ninth visit, \hat{p} increases to 83%, which is a strong indication of cognitive decline. Clinicians may wish to monitor this patient more often, prescribe enhanced medications, or recommend behavioral therapies. For patient 50108912, the cognitive trajectory fluctuates over a larger range. In particular, \hat{p} at the tenth visit is more than 85%, and it reaches nearly 90% at the eleventh visit, which is a convincing evidence of serious cognitive decline. Although this patient was not diagnosed as AD during the course of the study, our results suggest the potential for AD onset during the study. As a result, clinicians may want to adjust the treatment and monitoring schedule for this patient to delay or prevent the onset of AD.

Among the NCI patients, one patient (ID 5218242) might experience temporary cognitive decline at the sixth visit with $\hat{\beta}_3 = -0.11$, which is outside the null region, and $\hat{p} = 77\%$. These larger estimates may be also due to the learning stage of the model when we fit the model using a shorter trajectory. However, this patient's cognitive function gradually returned to normal, as $\hat{\beta}_3$ is gradually absorbed back into the null region and \hat{p} decreases. Clinicians may therefore consider monitoring this patient annually according to routine protocol. For one patient (ID 10100600), the cognitive function was relatively stable throughout the follow-up period: the posterior means ($\hat{\beta}_3$ s) are close to the boundary $\hat{\delta}$, and \hat{p} s are less than 60% for visits 7 to 11. Clinicians and AD researchers may choose to monitor this patient according to the current annual schedule until they see more severe signs of cognitive decline.

The last patient (ID 31813134) was diagnosed as NCI at the last visit. However, the model results suggest this patient may experience AD earlier in the study. In particular, the slope decrement $\hat{\beta}_3$ decreases steadily from the eighth visit to the eleventh visit, where $\hat{\beta}_3 = -0.14$ is well outside the null region. The estimated probability \hat{p} steadily increases over time from 82% at the tenth visit to 94% at the eleventh visit, which is a convincing sign of progression to AD. For this scenario, clinicians may consider monitoring this patient more frequently to avoid potential misdiagnosis, which may delay treatment and jeopardize the patient's quality of life. More generally, when there is a discrepancy between clinical diagnosis and what model suggests, clinicians may need to pay special attention to the patient.

6 Discussion

In this paper, we proposed a Bayesian hierarchical random change point model with a parameter constraint in which we allow for a random intercept, random change point time, random slope before the change point, and random slope after the change point. The parameter space for the difference in slope before and after the change point was constrained and partitioned into a null region, representing NCI, and a rejection region, representing diseased cognitive impairment. Using this model, we can (1) distinguish normal aging individuals from individuals experiencing diseased cognitive impairment, (2) characterize

the rate and timing of the cognitive decline if individuals' cognitive functional trajectories encounter change points, and (3) identify patients at high risk of progression to AD using both the difference in slope estimates and the probability of advancing to AD. The simulations and the real application illustrate that the proposed model works very well in terms of model fitting and prediction. The model is robust to different parameter values, and estimating the threshold is preferable to assuming a fixed value, as the latter may be misspecified in practice.

Most importantly, the proposed model provides an additional tool to assist clinical evaluation for AD patients. Using the model, we are able to predict the subject-specific probability of progression to AD at various time points by measurements available only from cognitive measurements. Using this predicted probability together with the subject-specific change point and the rate of cognitive decline, clinicians can be assisted to (1) identify individuals with high risk of progression to AD several years before the actual event, (2) adjust individuals' treatment plan and follow-up visit plan accordingly, which may hopefully delay disease progression and improve patients' quality of life, and (3) use the model results as an additional tool to assist the diagnosis, particularly when there is a discrepancy between the clinical diagnosis and the model's suggestion about the changes in cognitive trajectories, thus potentially avoiding misdiagnoses and improving patients' quality of life.

Because AD is a nonreversible disease with no cure, it is critical to identify individuals at high risk of progression to AD. Doing so may prompt patients and clinicians to better understand their patient's risk factors and identify effective medication regimens or behavioral therapies that may prevent or delay disease progression. For example, patients can increase their physical activity, eat a heart- (and brain-) healthy diet, quit smoking, and initiate cognitive training to enhance memory, reasoning, and speed of processing.³⁰

While the widespread clinical use of our method may be several years away, the approach might find more immediate application in AD research centers or memory clinics where extensive longitudinal cognitive data are available for clinicians or researchers involved in large cohort AD research studies. Thus, for the time being, the target population might be limited to participants in those large-scale AD research cohort studies and any patients with memory concerns referred by their primary care doctors to clinicians. Despite these limitations, we believe the method will have growing impact as a way to assess patients longitudinally in a noninvasive manner in order to make predictions for the potential onset of dementia. This goal aligns with NIH AD diagnostic guidelines (<https://www.nia.nih.gov/health/alzheimers-disease-diagnostic-guidelines>), which states that "clinicians should obtain long-term assessment of cognitive whenever possible to gain evidence of progressive decline."

There are a number potential extensions of this work. First, future extensions of this work could focus on testing these models using cognitive screening instruments (e.g. MMSE, MiniCog) which, while less sensitive than a neuropsychological test battery, are more widely available. Second, the model could be expanded to include two change points and more complex order constraints. This model could be useful since the cognitive measurements

could encounter two change points. Because the process of AD is from NCI to MCI, and eventually to AD, the first change point could happen at the time when the disease switches from NCI to MCI, and the second change point could happen at the time when the disease transfers from MCI to AD. In this case, our model can be extended to include three slopes: the slope before the first change point (α_1), the slope between the first and the second change point (α_2), and the slope after the second change point (α_3). We could add order constraints on these slopes, for example, $\alpha_1 > \alpha_2 > \alpha_3$. Third, we could incorporate imaging or other biomarker data together with clinical evaluation measurements and fit a joint model of multiple outcomes to improve the predictions. Finally, we could model the random change point, κ_i , as a function of certain predictors by fitting a truncated normal regression, with the mean as a function of predictors instead of a fixed value.

A complementary approach would be to model the data via a hidden Markov model (HMM)³¹ which could include measurements available beyond clinical evaluation (e.g. neuroimaging data). Under this approach, patients' disease stage could transition between NCI and MCI and to AD over the course of the study. Conditional on a patient's disease state at time t , we could model cognitive impairment score. The HMM model could be used to determine when patients transition to the "absorbing" AD state, which might be used to facilitate clinical practice.

There are some limitations of the proposed approach. First, in the Potential clinical usage: Personalized risk prediction section, we illustrate how to use the proposed method to predict disease progression using estimated values of $\hat{\beta}_3$ and $\hat{\rho}$. While our assessments of $\hat{\beta}_3$ and $\hat{\rho}$ are at present qualitative in nature, working with clinical investigators, it may be possible to derive clinically meaningful thresholds for these values. Second, the proposed model needs longitudinal data to run, and these data may not be widely available outside of large-scale longitudinal cohort studies, although some of these have publicly available data that might support the use of our model. Additionally, in memory disorder clinics where patients are annually seen with cognitive assessments, reviewing electronic health records or certain algorithms (for example, Natural language processing algorithm) may be used to extract relevant clinical data (e.g. cognitive test scores and other clinical data that are collected over time). In general, we anticipate that the model will find increasing value in settings where longitudinal data related to cognitive decline are available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Rush Alzheimer's Disease Center at Rush University Medical Center provided Religious Orders Study dataset. The study was supported by NIA grant P30AG10161. Data can be requested at <https://www.radc.rush.edu>.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by grant R21 LM012866 from the National Institutes of Health.

References

1. Toepper M. Dissociating normal aging from Alzheimer's disease: a view from cognitive neuroscience. *J Alzheimers Dis* 2017; 57: 331–352. [PubMed: 28269778]
2. Vandenberghe R and Tournoy J. Cognitive aging and Alzheimer's disease. *Postgrad Med J* 2005; 81: 343–352. [PubMed: 15937198]
3. Spaan P. Cognitive decline in normal aging and early Alzheimer's disease: a continuous or discontinuous transition? A historical review and future research proposal. *Cogent Psychol* 2016; 3: 1–12.
4. Perl DP. Neuropathology of Alzheimer's disease. *Mt Sinai J Med* 2010; 77: 32–42. [PubMed: 20101720]
5. Yu L, Boyle P, Wilson R, et al. A random change point model for cognitive decline in Alzheimer's disease and mild cognitive impairment. *Neuroepidemiology* 2012; 39: 73–83. [PubMed: 22814083]
6. Hall CB, Lipton RB, Sliwinski M, et al. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Stat Med* 2000; 19: 1555–1566. [PubMed: 10844718]
7. Komarova N and Thalhauser C. High degree of heterogeneity in Alzheimer's disease progression patterns. *PLoS Comput Biol* 2011; 7: 1–6.
8. Alzheimer's Disease Association. Alzheimer's disease facts and figures. Chicago, IL: Author, 2016.
9. Hall CB, Ying J, Juo L, et al. Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Comput Stat Data Anal* 2003; 42: 91–119.
10. Cloutier S, Chertkow H, Kergoat M, et al. Patterns of cognitive decline prior to dementia in persons with mild cognitive impairment. *J Alzheimers Dis* 2015; 47: 901–913. [PubMed: 26401770]
11. Jacqmin-Gadda H, Commenges D and Dartigues J. Random change point model for joint modeling of cognitive decline and dementia. *Biometrics* 2006; 62: 254–260. [PubMed: 16542253]
12. van den Hout A, Fox J-P and Muniz-Terrera G. Longitudinal mixed-effects models for latent cognitive function. *Stat Model* 2015; 15: 366–387.
13. Dominicus A, Ripatti S, Pedersen N, et al. A random change point model for assessing variability in repeated measures of cognitive function. *Stat Med* 2008; 27: 5786–5798. [PubMed: 18680123]
14. Ji M, Xiong C and Grundman M. Hypothesis testing of a change point during cognitive decline among Alzheimer's disease patients. *J Alzheimers Dis* 2003; 5: 375–382. [PubMed: 14646028]
15. Slate EH and Turnbull BW. Statistical models for longitudinal biomarkers of disease onset. *Stat Med* 2000; 19: 617–637. [PubMed: 10694740]
16. Bennett D, Schneider J, Arvanitakis Z, et al. Overview and finding from the religious orders study. *Curr Alzheimer Res* 2012; 9: 628–645. [PubMed: 22471860]
17. Beach T, Monsell S, Phillips L, et al. Accuracy of the clinical diagnosis of Alzheimer disease at national institute on aging Alzheimer's disease centers, 2005–2010. *J Neuropathol Exp Neurol* 2012; 71: 266–273. [PubMed: 22437338]
18. Gaugler JE, Kane RL, Johnston JA, et al. Sensitivity and specificity of diagnostic accuracy in Alzheimer's disease: a synthesis of existing evidence. *Am J Alzheimers Dis Other Dement* 2013; 28: 337–347.
19. Sabbagh M, Lue L, Fayard D, et al. Increasing precision of clinical diagnosis of Alzheimer's disease using a combined algorithm incorporating clinical and novel biomarker data. *Neurol Ther* 2017; 6(Suppl 1): S83–S95.
20. Chintamaneni M and Bhaskar M. Biomarkers in Alzheimer's disease: a review. *ISRN Pharmacol* 2012; 12: 1–12.
21. Moreno L, Osta R and Calvo A. New perspectives in the search for reliable biomarkers in Alzheimer's disease. *Eur J Psychiat* 2015; 29: 51–65.
22. Apostolova LG. Alzheimer's disease. *Continuum* 2016; 22: 419–434. [PubMed: 27042902]
23. Wilson R, Beckett A, Barnest L, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002; 17: 179–193. [PubMed: 12061405]
24. Marchand E and Strawderman WE. Estimation in restricted parameter spaces: a review. *Inst Math Stat Lect Notes Monogr Ser* 2004; 45: 21–44.

25. Neelon B and Dunson D. Bayesian isotonic regression and trend analysis. *Biometrics* 2004; 60: 398–406. [PubMed: 15180665]
26. Gelfand AE, Smith AFM and Lee T. Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *J Am Stat Assoc* 1992; 87: 523–532.
27. Richardson S and Green P. On Bayesian analysis of mixtures with an unknown number of components. *J R Statist Soc B* 1997; 59: 731–792.
28. Rajan K, Wilson R, Barnes B, et al. A cognitive turning point in development of clinical Alzheimer's disease dementia and mild cognitive impairment: a biracial population study. *J Gerontol Med Sci* 2017; 72: 424–430.
29. Chiu G, Lockhart R and Routledge R. Bent-cable regression theory and applications. *J Am Stat Assoc* 2006; 101: 542–553.
30. Mendiola-Precoma J, Berumen LC, Padilla K, et al. Therapies for prevention and treatment of Alzheimer's disease. *BioMed Res Int*; 2016; 2016: 2589276.
31. Yu L, Boyle PA, Leurgans S, et al. Effect of common neuropathologies on progression of late life cognitive impairment. *Neurobiol Aging* 2015; 36: 2224–2231.

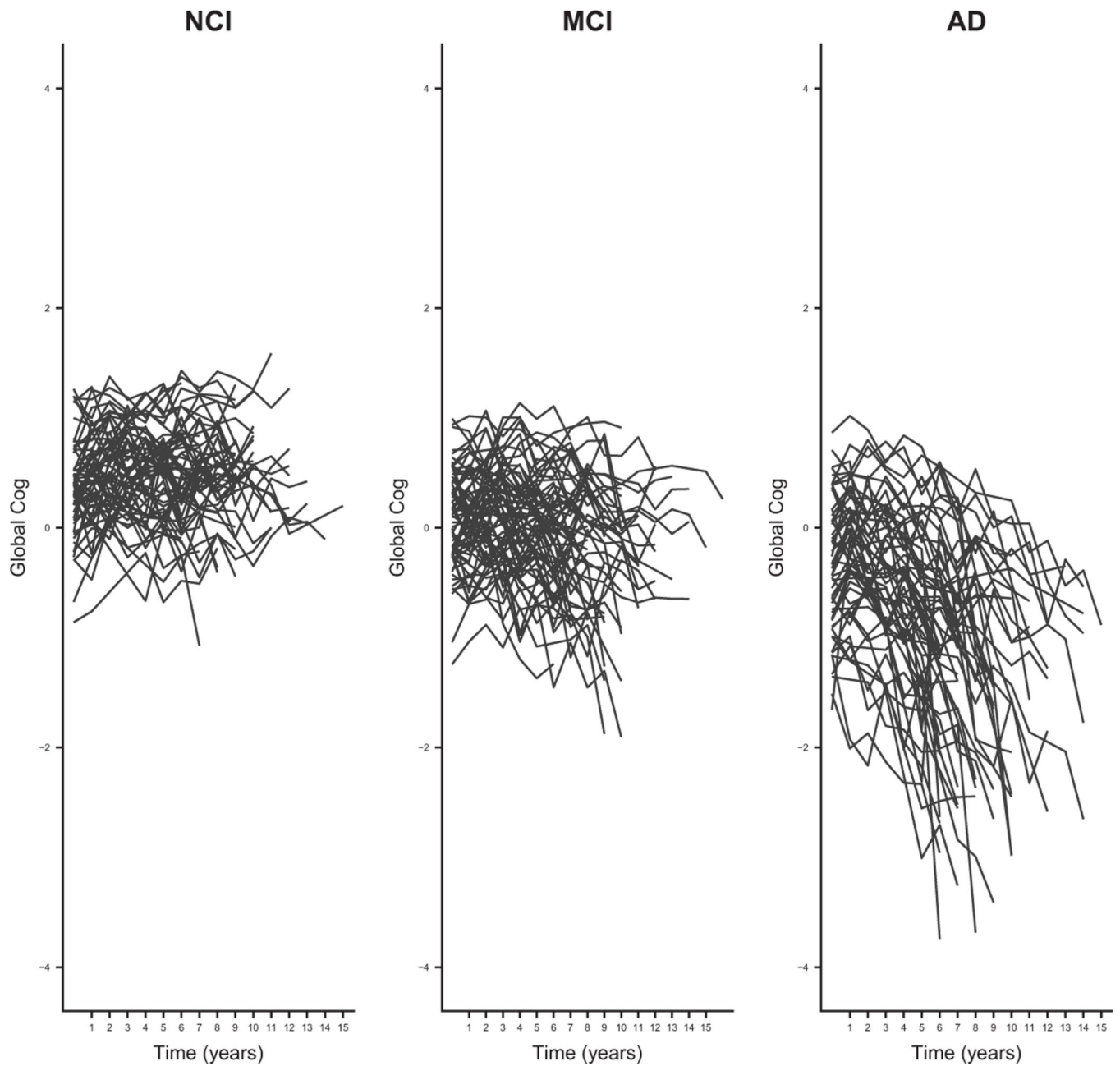


Figure 1. Cognitive function trajectories for patients diagnosed with NCI, MCI, and AD at the final clinical visit for the sample of 189 patients in the ROS data.
NCI: no cognitive impairment; MCI: mild cognitive impairment; AD: Alzheimer's disease.

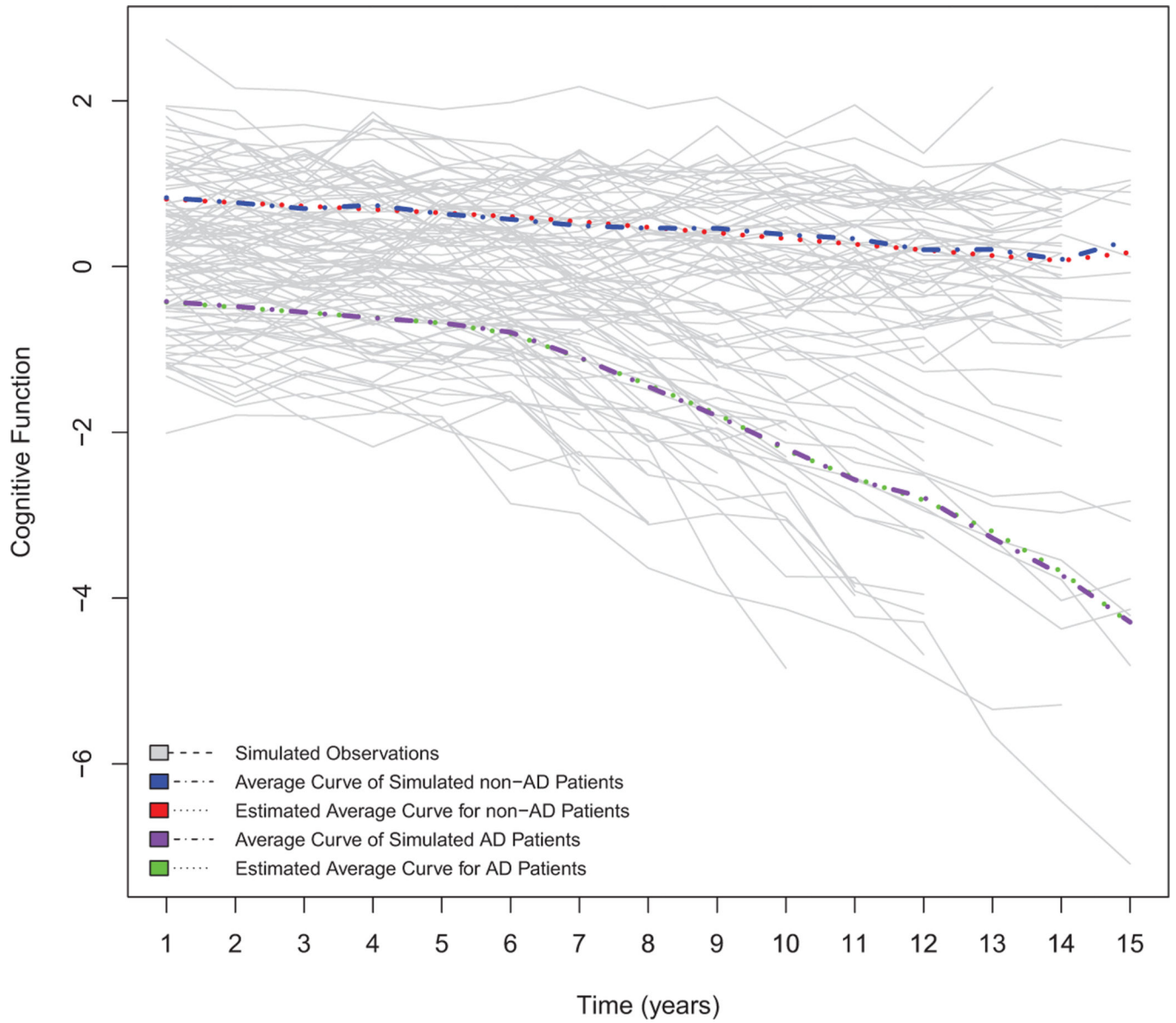


Figure 2.
 Estimated and simulated average cognitive trajectories for Scenario 1 in the simulation study.
 AD: Alzheimer’s disease

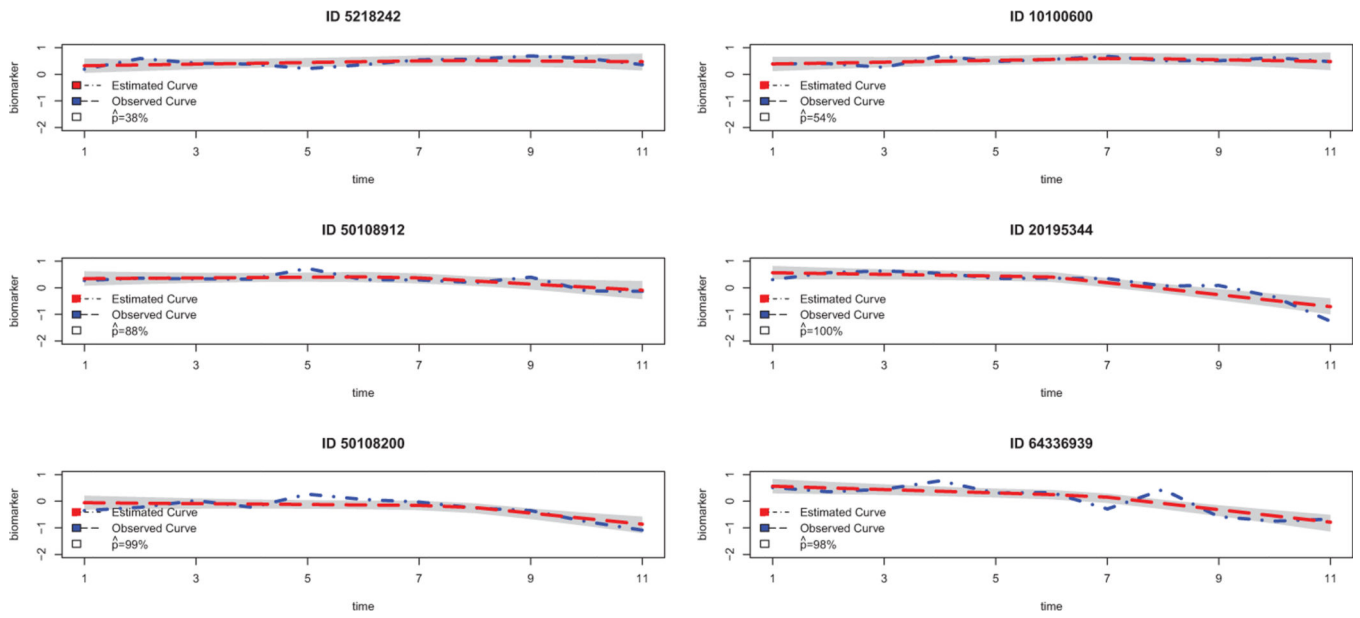


Figure 3. Observed and estimated trajectories for six selected patients from the ROS study. Shaded regions denote 95% credible bands.

Table 1.

Simulated and posterior mean parameter estimates for simulation study for Scenarios 1, 2, and 3.

Parameter	True value	Posterior mean (95% CI)
Scenario 1: $\delta = -0.05$, $\sigma_{\epsilon}^2 = 0.05$, $\rho = 0.1$, and Σ_{β} with Specification 1		
β_1	-0.18	-0.20 [-0.45, 0.06]
β_2	-0.05	-0.05 [-0.06, -0.04]
β_3	-0.30	-0.26 [-0.32, -0.21]
β_4	-0.04	-0.05 [-0.06, -0.035]
β_5	0.04	0.04 [0.02, 0.06]
σ_{ϵ}^2	0.05	0.05 [0.049, 0.06]
δ	-0.05	-0.05 [-0.07, -0.03]
Scenario 2: $\delta = -0.1$, $\sigma_{\epsilon}^2 = 0.5$, $\rho = 0.1$, and Σ_{β} with Specification 1		
β_1	-0.18	-0.16 [-0.40, 0.13]
β_2	-0.05	-0.06 [-0.08, -0.03]
β_3	-0.30	-0.28 [-0.32, -0.23]
β_4	-0.04	-0.05 [-0.06, -0.027]
β_5	0.04	0.04 [0.02, 0.06]
σ_{ϵ}^2	0.5	0.48 [0.45, 0.53]
Δ	-0.1	-0.10 [-0.11, -0.09]
Scenario 3: $\delta = -0.05$, $\sigma_{\epsilon}^2 = 0.05$, $\rho = 0.2$, and Σ_{β} with Specification 1		
β_1	-0.18	-0.18 [-0.52, 0.09]
β_2	-0.05	-0.05 [-0.06, -0.04]
β_3	-0.30	-0.26 [-0.31, -0.21]
β_4	-0.04	-0.05 [-0.06, -0.03]
β_5	0.04	0.04 [0.02, 0.06]
σ_{ϵ}^2	0.05	0.05 [0.048, 0.06]
δ	-0.05	-0.05 [-0.06, -0.04]

CI: credible interval. 95% CIs are given in parentheses.

Table 2.

Clinical usage: individual risk of progression to AD ($\hat{\beta}_3$ and $\hat{p}(\%)$) for 14 randomly selected patients from the ROS study.

ID	6th visit	7th visit	8th visit	9th visit	10th visit	11th visit	Time to first AD diagnosis
50108200	-0.06	-0.07	-0.09	-0.11	-0.15	-0.18	11th visit
	53	54	71	79	95	99	
50102790	-0.09	-0.12	-0.16	-0.18	-0.18	-	10th visit
	70	84	95	98	98		
50103679	-0.13	-0.15	-0.16	-0.18	-	-	9th visit
	85	91	96	98			
81874628	-0.13	-0.13	-0.14	-0.13	-0.17	-	10th visit
	86	89	94	90	98		
64336939	-0.11	-0.15	-0.11	-0.15	-0.17	-0.17	11th visit
	78	92	78	93	98	98	
20195344	-0.10	-0.09	-0.12	-0.12	-0.15	-0.19	11th visit
	76	69	84	85	95	100	
59796318	-0.08	-0.05	-0.09	-	-	-	8th visit
	61	46	72				
5632732	-0.05	-0.05	-0.10	-0.12	-0.06	-0.07	-
	47	44	74	83	60	61	
5577538	-0.10	-0.07	-0.06	-0.03	-0.02	-0.02	-
	73	56	50	34	24	22	
50108912	-0.09	-0.09	-0.10	-0.08	-0.12	-0.12	-
	70	64	79	62	86	88	
5218242	-0.11	-0.07	-0.05	-0.03	-0.03	-0.04	-
	77	54	46	27	30	40	
10100448	-0.08	-0.07	-0.04	-0.02	-0.04	-0.05	-
	59	57	37	27	37	52	
10100600	-0.08	-0.06	-0.07	-0.07	-0.05	-0.05	-
	65	47	55	58	49	54	
31813134	-0.07	-0.06	-0.08	-0.10	-0.11	-0.14	-
	60	52	64	77	82	94	

AD: Alzheimer's disease.