Published in final edited form as:

Biometrics. 2015 September ; 71(3): 625–635. doi:10.1111/biom.12313.

Quantile Regression with a Change-point Model for Longitudinal Data: An Application to the Study of Cognitive Changes in Preclinical Alzheimer's Disease

Chenxi Li^{1,*}, N. Maritza Dowling², and Rick Chappell^{2,3}

¹Department of Epidemiology and Biostatistics, Michigan State University, East Lansing, MI 48824, U.S.A

²Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, WI 53792, U.S.A

³Department of Statistics, University of Wisconsin-Madison, Madison, WI 53706, U.S.A

Summary

Progressive and insidious cognitive decline that interferes with daily life is the defining characteristic of Alzheimer's disease (AD). Epidemiological studies have found that the pathological process of AD begins years before a clinical diagnosis is made and can be highly variable within a given population. Characterizing cognitive decline in the preclinical phase of AD is critical for the development of early intervention strategies when disease-modifying therapies may be most effective. In the last decade, there has been an increased interest in the application of change-point models to longitudinal cognitive outcomes prior to and after diagnosis. Most of the proposed statistical methodology for describing decline relies upon distributional assumptions that may not hold. In this paper, we introduce a quantile regression with a change-point model for longitudinal data of cognitive function in persons bound to develop AD. A change-point in our model reflects the transition from the cognitive decline due to normal aging to the accelerated decline due to disease progression. Quantile regression avoids common distributional assumptions on cognitive outcomes and allows the covariate effects and the change-point to vary for different quantiles of the response. We provided an approach for estimating the model parameters, including the change-point, and presented inferential procedures based on the asymptotic properties of the estimators. A simulation study showed that the estimation and inferential procedures perform reasonably well in finite samples. The practical use of our model was illustrated by an application to longitudinal episodic memory outcomes from two cohort studies of aging and AD.

*cli@epi.msu.edu.

Supplementary Materials

Web Appendix A, referenced in Section 2.2, R code implementing the methods and a simulated data set are available with this paper at the *Biometrics* website on Wiley Online Library.

Keywords

Alzheimer's disease; Change-point model; Cognitive aging; Disease progression; Longitudinal data; Quantile regression

1. Introduction

Alzheimer's disease (AD) is an age-associated chronic neurodegenerative disorder, clinically characterized by a progressive loss of memory and other cognitive and functional abilities. Impairments in episodic memory (the ability to learn and retain new information) are among the earliest signs and symptoms of AD. Characterizing the trajectories of cognitive decline in persons bound to develop AD and the risk factors affecting rate of decline is critical for developing strategies for early diagnosis and potential treatments for delaying or preventing the onset of dementia. In the preclinical stage of AD, the cognitive decline is gradual and difficult to differentiate from normal aging or mild affective disorders (Sliwinski et al., 1996; Becker et al., 2009), but as the disease progresses, the decline in cognitive function accelerates and becomes distinguishable from the profile of typical cognitive aging. Prior literature suggests that the acceleration of cognitive decline might begin months or even years before a clinical diagnosis of AD is made (Amieva et al., 2005; Johnson et al., 2009; Grober et al., 2008; Howieson et al., 2008). Numerous random effects single change-point models have been formulated to examine this acceleration of cognitive loss and individual-level variations in the change trajectories in the pre-diagnosis phase of AD (see, e.g., Hall et al., 2000, 2003, Jacqmin-Gadda et al., 2006, Yu et al., 2012, van den Hout et al., 2010, and more recently, Yang and Gao, 2013). However, the conditional distribution of the response variable in a random effects change-point model is often assumed to be normal even though the response is discrete and/or shows boundary effects. To avoid this shortcoming, we propose a quantile regression change-point model to describe changes in cognitive function in the preclinical and prodromal stages of AD. In addition to being distribution free, as in other situations, rank-based methods are more useful than those which estimate means in change-point models when there may be outliers and, especially, when the scientific questions focus on extremes.

Quantile regression (QR) models the effects of covariates on a conditional quantile of the response variable. Unlike mean regression, QR can delineate the entire conditional distribution of a response given covariates, doesn't have to specify the error distribution for estimation and inference, and is robust against response outliers (for an overview, see Koenker, 2005). In the past two decades, QR has been extended to longitudinal data modeling in several ways. For example, approaches based on marginal models include those proposed by Jung (1996), He et al. (2002), Wang and Fygenon (2009), and Wang and Zhu (2011). Others have proposed methodological approaches based on subject-specific models (Koenker, 2004; Kim and Yang, 2011) or using Bayesian methods for model estimation (Reich et al., 2010). Recently, Li et al. (2011) studied a piecewise linear QR model with a single change-point. In this paper, we extend this model to the longitudinal framework and present an application to study the progression of cognitive impairment in incident AD subjects participating in two ongoing clinical-pathological cohort studies. Similar to Li et al.

(2011), this work focuses on dealing with single change-point situations. Extensions to handle two or more change-points in the model are prohibited by algorithmic issues and merit future research.

The rest of the paper is organized as follows. Section 2 presents the methodology and the associated asymptotic theory. Section 3 evaluates the finite sample performance of the methodology through simulation studies. In Section 4, we apply the proposed methodology to a real data set to examine the longitudinal course of cognitive function in preclinical AD and how decline varies with age, education, and genetic measures. We summarize the paper and discuss future research directions in Section 5. The proofs of the statistical properties of the model parameter estimators are given in the online supplementary material.

2. Quantile Regression with a Change-point Model for Longitudinal Data

2.1 Model Setup and Parameter Estimation

We consider the following general quantile regression with a change-point model (QRCPM) for longitudinal data:

$$Y_{ij} = \alpha_\tau + (\beta_{1,\tau} I\{X_{ij} \leq t_\tau\} + \beta_{2,\tau} I\{X_{ij} > t_\tau\})(X_{ij} - t_\tau) + \mathbf{z}_{ij}^\top \boldsymbol{\gamma}_\tau + u_{ij}^{(\tau)}, \quad i=1, \dots, N, \quad j=1, \dots, n_i, \quad (1)$$

where i indexes subjects, j indexes observations within a subject, Y_{ij} is the response variable, X_{ij} is the scalar covariate whose slope changes at an unknown change-point t_τ , $\mathbf{z}_{ij}$ is a q -dimensional vector of covariates with constant slopes, the τ th conditional quantile of $u_{ij}^{(\tau)}$

given X_{ij} and $\mathbf{z}_{ij}$ is 0, and $u_{ij}^{(\tau)}$ are independent across i (subjects) but are dependent within a subject. Model (1) with the above assumptions is a population-averaged model for longitudinal data. It means that in the population with covariate values X_{ij} and $\mathbf{z}_{ij}$, 100 τ % of its subjects have an outcome value no larger than

$\alpha_\tau + (\beta_{1,\tau} I\{X_{ij} \leq t_\tau\} + \beta_{2,\tau} I\{X_{ij} > t_\tau\})(X_{ij} - t_\tau) + \mathbf{z}_{ij}^\top \boldsymbol{\gamma}_\tau$. All model parameters ($\alpha_\tau, \beta_{1,\tau}, \beta_{2,\tau}, t_\tau, \boldsymbol{\gamma}_\tau$) depend on τ . In data analysis, τ is usually determined based on the subject-matter interest. Throughout the rest of the paper we assume that n_i (the number of observations within a subject) is bounded by some positive constant K , but N (the number of subjects) grows. Let $\boldsymbol{\theta}_0(\tau) = (\alpha_\tau, \beta_{1,\tau}, \beta_{2,\tau}, t_\tau, \boldsymbol{\gamma}_\tau^\top)^\top$ be the true parameter vector. We define

$$n = \sum_{i=1}^N n_i, \quad \boldsymbol{\theta} = (\alpha, \beta_1, \beta_2, t, \boldsymbol{\gamma}^\top)^\top, \quad \boldsymbol{\eta} = (\alpha, \beta_1, \beta_2, \boldsymbol{\gamma}^\top)^\top, \quad \boldsymbol{\eta}_0(\tau) = (\alpha_\tau, \beta_{1,\tau}, \beta_{2,\tau}, \boldsymbol{\gamma}_\tau^\top)^\top, \quad \mathbf{w}_{ij} = (1, X_{ij}, \mathbf{z}_{ij}^\top)^\top, \text{ and}$$

$$g(\mathbf{w}_{ij}; \boldsymbol{\theta}) = \alpha + (\beta_1 I\{X_{ij} \leq t\} + \beta_2 I\{X_{ij} > t\})(X_{ij} - t) + \mathbf{z}_{ij}^\top \boldsymbol{\gamma}.$$

To derive the estimator of $\boldsymbol{\theta}_0(\tau)$, $\hat{\boldsymbol{\theta}}_n(\tau)$, we adopt the working assumption of independence between observations and thus consider minimizing the objective function

$$Q_{n,\tau}(\boldsymbol{\theta}) = \sum_{ij} \rho_{\tau}(Y_{ij} - g(\mathbf{w}_{ij}; \boldsymbol{\theta})), \quad (2)$$

where $\rho_{\tau}(u) = u(\tau - I\{u < 0\})$ is the quantile loss function. Due to the non-convexity of $Q_{n,\tau}(\boldsymbol{\theta})$, $\hat{\boldsymbol{\theta}}_n(\tau)$ is obtained via profile estimation. A profile estimator of $\boldsymbol{\eta}$ at a fixed t is given by

$$\hat{\boldsymbol{\eta}}_{n,\tau}(t) = \arg \min_{\boldsymbol{\eta} \in \mathbb{R}^{3+q}} Q_{n,\tau}(\boldsymbol{\eta}, t),$$

which can be obtained by the function ‘rq’ in R package ‘quantreg’. An estimator of the change-point t_{τ} is given by

$$\hat{t}_{n,\tau} = \arg \min_{t \in (a,b) \cap (X_{n(2)}, X_{n(n-1)})} Q_{n,\tau}(\hat{\boldsymbol{\eta}}_{n,\tau}(t), t),$$

where a and b are two constants such that t_{τ} is thought to be in (a, b) , usually determined graphically. $\hat{\boldsymbol{\theta}}_n(\tau)$ is obtained from $\hat{\boldsymbol{\eta}}_{n,\tau}(\hat{t}_{n,\tau})$ and $\hat{t}_{n,\tau}$. The optimization method used to minimize $Q_{n,\tau}(\hat{\boldsymbol{\eta}}_{n,\tau}(t), t)$ is a combination of golden section search and successive parabolic interpolation, implemented by the function ‘optimize’ in R package ‘stats’. On a PC with a 3.40GHz CPU and 8GB RAM, the profile estimation method takes 0.73 seconds to fit a change-point quantile regression model with two covariates to a longitudinal data set of sample size 8000 (1600 subjects each having 5 observations).

2.2 Large Sample Properties

To derive the asymptotic properties of $\hat{\boldsymbol{\theta}}_n(\tau)$, we introduce some notations:

$$\begin{aligned} h(\mathbf{w}_{ij}; \boldsymbol{\theta}) &= (I\{X_{ij} \leq t\}, X_{ij}I\{X_{ij} \leq t\}, I\{X_{ij} > t\}, X_{ij}I\{X_{ij} > t\}, \mathbf{z}_{ij}^{\top})^{\top}, \\ g_1(\mathbf{w}_{ij}; \boldsymbol{\theta}) &= \alpha + \beta_1(X_{ij} - t) + \mathbf{z}_{ij}^{\top} \boldsymbol{\gamma}, \\ g_2(\mathbf{w}_{ij}; \boldsymbol{\theta}) &= \alpha + \beta_2(X_{ij} - t) + \mathbf{z}_{ij}^{\top} \boldsymbol{\gamma}, \\ \tilde{g}(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau)) &= \frac{\partial g_1(\mathbf{w}_{ij}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0(\tau)} I\{X_{ij} \leq t_{\tau}\} + \frac{\partial g_2(\mathbf{w}_{ij}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0(\tau)} I\{X_{ij} > t_{\tau}\}, \\ f_{ij}^{(\tau)} &\text{ is the conditional density of } u_{ij}^{(\tau)} \text{ given } \mathbf{w}_{ij}, \\ C_{n,\tau} &= n^{-1} \sum_{ij} h(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau)) h(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau))^{\top}, \\ D_{n,\tau} &= -n^{-1} \sum_{ij} f_{ij}^{(\tau)}(0) h(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau)) \tilde{g}(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau))^{\top}, \quad \text{and} \\ S_{n,\tau}(\boldsymbol{\eta}, t) &= n^{-1} \sum_{ij} [\rho_{\tau}(Y_{ij} - g(\mathbf{w}_{ij}; \boldsymbol{\eta}, t)) - \rho_{\tau}(Y_{ij})]. \end{aligned}$$

Note that if $f_{Y_{ij}}$ is the conditional density of Y_{ij} given $\mathbf{w}_{ij}$, we have

$f_{ij}^{(\tau)}(u) = f_{Y_{ij}}(u + g(\mathbf{w}_{ij}; \boldsymbol{\theta}_0(\tau)))$. The large sample properties of $\hat{\boldsymbol{\theta}}_n(\tau)$ are presented in the following two theorems. The full proofs are deferred to Web Appendix A.

Theorem 1—Under assumptions (A1)–(A5) in Web Appendix A and the condition that $\beta_{1,\tau} \quad \beta_{2,\tau} \quad \hat{\boldsymbol{\theta}}_n(\tau)$ is a strongly consistent estimator of $\boldsymbol{\theta}_0(\tau)$.

Theorem 2—Under assumptions (A1)–(A6), given in Web Appendix A, and the condition that $\beta_{1,\tau} \quad \beta_{2,\tau}$ we have

$$\Gamma_n(\delta^{(\tau)})^{-1/2} n^{1/2}(\hat{\theta}_n(\tau) - \theta_0(\tau)) \xrightarrow{d} N(0, I), \quad (3)$$

where

$$\begin{aligned} \Gamma_n(\delta^{(\tau)}) = & n^{-1} D_{n,\tau}^{-1} \{ \sum_{ij} h(\mathbf{w}_{ij}; \theta_0(\tau)) h(\mathbf{w}_{ij}; \theta_0(\tau))^{\top} \tau(1-\tau) \\ & + \sum_i \sum_{j \neq j'} h(\mathbf{w}_{ij}; \theta_0(\tau)) h(\mathbf{w}_{ij'}; \theta_0(\tau))^{\top} (\delta_{ijj'}^{(\tau)} - \tau^2) \} D_{n,\tau}^{-\top}, \\ \delta_{ijj'}^{(\tau)} = & P(u_{ij}^{(\tau)} < 0, u_{ij'}^{(\tau)} < 0) \text{ and } \delta^{(\tau)} = \{ \delta_{ijj'}^{(\tau)} | i=1, \dots, N, 1 \leq j, j' \leq n_i, j \neq j' \}. \end{aligned}$$

It is easy to extend the foregoing result to QRCPM at multiple quantile levels. For $0 < \tau_1 <$

$\tau_2 < \dots < \tau_m < 1$, $m \in \mathbb{Z}_+$, we set $\theta_0(\tau_j) = (\alpha_{\tau_j}, \beta_{1,\tau_j}, \beta_{2,\tau_j}, t_{\tau_j}, \gamma_{\tau_j}^{\top})^{\top}$, $1 \leq j \leq m$. We assume that model (1) is true for all the τ_j 's. Let $\hat{\theta}_n = (\hat{\theta}_n(\tau_1)^{\top}, \hat{\theta}_n(\tau_2)^{\top}, \dots, \hat{\theta}_n(\tau_m)^{\top})^{\top}$ be the vector of estimated parameters with $\hat{\theta}_0 = (\theta_0(\tau_1)^{\top}, \theta_0(\tau_2)^{\top}, \dots, \theta_0(\tau_m)^{\top})^{\top}$. To describe the asymptotic properties of $\hat{\theta}_n$, we introduce three more notations: $H_n = \text{diag}[D_{n,\tau_1}, \dots, D_{n,\tau_m}]$, $\zeta_{ijj'}^{(\tau_k, \tau_l)} = P(u_{ij}^{(\tau_k)} < 0, u_{ij'}^{(\tau_l)} < 0)$ for $j \neq j'$ and $k \neq l$, and

$$J_n(\tau_k, \tau_l) = \begin{cases} n^{-1} \{ \sum_{ij} h(\mathbf{w}_{ij}; \theta_0(\tau_k)) h(\mathbf{w}_{ij}; \theta_0(\tau_k))^{\top} \tau_k(1-\tau_k) \\ + \sum_i \sum_{j \neq j'} h(\mathbf{w}_{ij}; \theta_0(\tau_k)) h(\mathbf{w}_{ij'}; \theta_0(\tau_k))^{\top} (\delta_{ijj'}^{(\tau_k)} - \tau_k^2) \} & (k=l), \\ n^{-1} \{ \sum_{ij} h(\mathbf{w}_{ij}; \theta_0(\tau_k)) h(\mathbf{w}_{ij}; \theta_0(\tau_l))^{\top} (\tau_k \wedge \tau_l - \tau_k \tau_l) \\ + \sum_i \sum_{j \neq j'} h(\mathbf{w}_{ij}; \theta_0(\tau_k)) h(\mathbf{w}_{ij'}; \theta_0(\tau_l))^{\top} (\zeta_{ijj'}^{(\tau_k, \tau_l)} - \tau_k \tau_l) \} & (k \neq l), \end{cases}$$

By Lemma 1 in Web Appendix A, we have with probability one that

$$n^{1/2}(\hat{\theta}_n(\tau_k) - \theta_0(\tau_k)) = -n^{-1/2} D_{n,\tau_k}^{-1} \sum_{ij} \varphi_{\tau_k}(u_{ij}^{(\tau_k)}) h(\mathbf{w}_{ij}; \theta_0(\tau_k)) + o_p(1) \quad k=1, \dots, m, \quad (4)$$

where $\varphi_{\tau}(r) = \tau - \mathbf{I}\{r < 0\}$. Stacking (4) from $k=1$ to m and then applying the multivariate central limit theorem, we obtain

$$(H_n^{-1} \Lambda_n H_n^{-\top})^{-1/2} n^{1/2}(\hat{\theta}_n - \theta_0) \xrightarrow{d} N(0, I), \quad (5)$$

where Λ_n is a $m(4+q) \times m(4+q)$ matrix with kl th block being $J_n(\tau_k, \tau_l)$ ($k, l = 1, \dots, m$).

2.3 Asymptotic Covariance Matrix Estimation

If the asymptotic covariance matrix $H_n^{-1} \Lambda_n H_n^{-\top}$ in (5) can be consistently estimated, confidence intervals and hypothesis tests on θ_0 can be derived when the sample size n is

large. The asymptotic covariance matrix estimation involves estimating $f_{ij}^{(\tau_k)}(0)$, $\delta_{ijj'}^{(\tau_k)}$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)}(k, l=1, \dots, m)$. The density estimation method we use is the so-called *Hendricks-Koenker Sandwich* described in Section 3.4.2 of Koenker (2005). $\delta_{ijj'}^{(\tau_k)}$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)}$ are estimated in different ways depending on the assumed correlation structure of $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^\top$.

We consider the estimation of $\delta_{ijj'}^{(\tau_k)}$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)}$ for four specific correlation structures of $\mathbf{Y}_i$:

(1) Compound Symmetry Form—The correlation matrix of $\mathbf{Y}_i$ is of the compound symmetry form, i.e.,

$$\text{corr}(\mathbf{Y}_i) = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \cdots & 1 \end{pmatrix},$$

where ρ is constant across subjects. Under this form of correlation matrix, it is reasonable to assume that $\delta_{ijj'}^{(\tau_k)} = \delta^{(\tau_k)} := P(u_{11}^{(\tau_k)} < 0, u_{12}^{(\tau_k)} < 0)$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)} = \zeta^{(\tau_k, \tau_l)} := P(u_{11}^{(\tau_k)} < 0, u_{12}^{(\tau_l)} < 0)$ for $i = 1, \dots, N, j, j' = 1, \dots, n_i$ and $j \neq j'$. This assumption holds if, e.g., $u_{ij}^{(0.5)} = a_i^{(0.5)} + \varepsilon_{ij}^{(0.5)}$, where $a_i^{(0.5)}$'s are i.i.d. random subject effects that are independent of the i.i.d. within-subject random errors $\varepsilon_{ij}^{(0.5)}$'s. So two naturally reasonable estimators of $\delta^{(\tau_k)}$ and $\zeta^{(\tau_k, \tau_l)}$ are respectively

$$\hat{\delta}^{(\tau_k)} = \left\{ \sum_i n_i(n_i - 1)/2 \right\}^{-1} \sum_i \sum_{j < j'} I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_k)} < 0\} \quad (6)$$

and

$$\hat{\zeta}^{(\tau_k, \tau_l)} = \left\{ \sum_i n_i(n_i - 1) \right\}^{-1} \sum_i \sum_{j \neq j'} I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_l)} < 0\}, \quad (7)$$

where $\hat{u}_{ij}^{(\tau)}$'s are the fitted residuals from the QRCPM at quantile level τ .

(2) AR(1) Structure—The correlation matrix of $\mathbf{Y}_i$ has AR(1) structure, i.e.,

$$\text{corr}(\mathbf{Y}_i) = \begin{pmatrix} 1 & \rho & \cdots & \rho^{n_i-1} \\ \rho & 1 & \cdots & \rho^{n_i-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{n_i-1} & \rho^{n_i-2} & \cdots & 1 \end{pmatrix},$$

where ρ is constant across subjects. Under this correlation structure, it is reasonable to assume that $\delta_{ijj'}^{(\tau_k)} = \delta_{\text{ar1}, |j-j'|}^{(\tau_k)} := P(u_{11}^{(\tau_k)} < 0, u_{1,1+|j-j'|}^{(\tau_k)} < 0)$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)} = \zeta_{\text{ar1}, |j-j'|}^{(\tau_k, \tau_l)} := P(u_{11}^{(\tau_k)} < 0, u_{1,1+|j-j'|}^{(\tau_l)} < 0)$ for $i = 1, \dots, N, j, j' = 1, \dots, n_i$ and $j \neq j'$. This assumption holds if, e.g., $u_{ij}^{(0.5)} = v(X_{ij}, \mathbf{z}_{ij})u_{ij}$, where $v(\cdot, \cdot)$ is a bivariate linear function and u_{ij} ($j = 1, \dots, n_i$) follow a stationary AR(1) time series model. Therefore, two naturally reasonable estimators of $\delta_{\text{ar1}, j}^{(\tau_k)}$ and $\zeta_{\text{ar1}, j}^{(\tau_k, \tau_l)}$ ($j=1, \dots, \max_i n_i - 1$) are respectively

$$\hat{\delta}_{\text{ar1}, j}^{(\tau_k)} = \frac{1}{\sum_i (n_i - j) I\{n_i > j\}} \sum_i \sum_{q=1}^{n_i-j} I\{\hat{u}_{iq}^{(\tau_k)} < 0, \hat{u}_{i, q+j}^{(\tau_k)} < 0\} \quad (8)$$

and

$$\hat{\zeta}_{\text{ar1}, j}^{(\tau_k, \tau_l)} = \frac{1}{\sum_i 2(n_i - j) I\{n_i > j\}} \sum_i \sum_{q=1}^{n_i-j} \left[I\{\hat{u}_{iq}^{(\tau_k)} < 0, \hat{u}_{i, q+j}^{(\tau_l)} < 0\} + I\{\hat{u}_{i, q+j}^{(\tau_k)} < 0, \hat{u}_{iq}^{(\tau_l)} < 0\} \right]. \quad (9)$$

(3) Heteroscedastic Correlation— $u_{ij}^{(0.5)} = a_i^{(0.5)} + g(X_{ij})\varepsilon_{ij}^{(0.5)}$, where $g(\cdot)$ is a non-constant function, $a_i^{(0.5)}$'s are independent and identically distributed, $\varepsilon_{ij}^{(0.5)}$'s are independent and identically distributed, and $a_i^{(0.5)}$ and $\varepsilon_{ij}^{(0.5)}$ are independent. Under this assumption, it is reasonable to estimate $\delta_{ijj'}^{(\tau_k)}$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)}$ with generalized additive (GA) models (Hastie and Tibshirani, 1986). Specifically, we fit a GA logistic model to the binary response $I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_k)} < 0\}$ ($j \neq j'$) for estimating $\delta_{ijj'}^{(\tau_k)}$:

$$\log \frac{\delta_{ijj'}^{(\tau_k)}}{1 - \delta_{ijj'}^{(\tau_k)}} = \alpha + S_1(|X_{ij} - X_{ij'}|) + S_2(X_{ij} + X_{ij'}), \quad (10)$$

where $S_1(\cdot)$ and $S_2(\cdot)$ are two unspecified functions that are related with $g(\cdot)$. This form of GA model embodies the feature that $\delta_{ijj'}^{(\tau_k)} = \delta_{ij'j}^{(\tau_k)}$. $S_1(\cdot)$ and $S_2(\cdot)$ are estimated by smoothing splines with four degrees of freedom. Therefore, the estimator of $\delta_{ijj'}^{(\tau_k)}$ is

$$\hat{\delta}_{ijj'}^{(\tau_k)} = \frac{\exp \left\{ \hat{\alpha} + \hat{S}_1(|X_{ij} - X_{ij'}|) + \hat{S}_2(X_{ij} + X_{ij'}) \right\}}{1 + \exp \left\{ \hat{\alpha} + \hat{S}_1(|X_{ij} - X_{ij'}|) + \hat{S}_2(X_{ij} + X_{ij'}) \right\}}, \quad (11)$$

where $\hat{\alpha}$, $\hat{S}_1(\cdot)$ and $\hat{S}_2(\cdot)$ are the estimators of α , $S_1(\cdot)$ and $S_2(\cdot)$, respectively. We also fit a GA logistic model to the binary response $I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_l)} < 0\}$ ($j \neq j'$) for estimating $\zeta_{ijj'}^{(\tau_k, \tau_l)}$:

$$\log \frac{\zeta_{ijj'}^{(\tau_k, \tau_l)}}{1 - \zeta_{ijj'}^{(\tau_k, \tau_l)}} = \eta + Q_1(X_{ij}) + Q_2(X_{ij'}), \quad (12)$$

where $Q_1(\cdot)$ and $Q_2(\cdot)$ are another two unspecified functions that are related with $g(\cdot)$. Note that we don't necessarily have $\zeta_{ijj'}^{(\tau_k, \tau_l)} = \zeta_{ij'j}^{(\tau_k, \tau_l)}$. Again, $Q_1(\cdot)$ and $Q_2(\cdot)$ are estimated by smoothing splines with four degrees of freedom. Hence, the estimator of $\zeta_{ijj'}^{(\tau_k, \tau_l)}$ is

$$\hat{\zeta}_{ijj'}^{(\tau_k, \tau_l)} = \frac{\exp \left\{ \hat{\eta} + \hat{Q}_1(X_{ij}) + \hat{Q}_2(X_{ij'}) \right\}}{1 + \exp \left\{ \hat{\eta} + \hat{Q}_1(X_{ij}) + \hat{Q}_2(X_{ij'}) \right\}}, \quad (13)$$

where $\hat{\eta}$, $\hat{Q}_1(\cdot)$ and $\hat{Q}_2(\cdot)$ are the estimators of η , $Q_1(\cdot)$ and $Q_2(\cdot)$, respectively.

(4) Unstructured Correlation—For an unstructured correlation matrix of $\mathbf{Y}_i$, we propose two general estimators for $\delta_{ijj'}^{(\tau_k)} - \tau_k^2$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)} - \tau_k \tau_l$ respectively:

$$\hat{\delta}_{ijj'}^{(\tau_k)} - \tau_k^2 = I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_k)} < 0\} - \frac{1}{2} \tau_k \left[I\{\hat{u}_{ij}^{(\tau_k)} < 0\} + I\{\hat{u}_{ij'}^{(\tau_k)} < 0\} \right] \quad (14)$$

and

$$\hat{\zeta}_{ijj'}^{(\tau_k, \tau_l)} - \tau_k \tau_l = I\{\hat{u}_{ij}^{(\tau_k)} < 0, \hat{u}_{ij'}^{(\tau_l)} < 0\} - \frac{1}{2} \left[I\{\hat{u}_{ij}^{(\tau_k)} < 0\} \tau_l + \tau_k I\{\hat{u}_{ij'}^{(\tau_l)} < 0\} \right]. \quad (15)$$

These two general estimators can also be applied to the previous three special cases.

It is worth pointing out that the above covariance matrix estimation methods are also applicable to the following linear quantile regression model for longitudinal data,

$$Y_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\beta}_\tau + u_{ij}^{(\tau)}, \quad i=1, \dots, N, \quad j=1, \dots, n_i, \quad (16)$$

where $\mathbf{X}_{ij} \in \mathbb{R}^p$, i indexes subjects, j indexes observations within a subject, the τ th conditional quantile of $u_{ij}^{(\tau)}$ given $\mathbf{X}_{ij}$ is 0, and $u_{ij}^{(\tau)}$ s are independent across i (subjects) but are dependent within a subject. The simplest estimation approach for (16) is to compute the quantile regression as if the data were all independent, i.e., estimating $\boldsymbol{\beta}_\tau$ by

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^N \sum_{j=1}^{n_i} \rho_{\tau}(Y_{ij} - \mathbf{X}_{ij}^{\top} \beta). \quad (17)$$

He et al. (2002) and Wang and Fygenon (2009), among others, showed that this simple method still consistently estimates β_{τ} and that the estimator is asymptotically normal with a covariance matrix containing the same terms $f_{ij}^{(\tau)}(0)$, $\delta_{ijj'}^{(\tau)}$ and $\zeta_{ijj'}^{(\tau_k, \tau_l)}$ as defined for (5).

3. Simulation study of QRCPM for longitudinal data

We investigated the finite sample performance of the estimators as well as Wald confidence intervals based on (5) for slopes and change-point in QRCPM through a simulation study. The simulated data were generated from the following three scenarios:

1. Scenario of Compound Symmetry Correlation Structure:

$$y_{ij} = 2.5 + (0.5I\{x_{ij} \leq 5\} - 0.5I\{x_{ij} > 5\})(x_{ij} - 5) - 0.2z_{ij} + a_i + \varepsilon_{ij}, \quad (18)$$

where $a_i \stackrel{\text{iid}}{\sim} N(0, 1)$ and $\varepsilon_{ij} \stackrel{\text{iid}}{\sim} N(0, 1)$;

2. Scenario of AR(1) Correlation Structure:

$$y_{ij} = 2.5 + (0.5I\{x_{ij} \leq 5\} - 0.5I\{x_{ij} > 5\})(x_{ij} - 5) - 0.2z_{ij} + v(x_{ij})u_{ij}, \quad (19)$$

where $v(x) = 3.2 - 0.2x$, $u_{ij} = 0.5u_{ij-1} + \varepsilon_{ij}$ and $\varepsilon_{ij} \stackrel{\text{iid}}{\sim} N(0, 1)$;

3. Scenario of Heteroscedastic Correlation Structure:

$$y_{ij} = 2.5 + (0.5I\{x_{ij} \leq 5\} - 0.5I\{x_{ij} > 5\})(x_{ij} - 5) - 0.2z_{ij} + a_i + g(x_{ij})\varepsilon_{ij} \quad (20)$$

where $g(x) = \sqrt{(3.2 - 0.2x)^2 - 1}$, $a_i \stackrel{\text{iid}}{\sim} N(0, 1)$ and $\varepsilon_{ij} \stackrel{\text{iid}}{\sim} N(0, 1)$.

The three corresponding covariance matrix estimators described in Section 2.3 were used in the above three scenarios respectively. As a comparison, the general estimator of the covariance matrix for the unstructured correlation of responses was also used in each scenario. To add some variation to the number of observations per subject in all three scenarios, we set $n_i = 5$ for $i = 1, \dots, N-2$, $n_{N-1} = 4$ and $n_N = 6$, where N was the number of subjects in a data set. The number of observations n in a data set was thus $5N$. In the scenarios of compound symmetry and heteroscedastic correlation structures, $x_{ij} \stackrel{\text{iid}}{\sim} U(0, 10)$ and $z_{ij} \stackrel{\text{iid}}{\sim} U(0, 10)$. In the scenario of AR(1) correlation structure, $x_{i1} \stackrel{\text{iid}}{\sim} U(0.5, 7.5)$, $x_{ij} = x_{i,j-1} + 0.5$ for $j > 1$, and $z_{ij} \stackrel{\text{iid}}{\sim} U(0, 10)$. The number of simulation runs was 1000. The simulation results are presented in Table 1. In all cases, we reported results for Wald confidence intervals with a nominal coverage probability of 0.95.

As shown in Table 1, the average of the standard error estimates is smaller than the empirical standard error for the change-point estimator in every case. As a result, the coverage rate of every Wald confidence interval for the change-point is lower than the nominal level of 95%. This might be due to the density estimation involved in estimating the asymptotic covariance matrix, which is not as efficient as the estimation of $\theta_0(\tau)$. Another explanation could be related to the finding by Hinkley (1969), based on an empirical study, that the asymptotic normality of the change-point estimator is not a good approximation for small sample sizes in two-phase broken line regression, although it was proved that, in this special case of segmented regression, the change-point estimator does have an asymptotic normal distribution (Feder, 1975). Notwithstanding the underestimation of the change-point estimator's standard error for small sample sizes, our simulation showed that, as the sample size increases, the estimate of theoretical standard error approaches the empirical standard error and the empirical coverage probability of Wald confidence interval goes up towards 95%, especially in the compound symmetry scenario. The estimators and Wald confidence intervals for regression coefficients performed reasonably well except a little underestimation of the standard errors in AR(1) and heteroscedastic scenarios when the number of observations is smaller than or equal to 4000. Again, this underestimation might be due to the less efficient density estimation involved in estimating the asymptotic covariance matrix. Similar findings (Wald confidence intervals' coverage rates being smaller than the nominal level) were reported by Koenker (2005), page 111, for linear quantile regression with a sample size of 500. In every correlation structure scenario, the

standard error estimate based on the general estimator of $\delta_{ijj}^{(\tau)}$ is equivalent to the one based on its specific estimator suitable for that scenario, so is the empirical coverage probability of

Wald confidence interval. This suggests that the general estimator of $\delta_{ijj}^{(\tau)}$ might be used universally.

4. Application to the Longitudinal Assessment of Cognitive Function in Preclinical Alzheimer's Disease

4.1 Background and Data

Section 4 presents an application of the QRCPM to the combined data from two ongoing cohort studies of incident Alzheimer's disease: the Religious Orders Study (ROS) (Bennett et al., 2012) and the Memory and Aging Project (MAP) (Bennett et al., 2012). Recruitment, exclusion, and inclusion criteria for these studies and subject evaluations have been previously described in detail (Wilson et al., 2002). Briefly, both studies recruit older individuals without dementia who agree to receive clinical and psychological evaluation each year and to donate their brain for postmortem examination. Enrollment for ROS began in 1994 and includes the participation of over 1,100 older religious clergy (priests, brothers, and nuns). MAP study began in 1997 and currently includes more than 1,400 participants from about 40 retirement communities and senior housing facilities in the Chicago metropolitan area. Besides sharing similar clinical and pathologic findings, these studies also share a common 17- test neuropsychological test battery and follow the same standard protocol and criteria for clinical diagnosis. Both studies were approved by the Institutional Review Board of Rush University Medical Center. Written informed consent was obtained

from all study participants. We used data from MAP and ROS studies to estimate the time point at which the acceleration of cognitive decline began in different cognition quantile groups of persons who developed AD and to study whether the shape of cognitive decline depends on gender, education level, and genetic risk factors.

Outcome Measure—The dependent variable of interest in this illustration, episodic memory, was a composite score comprised of six widely-used clinical neuropsychological tests intended for measuring conscious retrieval or recognition of information acquired at a particular time and place (Wilson et al., 2002). Although other psychological functions also may be affected in the early stages of the disease process, episodic memory has been frequently associated with cognitive deficits and early brain changes in preclinical AD (Salmon and Bondi, 1999). Indeed, progressive impairment in episodic memory lasting more than six months is part of the diagnostic criteria for probable AD (Dubois et al., 2007). Therefore, we focused on composite score outcomes defining the domain of episodic memory collected annually on both studies for up to 12 years. The target population for analysis was defined as subjects who were diagnosed with AD during the follow-up visits.

Covariates—Explanatory covariates were included in the model to study predictors of inter-individual differences in change-points and trajectories of cognition quantile over time at different quantile levels. Participant's gender, years of education (categorized into two levels: 12 years and > 12 years), and the presence of the apolipoprotein E ϵ 4 allele (ApoE4), a genetic risk factor for cognitive decline and AD, were included in the analysis as time-independent covariates. All these variables have been previously shown to be associated with changes in cognitive function.

Clinical Diagnoses—Annual diagnostic classifications followed a three step process. First, a neuropsychological test battery was administered and scored. Second, a clinical neuropsychologist assessed the presence of cognitive impairment after a careful review of all cognitive scores results and relevant demographic information. Third, a clinician determined whether standard criteria for dementia and AD (McKhann et al., 1984) were met. At this step, individuals were classified with respect to AD, mild cognitive impairment, and other neurologic disorders.

The sample for analysis consisted of 429 incident AD cases with at least one follow-up measure of episodic memory and at least one non-missing explanatory covariate. Participant's age at baseline ranged from 64.78 yrs to 102.10 yrs. A summary of characteristics of the sample is presented in Table 2.

4.2 Model Building

As in modeling the change of mean cognitive score over time in AD patients (Yu et al., 2012; Hall et al., 2000), we characterized the evolution of episodic memory quantile by a piecewise linear trajectory with a change-point in the time at which the decline of episodic memory quantile accelerated. The inspection of spaghetti plots of episodic memory score against age and time to AD diagnosis (Figures 1(a) and 1(b)) suggested that the change-

point might be different for different individuals on the scale of age, but common on the scale of time to AD diagnosis.

Thus, we fitted the following model to the data of the 429 incident AD cases:

$$Y_{ij} = \alpha_\tau + \beta_\tau I\{X_{ij} > t_\tau\} (X_{ij} - t_\tau) + \gamma_\tau (Z_{ij} - 65) + u_{ij}^{(\tau)}, \quad i=1, \dots, N, \quad j=1, \dots, n_i, \quad (21)$$

where Y_{ij} , X_{ij} and Z_{ij} are episodic memory score, the time to AD diagnosis and the age of subject i at the j th measurement, respectively, t_τ is the time point when the decline of episodic memory quantile begins to accelerate on the scale of time to diagnosis, and $u_{ij}^{(\tau)}$ has τ th quantile at zero. γ_τ represents the normal aging effect on episodic memory. β_τ reflects the change in decline rate of episodic memory quantile after the change-point. The subtraction of 65 in the age terms was done for interpretability of the intercept. Model (21) is equivalent to the following piecewise linear model on the age scale with a random change-point depending on the age of AD occurrence.

$$Y_{ij} = (\alpha_\tau - 65\gamma_\tau) + \gamma_\tau I\{Z_{ij} \leq R_i + t_\tau\} \{Z_{ij} - (R_i + t_\tau)\} + (\beta_\tau + \gamma_\tau) I\{Z_{ij} > R_i + t_\tau\} \{Z_{ij} - (R_i + t_\tau)\} + \gamma_\tau (R_i + t_\tau) + u_{ij}^{(\tau)}, \\ i=1, \dots, N, \quad j=1, \dots, n_i,$$

where R_i is subject i 's age of AD occurrence. Model (21) is also the same as the final statistical model used in Hall et al. (2000) except that the former models the quantile and the latter models the mean. We performed statistical inference assuming an unstructured correlation matrix for $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})$, that is, we used the covariance matrix estimator for the unspecified correlation structure described in Section 2.3. All the hypothesis tests in the analysis were two-sided with a significance level of 0.05.

To study whether the shape of episodic memory decline was related to gender, education level or genetic risk factors and whether there was a cohort effect from the two different study populations, we fitted model (21) with $\tau = 0.5$ to the data of the subsets of males, females, subjects who attended college, subjects who didn't attend college, subjects with ApoE4, subjects without ApoE4, subjects in MAP study and subjects in ROS study separately. Here we assume that the quantile functions (21) at different levels are affected by the same set of significant covariates in a common way across quantile levels. Because the model at $\tau = 0.5$ is the most robust against outliers and measurement errors in the models of all quantile levels, it was used to determine the significant covariates and how they affect the quantile curves. Through the 4-degrees-of-freedom Wald test and the 1-degree-of-freedom Wald test, we tested the difference in the full set of parameters and the individual parameters of model (21) between males and females, subjects who attended college and subjects who didn't attend college, subjects with ApoE4 and subjects without ApoE4, and subjects in MAP study and subjects in ROS study. We found that ApoE4 status produced significant differences in only $\alpha_{0.5}$ and $\gamma_{0.5}$, but gender and study cohort didn't yield any significant difference in any model parameter. In particular, the change-point $t_{(0.5)}$ doesn't differ

significantly for different cohorts and gender/education/genetic risk groups. Therefore, we decided to fit the following model to the data

$$Y_{ij} = \alpha_{\tau} + \beta_{\tau} I\{X_{ij} > t_{\tau}\} (X_{ij} - t_{\tau}) + \gamma_{\tau} (Z_{ij} - 65) + a_{\tau} V_i + b_{\tau} V_i * (Z_{ij} - 65) + u_{ij}^{(\tau)}, \quad (22)$$

where V_i is the indicator of whether to carry ApoE4.

4.3 Analysis Results

Model (22) was estimated at three quantile levels: 0.2, 0.5, and 0.8. The parameter estimates and their asymptotic confidence intervals are tabulated in Table 3. The 95% confidence intervals for the parameters showed that after the change-point, episodic memory declined significantly faster than before the onset of change at all the three quantile levels. The confidence intervals also showed that the presence of at least one ApoE4 allele significantly lowered the 20th and 50th percentiles of the episodic memory score and significantly decreased the decline rate of the 20th percentile. This finding may be due to the fact that subjects with episodic memory score at 20th percentile in the ApoE4 carrier group started follow-up with lower scores than their counterparts in the ApoE4 non-carrier group.

The three change-points for the 20th, 50th and 80th percentiles of the episodic memory score were 4.4, 4.6 and 5.0 years prior to the diagnosis of AD respectively. These results were similar to the estimates of change-points for the Buschke Selective Reminding Test score (Masur et al., 1994) obtained by Hall et al. (2000) (4.3 and 5.1 years prior to diagnosis in AD patients and dementia patients respectively). For comparison purposes, we also applied Hall et al. (2000) change-point model to our example data using the same outcome, explanatory variables, and general model set-up. The change-point estimate was 3.4 years prior to diagnosis; closest to the lowest (20th) percentile change-point obtained with the QRCPM.

Results of Wald tests for the QRCPM didn't show that change-points were significantly different across the three quantile levels. Furthermore, we compared the decline rate of episodic memory quantile across the three quantile levels using Wald tests. The decline rates before the change-point were not significantly different across the quantile levels for either AD patients with ApoE4 or AD patients without ApoE4. The decline rates after the change-point were significantly different across the quantile levels for both AD patients with ApoE4 and AD patients without ApoE4 (both $p \approx 0$). As illustrated in Figure 2, after the acceleration of episodic memory decline starts, the lower percentiles of the episodic memory score decline faster than the higher percentiles for both ApoE4 carriers and non-carriers.

5. Discussion

In this paper, we proposed a quantile regression with a change-point modeling framework for longitudinal data. The proposed methodology is useful in describing the trajectories of disease progression and studying the effect of potential correlates of observed clinical outcomes. We applied the model to data from two longitudinal clinical-pathologic cohort studies of aging and AD to study cognitive decline in preclinical AD. The model presented

avoids the possibly unrealistic normality assumptions in the distribution of cognitive scores, which is required by random effects change-point models. It also allows risk factors or disease biomarkers to have different effects on different quantiles of the distribution of cognition. We provided the asymptotic properties of the model parameter estimators as well as the covariance matrix estimation methods for statistical inference. The proposed methodology has numerous applications in epidemiology and medicine to study the course of changes in disease progression. Other examples include the growth pattern in body mass index of children (Sorva et al., 1990) and T4 cell counts prior to AIDS (Kiuchi et al., 1995).

There are at least three future research topics on quantile regression change-point models. One is to model the effects of covariates on the change-point and its associated covariate's coefficients explicitly. This will lead to a unified model for the entire data rather than separate models for data sets separated by covariates that significantly affect the change-point and/or its associated covariate's coefficients. However, we expect that many research efforts are needed to develop the asymptotic results for the unified model. Instead, one could make inferences by bootstrapping individuals. Another direction is to study subject-specific quantile regression change-point models in which the within-subject dependence among data is taken into account through the incorporation of random effects. Accounting for random effects in quantile regression is difficult due to the lack of a likelihood. However, using the connection between the asymmetric Laplace distribution (ALD) and quantile regression (Yu and Moyeed, 2001), one could develop a likelihood-based inferential approach by assuming an ALD for residual errors and a symmetric multivariate Laplace distribution (Kotz et al., 2001) for random effects. This working assumption has been adopted by Liu and Bottai (2009) for mixed effects quantile regression models without a change-point. The analysis of the data from ROS and MAP studies implicitly assumed that all the missing outcome measures were missing at random. This assumption is possibly not true. Therefore, the third research topic would be how to treat missing data in quantile regression with a change-point model for longitudinal data, especially when the missing data are nonignorable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The data for the example presented in this work were provided by these two studies supported by the NIA: the Religious Orders Study (P30AG10161, R01AG15819), and the Rush Memory and Aging Project (NIA Grant R01AG17917). We are grateful to the principal investigator of these two projects, David A. Bennett, MD, for the authorization to use the data. This work is supported in part by grants from the Alzheimer's Association (NIRG-12-242799, Dowling) and the NIA (5 P50 AG033514-05, Dowling and Chappell).

References

Amieva H, Jacqmin-Gadda H, Orgogozo JM, Le Carret N, Helmer C, Letenneur C, Barberger-Gateau P, Fabrigoule C, Dartigues JF. The 9 year cognitive decline before dementia of the Alzheimer's type: A prospective population-based study. *Brain*. 2005; 128:1093–1101. [PubMed: 15774508]

- Becker J, Chang Y, Lopez O, Dew M, Sweet R, Barnes D, Yaffe K, Young J, Kuller L, Reynolds CF. Depressed mood is not a risk factor for incident dementia in a community-based cohort. *American Journal of Geriatric Psychiatry*. 2009; 17:1653–663.
- Bennett D, Schneider J, Arvanitakis Z, Wilson R. Overview and findings from the Religious Orders study. *Current Alzheimer Research*. 2012; 9:628–645. [PubMed: 22471860]
- Bennett D, Schneider J, Buchman A, Barnes L, Boyle AP, Wilson R. Overview and findings from the Rush Memory and Aging project. *Current Alzheimer Research*. 2012; 9:646–663. [PubMed: 22471867]
- Dubois B, Feldman H, Jacova C, DeKosky S, Barberger-Gateau P, Cummings J, Delacourte A, Galasko D, Gauthier G, Jicha G, Meguro J, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Stern Y, Visser P, Scheltens P. Research criteria for the diagnosis of Alzheimer's disease: Revising the nincdscdrda criteria. *The Lancet Neurology*. 2007; 6:734–746. [PubMed: 17616482]
- Feder PI. On asymptotic distribution theory in segmented regression problems—identified case. *Annals of Statistics*. 1975; 3:49–83.
- Grober E, Hall CB, Lipton RB, Zonderman AB, Resnick SM, Kawas C, Albert M, Moss M, Tanzi R, Jones K, et al. Memory impairment, executive dysfunction, and intellectual decline in preclinical Alzheimer's disease. *Journal of the International Neuropsychological Society*. 2008; 14:266–278. [PubMed: 18282324]
- Hall C, Ying J, Kuo L, Lipton R. Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Computational Statistics and Data Analysis*. 2003; 42:91–109.
- Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Statistics in Medicine*. 2000; 19:1555–1566. [PubMed: 10844718]
- Hastie T, Tibshirani R. Generalized additive models. *Statistical Science*. 1986; 1:297–318.
- He X, Zhu ZY, Fung WK. Estimation in a semiparametric model for longitudinal data with unspecified dependence structure. *Biometrika*. 2002; 89:579–590.
- Hinkley DV. Inference about the intersection in two-phase regression. *Biometrika*. 1969; 56:495–504.
- Howieson DB, Carlson NE, Moore MM, Wasserman D, Abendroth CD, Payne-Murphy J, Kaye JA. Trajectory of mild cognitive impairment onset. *Journal of the International Neuropsychological Society*. 2008; 14:192–198. [PubMed: 18282317]
- Jacqmin-Gadda H, Commenges D, Dartigues JF. Random changepoint model for joint modeling of cognitive decline and dementia. *Biometrics*. 2006; 62:254–260. [PubMed: 16542253]
- Johnson D, Storandt M, Morris J, Galvin J. Longitudinal study of the transition from healthy aging to Alzheimer's disease. *Archives of Neurology*. 2009; 66:1254–1259. [PubMed: 19822781]
- Jung S. Quasi-likelihood for median regression models. *Journal of the American Statistical Association*. 1996; 91:251–257.
- Kim MO, Yang Y. Semiparametric approach to a random effects quantile regression model. *Journal of the American Statistical Association*. 2011; 106:1405–1417. [PubMed: 22347760]
- Kiuchi AS, Hartigan JA, Holford TR, Rubinstein P, Stevens CE. Change points in the series of t4 counts prior to aids. *Biometrics*. 1995; 51:236–248. [PubMed: 7766779]
- Koenker R. Quantile regression for longitudinal data. *Journal of Multivariate Analysis*. 2004; 91:74–89.
- Koenker, R. *Quantile Regression*. Cambridge; New York: 2005.
- Kotz, S.; Kozubowski, TJ.; Podgorski, K. *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser; Boston: 2001.
- Li C, Wei Y, Chappell R, He X. Bent line quantile regression with application to an allometric study of land mammals' speed and mass. *Biometrics*. 2011; 67:242–249. [PubMed: 20528859]
- Liu Y, Bottai M. Mixed-effects models for conditional quantiles with longitudinal data. *The International Journal of Biostatistics*. 2009; 5:1–24.
- Masur D, Sliwinski M, Lipton R, Blau A, Crystal H. Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology*. 1994; 44:1427. [PubMed: 8058143]

- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan E. Clinical diagnosis of Alzheimer's disease report of the Nincds-Adrda work group under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology*. 1984; 34:939–939. [PubMed: 6610841]
- Reich BJ, Bondell HD, Wang HJ. Flexible bayesian quantile regression for independent and clustered data. *Biostatistics*. 2010; 11:337–352. [PubMed: 19948746]
- Salmon, D.; Bondi, M. Neuropsychology of alzheimer disease. In: Terry, R.; Katzman, R.; Bick, K.; Sisodia, S., editors. *Alzheimer disease*. 2. Lippincott Williams & Wilkins; Philadelphia: 1999.
- Sliwinski M, Lipton R, Buschke H, Stewart W. The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 1996; 51:217–225.
- Sorva R, Lankinen S, Tolpanen E, Perheentupa J. Variation of growth in height and weight of children after infancy. *Acta Paediatrica*. 1990; 79:498–506.
- van den Hout A, Muniz-Terrera G, Matthews F. Smooth random change point models. *Statistics in Medicine*. 2010; 30:599–610. [PubMed: 21337356]
- Wang H, Fygenon M. Inference for censored quantile regression models in longitudinal studies. *Annals of Statistics*. 2009; 37:756–781.
- Wang HJ, Zhu Z. Empirical likelihood for quantile regression models with longitudinal data. *Journal of Statistical Planning and Inference*. 2011; 141:1603–1615.
- Wilson R, Beckett L, Barnes L, Schneider JA, Bach J, Evans D, Bennett D. Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging*. 2002; 17:179. [PubMed: 12061405]
- Wilson R, De Leon CF, Barnes L, Schneider JA, Bienias JL, Evans DA, Bennett D. Participation in cognitively stimulating activities and risk of incident Alzheimer's disease. *Journal of the American Medical Association*. 2002; 287:742–748. [PubMed: 11851541]
- Yang L, Gao S. Bivariate random change point models for longitudinal outcomes. *Statistics in Medicine*. 2013; 32:1038–1053. [PubMed: 22893450]
- Yu K, Moyeed RA. Bayesian quantile regression. *Statistics and Probability Letters*. 2001; 54:437–447.
- Yu L, Boyle P, Wilson RS, Segawa E, Leurgans S, De Jager PL, Bennett DA. A random change point model for cognitive decline in Alzheimer's disease and mild cognitive impairment. *Neuroepidemiology*. 2012; 39:73–83. [PubMed: 22814083]

(a)

(b)

Figure 1. Spaghetti plots of episodic memory score against age (a) and time to AD diagnosis (b)

Figure 2.

Episodic memory score against time to diagnosis of AD conditional on ApoE4 status with fitted 80th (dotted), 50th (dashed) and 20th (solid) quantile regression lines. The fitted quantile regression lines are for the sample mean of age at diagnosis in the sample used to fit model (22): 86.1 years old.

Table 1

Simulation results for estimating $\beta_{1,\tau}$, $\beta_{2,\tau}$, t_τ and γ_τ

<i>n</i>	Corr. Struct.	Parameter	τ	Bias	SE	SEE		CP	
						Specific	General	Specific	General
2000	Comp. Sym.	$\beta_{1,\tau}$	0.5	0.002	0.041	0.040	0.040	0.941	0.943
			0.7	0.003	0.043	0.042	0.042	0.948	0.948
		$\beta_{2,\tau}$	0.5	0.001	0.038	0.038	0.038	0.939	0.938
			0.7	-0.001	0.040	0.040	0.040	0.951	0.951
		t_τ	0.5	-0.008	0.169	0.159	0.159	0.937	0.936
			0.7	-0.003	0.177	0.167	0.168	0.938	0.939
		γ_τ	0.5	0.000	0.013	0.014	0.014	0.960	0.961
			0.7	0.000	0.014	0.015	0.015	0.953	0.954
	AR(1)	$\beta_{1,\tau}$	0.5	0.018	0.153	0.130	0.129	0.914	0.913
			0.7	0.015	0.147	0.134	0.134	0.923	0.919
		$\beta_{2,\tau}$	0.5	-0.010	0.098	0.089	0.089	0.932	0.928
			0.7	-0.009	0.100	0.093	0.093	0.927	0.927
		t_τ	0.5	-0.002	0.429	0.317	0.317	0.865	0.860
			0.7	0.005	0.423	0.331	0.331	0.875	0.874
		γ_τ	0.5	0.000	0.025	0.024	0.024	0.932	0.932
			0.7	0.000	0.026	0.025	0.025	0.944	0.936
	Hetero.	$\beta_{1,\tau}$	0.5	0.003	0.080	0.075	0.075	0.929	0.930
			0.7	0.005	0.085	0.079	0.079	0.929	0.930
		$\beta_{2,\tau}$	0.5	-0.006	0.048	0.046	0.046	0.936	0.934
			0.7	-0.005	0.053	0.048	0.048	0.925	0.922
		t_τ	0.5	0.014	0.282	0.247	0.247	0.906	0.908
			0.7	0.009	0.305	0.261	0.260	0.894	0.895
		γ_τ	0.5	0.001	0.019	0.020	0.020	0.951	0.956
			0.7	0.000	0.020	0.021	0.021	0.957	0.953
4000	Comp. Sym.	$\beta_{1,\tau}$	0.5	0.001	0.028	0.028	0.028	0.956	0.953
			0.7	0.000	0.031	0.029	0.029	0.935	0.939

<i>n</i>	Corr. Struct.	Parameter	τ	Bias	SE	SEE		CP	
						Specific	General	Specific	General
8000	AR(1)	$\beta_{2,\tau}$	0.5	-0.002	0.028	0.027	0.027	0.943	0.944
			0.7	-0.003	0.029	0.029	0.029	0.955	0.959
		t_τ	0.5	0.003	0.124	0.112	0.113	0.923	0.922
			0.7	0.007	0.128	0.118	0.119	0.927	0.926
		γ_τ	0.5	0.001	0.010	0.010	0.010	0.956	0.954
			0.7	0.000	0.010	0.010	0.010	0.938	0.942
		$\beta_{1,\tau}$	0.5	0.004	0.097	0.090	0.090	0.925	0.923
			0.7	0.008	0.102	0.093	0.093	0.930	0.932
		$\beta_{2,\tau}$	0.5	-0.004	0.068	0.062	0.062	0.924	0.926
			0.7	-0.004	0.069	0.064	0.064	0.925	0.929
		t_τ	0.5	0.002	0.272	0.221	0.221	0.894	0.897
			0.7	0.002	0.280	0.232	0.232	0.897	0.900
8000	Hetero.	$\beta_{1,\tau}$	0.5	0.001	0.017	0.017	0.017	0.949	0.951
			0.7	-0.001	0.018	0.018	0.018	0.951	0.953
		$\beta_{2,\tau}$	0.5	0.003	0.053	0.053	0.053	0.948	0.947
			0.7	0.002	0.058	0.055	0.055	0.935	0.936
		t_τ	0.5	0.000	0.033	0.033	0.033	0.944	0.943
			0.7	-0.002	0.036	0.034	0.035	0.943	0.942
		γ_τ	0.5	-0.005	0.190	0.176	0.175	0.920	0.920
			0.7	0.004	0.200	0.184	0.184	0.922	0.926
		$\beta_{1,\tau}$	0.5	0.000	0.014	0.014	0.014	0.953	0.957
			0.7	0.000	0.014	0.015	0.015	0.943	0.948
8000	Comp. Sym.	$\beta_{1,\tau}$	0.5	0.000	0.020	0.019	0.019	0.936	0.939
			0.7	0.000	0.020	0.020	0.020	0.941	0.942
		$\beta_{2,\tau}$	0.5	-0.001	0.020	0.019	0.019	0.949	0.947
			0.7	-0.001	0.021	0.021	0.021	0.939	0.938
		t_τ	0.5	0.004	0.084	0.079	0.079	0.944	0.945
			0.7	0.000	0.086	0.084	0.084	0.937	0.935
		γ_τ	0.5	0.000	0.007	0.007	0.007	0.935	0.933

<i>n</i>	Corr. Struct.	Parameter	τ	Bias	SE	SEE		CP	
						Specific	General	Specific	General
AR(1)		$\beta_{1,\tau}$	0.7	0.000	0.007	0.007	0.007	0.946	0.947
			0.5	0.003	0.068	0.066	0.065	0.933	0.934
		$\beta_{2,\tau}$	0.7	0.004	0.073	0.068	0.068	0.926	0.925
			0.5	-0.003	0.044	0.043	0.043	0.952	0.951
		t_τ	0.7	-0.004	0.046	0.045	0.045	0.951	0.951
			0.5	0.003	0.172	0.151	0.151	0.910	0.908
		γ_τ	0.7	0.005	0.181	0.158	0.158	0.909	0.910
			0.5	0.000	0.011	0.012	0.012	0.958	0.958
			0.7	0.000	0.012	0.012	0.012	0.960	0.958
Hetero.		$\beta_{1,\tau}$	0.5	0.000	0.038	0.037	0.037	0.947	0.944
			0.7	0.001	0.039	0.039	0.039	0.951	0.949
		$\beta_{2,\tau}$	0.5	-0.001	0.024	0.023	0.023	0.942	0.943
			0.7	-0.001	0.025	0.024	0.024	0.942	0.946
		t_τ	0.5	0.000	0.131	0.124	0.124	0.935	0.938
			0.7	-0.004	0.140	0.130	0.130	0.928	0.927
		γ_τ	0.5	0.000	0.010	0.010	0.010	0.944	0.945
			0.7	0.000	0.010	0.010	0.010	0.959	0.961

n: total number of observations, $n = 5N$ with N being the number of subjects; Corr. Struct.: true correlation structure of $\mathbf{Y}_i$; Comp. Sym.: compound symmetry; Hetero.: heteroscedastic. Bias and SE are bias and standard error of the parameter estimator, SEE is the mean of the standard error estimator, and CP is the coverage rate of the 95% confidence interval. Specific means using the estimator of $\delta_{ijj}^{(\tau)}$ that is specific to the correlation structure of $\mathbf{Y}_i$; General means using the estimator of $\delta_{ijj}^{(\tau)}$ that does not depend on the correlation structure of $\mathbf{Y}_i$.

Table 2

Characteristics of the sample of 429 subjects for analysis

Characteristic	Analysis
Gender—no. (%)	
Male	115 (26.8%)
Female	314 (73.2%)
Years of education—no. (%)	
> 12 (attended college)	89 (20.7%)
12 (not attended college)	340 (79.3%)
ApoE4 status—no. (%)	
carrier	127 (29.6%)
non-carrier	274 (63.9%)
unknown	28 (6.5%)
Study—no. (%)	
MAP	189 (44.1%)
ROS	240 (55.9%)
Age at baseline (yrs)—mean (sd)	81.48 (6.57)
Age at AD diagnosis (yrs)—mean (sd)	85.79 (6.62)
No. of cognitive evaluations—mean (sd)	7.1 (2.8)

Table 3

Estimates and 95% CIs for the Parameters of Model (22)

	τ		
	0.2	0.5	0.8
α_τ	0.12 [-0.09, 0.33]	0.39 [0.24, 0.54]	0.70 [0.59, 0.80]
β_τ	-0.28 [-0.32, -0.25]	-0.23 [-0.26, -0.21]	-0.16 [-0.19, -0.13]
t_τ	-4.4 [-4.9, -3.9]	-4.6 [-5.0, -4.1]	-5.0 [-5.7, -4.2]
γ_τ	-0.025 [-0.037, -0.013]	-0.013 [-0.021, -0.005]	-0.010 [-0.016, -0.004]
a_τ	-0.73 [-1.13, -0.33]	-0.53 [-0.92, -0.15]	-0.18 [-0.55, 0.20]
b_τ	0.023 [0.002, 0.044]	0.011 [-0.008, 0.031]	0.000 [-0.021, 0.021]