

A functional proportional hazard cure rate model for interval-censored data

Statistical Methods in Medical Research

2022, Vol. 31(1) 154–168

© The Author(s) 2021

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/09622802211052972

journals.sagepub.com/home/smm

Haolun Shi¹, Da Ma², Mirza Faisal Beg², and Jiguo Cao¹

Abstract

Existing survival models involving functional covariates typically rely on the Cox proportional hazards structure and the assumption of right censorship. Motivated by the aim of predicting the time of conversion to Alzheimer's disease from sparse biomarker trajectories in patients with mild cognitive impairment, we propose a functional mixture cure rate model with both functional and scalar covariates for interval censoring and sparsely sampled functional data. To estimate the nonparametric coefficient function that depicts the effect of the shape of the trajectories on the survival outcome and cure probability, we utilize the functional principal component analysis to extract the functional features from the sparsely and irregularly sampled trajectories. To obtain parameter estimates from the mixture cure rate model with interval censoring, we apply the expectation-maximization algorithm based on Poisson data augmentation. The estimation accuracy of our method is assessed via a simulation study and we apply our model on Alzheimer's disease Neuroimaging Initiative data set.

Keywords

Cox's model, functional principal component analysis, functional data analysis, proportional hazards model, survival analysis, interval-censored data

Introduction

Alzheimer's disease (AD) is one of the most prevalent age-related neurodegenerative disease worldwide, with no available cure yet. Early prognosis is therefore crucial for planning proper clinical intervention. It is especially true for people diagnosed with mild cognitive impairment (MCI), to whom the prediction of whether and when the future disease onset would happen is particularly valuable. MCI is defined as a transitional stage between the normal age-related decline in cognitive function, and the problematic dementia of a much more severe degree, such as AD. Subjects diagnosed with MCI experience problems with memory/mental function that are more serious than normal age-related changes. For patients with AD, a retrospective study would regard the MCI as a stage of subtle onset which is followed by gradual neurodegenerative progression into AD. However, not all MCI patients would eventually suffer from AD; symptoms of MCI may remain stable until death, develop into other types of dementia, or even ameliorate over time.

One research question that is of particular interest and has attracted considerable attention among AD research community is how to obtain an accurate prognostic prediction of AD onset for MCI patients. Studies have been conducted to investigate this question from multiple facets, which utilize various volumetric, textural, and behavioral data to predict the conversion from MCI to AD. However, such prognostic prediction has been proven to be challenging, and previous studies have only achieved limited success^{1–3}.

To predict the time to conversion from MCI to AD from longitudinal data, various functional survival models have been proposed. Kong et al.⁴ proposed a functional linear Cox regression model that utilizes functional principal component analysis (FPCA) to extract functional features from the surface data. Lee et al.⁵ developed a Bayesian extension of the functional proportional hazards regression model and applied the method on AD Neuroimaging Initiative (ADNI) data set. Gellar et al.⁶ proposed a combination of the penalized signal regression model and mixed-effects proportional hazards

¹Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada

²School of Engineering, Simon Fraser University, Burnaby, BC, Canada

Corresponding author:

Jiguo Cao, Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada.

Email: jiguo_cao@sfu.ca

model, which is able to handle functional covariates. Qu et al.⁷ established the asymptotical properties of the maximum partial likelihood estimator in the functional proportional hazards model.

Existing functional survival models in the literature rely on three assumptions. First, the Cox proportional hazards structure is most commonly adopted as the regression framework. However, not all the stages of MCI would eventually lead to AD, as the state of MCI could also remain stable, develop into other types of neurodegenerative diseases, or even improve and revert back to normal over time. This would suggest that the cure rate model, which assumes the existence of a certain proportion of patients that are insusceptible to AD, would be a more proper choice of modeling.

Second, the time to MCI-to-AD conversion is assumed to be subject to right censoring, whereas the assumption of interval censoring is more appropriate for the application on the ADNI data set. In each follow-up visit, the patient would be diagnosed as one of the three categories: cognitive normal (CN), MCI, or AD. The actual time point of MCI-to-AD conversion is located between the first diagnosis of AD and the last diagnosis of MCI, thus censored within the interval of these two times of follow-up visit. Therefore, the assumption of interval censoring would thus more realistically reflect the actual scheme of censoring of the time to MCI-to-AD conversion.

Third, the employed method for FPCA in existing functional survival models relies on the assumption of a dense sampling scheme, i.e., each data point on the grid along the trajectories is not missing and, ideally, evenly spaced. In the ADNI study, patients were followed up around every 6 months in the first 2 years and subsequently around every 12 months until the end of their follow-up periods, which usually last over 10 years. It is worth emphasizing that the actual time of follow-up and diagnosis is not exactly spaced on a 12-month grid and may differ by the nominal time by several months, thus the assumption of a sparse sampling scheme on a monthly grid, rather than a dense sampling scheme on a yearly grid, would render a more accurate and appropriate depiction of the biomarker trajectories. Therefore, a new method for extracting functional features under such a sparse sampling scheme is needed.

In light of the aforementioned issues, we propose a functional cure rate model with interval censoring and sparse sampling. To estimate the nonparametric coefficient function that depicts the effect of the shape of the trajectories on the survival outcome, we utilize the FPCA through conditional expectation to extract the functional features from the sparsely and irregularly sampled trajectories⁸. To obtain an estimate from the mixture cure model with interval censoring, we apply the expectation-maximization (EM) algorithm based on Poisson data augmentation^{9,10}.

FPCA is a useful tool for extracting the primary mode of variations in the trajectories, thereby decomposing the underlying stochastic process into several orthogonal functional components. These uncorrelated components, referred to as the functional principal components (FPCs), can be further used as a foundation for functional regression under various modeling frameworks. When the functional data are subject to a sparse and irregular sampling design, the principal analysis by conditional expectation (PACE) proposed by Yao et al.⁸ utilized the two-dimensional local smoother to estimate the covariance structure and the standard deviation of the measurement error, followed by eigendecomposition of the estimated covariance function to obtain the estimates of the FPCs, and the calculation of the FPC scores is based on conditional expectation. Hall et al.¹¹ established the validity of the PACE method in terms of the asymptotical statistical properties. After extracting the FPCs, functional regression can then be used to model the relationship between the response of interest and the functional covariates. Hall and Horowitz¹⁴ established the theoretical properties of the estimated covariate effects in the functional regression model. Yao et al.¹³ proposed a functional regression model for sparse data and developed a flexible method for prediction based on FPCA. Yu et al.¹² proposed a partial functional quantile regression for the analysis of neuroimaging data. Zhu et al.¹⁸ proposed a multivariate varying coefficient model that uses the kernel smoothing methods for capturing the relationship between multiple functional responses and covariates, which is particularly suited for the neuroimaging data. A comprehensive review on functional regression can be found in Morris¹⁵, and various generalized models for functional regression can be found in Escabias et al.¹⁶, Cardot and Sarda¹⁷, and Yao and Muller¹⁹.

The estimation of the effects of both baseline and functional covariates is conducted under the mixture cure rate model framework. The mixture cure rate model separates the patient population into susceptible (typical survival data) and insusceptible (with survival probability always equal to one) subgroups. The classic two-component mixture model can be traced back to Berkson and Gage²⁰, where a cure fraction is incorporated in the survival function. Kuk and Chen²¹ combined the Cox regression model with the logistic model for the joint modeling of survival times and cure probabilities. Peng and Dear²² applied the EM algorithm to obtain the maximum likelihood estimator under the Cox proportional hazards cure model. When interval censoring is assumed, Wang et al.⁹ proposed an efficient and flexible proportional hazards model using data augmentation via introducing latent Poisson random variables. Zhou et al.¹⁰ proposed a generalized odds rate model for interval-censored data by incorporating the gamma-Poisson data augmentation. Kim and Jhun²³ proposed a cure rate model based on interval-censored data with frailty. Lam et al.²⁴ proposed a semiparametric cure rate model with frailty and utilized a data augmentation approach for parameter estimation, and Lam and Wong²⁵ extended the model to incorporate clustered data.

The rest of the article is organized as follows: In the “Methodology” section, we present the methodology for the functional cure rate model with interval censoring. In the “Simulation study” section, a simulation study is conducted to assess the empirical

accuracy of the estimation of the covariate effects. In the “Application” section, we apply the proposed method on the ADNI data set to study the effects of the shape of the trajectories of the hippocampus volume on the time to MCI-to-AD conversion and provide interpretation to the results. Finally, the “Discussion” section concludes the paper with a discussion.

Methodology

In the functional mixture cure model, each subject i is characterized with underlying true time-varying covariate trajectories $X_i(s)$ and/or $Z_i(s)$, $s \in \mathcal{S}$, where $\mathcal{S}$ is the observation time window of the trajectories, as well as a non-time-varying covariate vector $\mathbf{x}_i^0$ and/or $\mathbf{z}_i^0$. The $X_i(s)$ and $\mathbf{x}_i^0$ are related to the survival times and the $Z_i(s)$ and $\mathbf{z}_i^0$ related to the cure probabilities. The survival function for subject i is modeled by the two-component mixture cure model

$$S(t|\mathbf{x}_i^0, \mathbf{z}_i^0, X_i, Z_i) = 1 - \pi(\mathbf{z}_i^0, Z_i) + \pi(\mathbf{z}_i^0, Z_i) \cdot S_u(t|\mathbf{x}_i^0, X_i).$$

The cure submodel for the cure probability is

$$1 - \pi(\mathbf{z}_i^0, Z_i) = \text{logistic}\left(\boldsymbol{\eta}_0^\top \mathbf{z}_i^0 + \int_{\mathcal{S}} Z_i(s)\eta(s)ds\right), \quad (1)$$

where $\text{logistic}(x) = \exp(x)/\{\exp(x) + 1\}$, $\boldsymbol{\eta}_0$ is the coefficient vector for the baseline covariates, and $\eta(\cdot)$ the coefficient function for the effect of the biomarker trajectory on the cure probability.

The survival function $S_u(\cdot|\mathbf{x}_i^0, X_i)$ depicts the survival probability of susceptible patients under the proportional hazards assumption

$$S_u(t|\mathbf{x}_i^0, X_i) = \exp\left\{-\exp\left\{H(t) + \boldsymbol{\beta}_0^\top \mathbf{x}_i^0 + \int_{\mathcal{S}} X_i(s)\beta(s)ds\right\}\right\}, \quad (2)$$

where $H(\cdot)$ is a nondecreasing transformation function satisfying $H(0) = -\infty$, $\boldsymbol{\beta}_0$ is the coefficient vector for the baseline covariates, and $\beta(\cdot)$ the coefficient function for the effect of the trajectory on the survival hazard.

Let $H_e(t) = \exp\{H(t)\}$. Following Zhou et al.¹⁰, we adopt the integrated spline expansion for modeling the monotone function $H_e(t)$ as

$$H_e(t) = \sum_{l=1}^k \gamma_l b_l(t), \quad (3)$$

where $b_l(\cdot)$ is a series of nondecreasing integrated spline basis functions that ranges from 0 to 1 and the $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)$ is non-negative to ensure monotonicity²⁶.

Functional principal component analysis

For the ease of notation and without loss of generality, we assume that $X_i(s)$ and $Z_i(s)$ are the same, where $s \in \mathcal{S}$. Let $Y_i(s_{ij}) = X_i(s_{ij}) + e_{ij}$ be the measurement at the j th time point ($j = 1, \dots, n_i$) of the i th subject's trajectory, and e_{ij} is assumed to be a measurement error term following normal distribution with mean 0 and variance σ^2 .

We propose to model the trajectories $X_i(s)$ or $Z_i(s)$ and the coefficients functions $\beta(\cdot)$ and $\eta(\cdot)$ using FPCA. The longitudinal trajectories $X_i(s)$ are modeled as independent realizations from an underlying stochastic process $X(s)$. Let $\mu(s) = E(X(s))$ and $G(r, s) = \text{Cov}(X(r) - \mu(r), X(s) - \mu(s))$ denote the mean function and the covariance function, respectively. Based on the Karhunen-Lovève decomposition, $X_i(s)$ can be expanded as

$$X_i(s) = \mu(s) + \sum_{k=1}^{\infty} \xi_{ik} \phi_k(s), \quad (4)$$

where $\phi_k(s)$ is the k th FPC, and ξ_{ik} is the associated FPC score for the i th subject. The FPCs should satisfy

$$\int_{\mathcal{S}} \phi_k(s)\phi_j(s)ds = \delta_{kj},$$

where $\delta_{kj} = 1$ if $k = j$ and 0 otherwise.

The FPC score is defined as

$$\xi_{ik} = \int_{\mathcal{S}} (X_i(s) - \mu(s))\phi_k(s)ds. \quad (5)$$

The magnitude of ξ_{ik} represents the degree of similarity between the $X_i(s) - \mu(s)$ and the FPC $\phi_k(s)$. The mean and variance of the distribution of ξ_{ik} are $E(\xi_{ik}) = 0$ and $\text{Var}(\xi_{ik}) = \lambda_k$, where $\lambda_1 \geq \lambda_2 \geq 0$.

By Mercer's theorem, the covariance function $G(r, s)$ can be expressed as

$$G(r, s) = \sum_{k=1}^{\infty} \lambda_k \phi_k(r) \phi_k(s). \quad (6)$$

As it is not realistic to obtain estimates of infinite FPCs, we approximate $X_i(s)$ by keeping only the first K FPCs and the corresponding FPC scores.

To estimate the FPCs, the first step is to establish smoothed estimates of the mean and covariance functions. The estimated mean function is obtained by smoothing the data from all the trajectories using the local linear smoother, and the estimated covariance function $G(r, s)$ is obtained using a two-dimensional kernel smoother. We denote the estimated mean function and the estimated smoothed covariance function as $\hat{\mu}(s)$ and $\hat{G}(r, s)$, respectively.

Estimates for $\phi_k(s)$ and λ_k can be solved through the eigenequation

$$\int_S \hat{G}(r, s) \phi_k(r) dr = \lambda_k \phi_k(s), \quad (7)$$

with the constraints $\|\phi_k\|^2 = 1$ and $\langle \phi_k, \phi_j \rangle = 1$ if $k = j$, and 0 otherwise. The solution to such an eigenequation $\hat{\phi}_k(s)$ can be found by applying spectral decomposition to the discretized covariance function $\hat{G}(r, s)$.

Let $\hat{\Phi}_{ik}$ and $\hat{\mu}_i$ denote the vectors of values of $\hat{\phi}_k(s)$ and $\hat{\mu}(s)$ evaluated at time points s_{ij} , and let $\tilde{\mathbf{G}}_i$ denote the matrix of values of $\hat{G}(r, s)$ evaluated at the two-dimensional grid consisting of s_{ij} . Using the approach of conditional expectation⁸, the FPC score of the i th subject and the k th FPC is computed as

$$\hat{\xi}_{ik} = \hat{\mathbf{E}}(\xi_{ik} | \mathbf{Y}_i) = \lambda_k \hat{\Phi}_{ik}^T \hat{\Sigma}_{\mathbf{Y}_i}^{-1} (\mathbf{Y}_i - \hat{\mu}_i),$$

where $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^T$, $\hat{\Sigma}_{\mathbf{Y}_i} = \tilde{\mathbf{G}}_i + \hat{\sigma}^2 \mathbf{I}_{n_i}$.

To model the integral $\int_S X_i(s) \beta(s) ds$ and $\int_S Z_i(s) \eta(s) ds$ in (1) and (2), we consider expanding the coefficient functions $\beta(s)$ and $\eta(s)$ on the bases of FPCs $\{\phi_k(s), k = 1, \dots, \infty\}$, that is,

$$\beta(s) = \sum_{k=1}^{\infty} \beta_k \phi_k(s),$$

$$\eta(s) = \sum_{k=1}^{\infty} \eta_k \phi_k(s),$$

where $\beta_k = \int_S \beta(s) \phi_k(s) ds$ and $\eta_k = \int_S \eta(s) \phi_k(s) ds$ are the basis coefficients. Therefore, the integral can be rewritten as

$$\begin{aligned} \int_S X_i(s) \beta(s) ds &= \int_S \mu(s) \beta(s) ds + \sum_{k=1}^{\infty} \xi_{ik} \beta_k, \\ \int_S Z_i(s) \eta(s) ds &= \int_S \mu(s) \eta(s) ds + \sum_{k=1}^{\infty} \xi_{ik} \eta_k. \end{aligned}$$

The first term, i.e., the integral of the mean function over $\beta(\cdot)$ or $\eta(\cdot)$, is a nuisance term that can be merged into the constant intercept in the cure probability submodel in (1), or merged into $H(t)$ in the survival submodel in (2). We denote $\beta_1 = (\beta_1, \dots, \beta_K)$ and $\eta_1 = (\eta_1, \dots, \eta_K)$.

Mixture cure model

After obtaining the FPC scores $\xi_{X,i} = (\xi_{X,i1}, \dots, \xi_{X,iK})$ and $\xi_{Z,i} = (\xi_{Z,i1}, \dots, \xi_{Z,iK})$ such that

$$\begin{aligned} X_i(s) &= \mu_X(s) + \sum_{k=1}^K \xi_{X,ik} \phi_{X,k}(s), \\ Z_i(s) &= \mu_Z(s) + \sum_{k=1}^K \xi_{Z,ik} \phi_{Z,k}(s), \end{aligned}$$

we rewrite $\mathbf{x}_i = (\mathbf{x}_i^0, \boldsymbol{\xi}_{X,i})$ and $\mathbf{z}_i = (\mathbf{z}_i^0, \boldsymbol{\xi}_{Z,i})$. We denote the observed data as $\mathbf{D} = \{(L_i, R_i, \delta_{L,i}, \delta_{R,i}, \delta_{I,i}, \mathbf{x}_i, \mathbf{z}_i), i = 1, 2, \dots, n\}$, where $\delta_{L,i}$, $\delta_{R,i}$, $\delta_{I,i}$ are the censoring indicators. In particular, $\delta_{L,i} = 1$ when the event time is left censored, $\delta_{R,i} = 1$ when it is right censored and $\delta_{I,i} = 1$ when it is interval censored between L_i and R_i .

Let $\boldsymbol{\beta} = (\boldsymbol{\beta}_0, \boldsymbol{\beta}_1)$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}_0, \boldsymbol{\eta}_1)$. The parameters to be estimated are denoted as $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\gamma})$. The likelihood is

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}|\mathbf{D}) &= \prod_{i=1}^n \pi(\mathbf{z}_i)^{1-\delta_{R,i}} \{1 - \pi(\mathbf{z}_i) + \pi(\mathbf{z}_i)S_u(L_i|\mathbf{x}_i)\}^{\delta_{R,i}} \\ &\quad \times \{1 - S_u(R_i|\mathbf{x}_i)\}^{\delta_{L,i}} \{S_u(L_i|\mathbf{x}_i) - S_u(R_i|\mathbf{x}_i)\}^{\delta_{I,i}}, \end{aligned}$$

where the $S_u(\cdot|\mathbf{x}_i)$ is given in (2).

The estimation procedure for the functional proportional hazards cure rate model can be seen as an extension of the regular proportional hazards model by Wang et al.⁹ or as a special case of the generalized odds rate model proposed by Zhou et al.¹⁰ extended to the functional covariates. We first introduce the latent variables u_i for the cure indicator, which equals 1 if subject i is susceptible and 0 otherwise, that is, $u_i \sim \text{Ber}(\pi(\mathbf{z}_i))$. Denote $\mathbf{u} = (u_1, \dots, u_n)$. Given the latent variables $\mathbf{u}$, the complete data likelihood is

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}|\mathbf{D}, \mathbf{u}) &= \prod_{i=1}^n \pi(\mathbf{z}_i)^{u_i} [1 - \pi(\mathbf{z}_i)]^{1-u_i} \{1 - S_u(R_i|\mathbf{x}_i)\}^{\delta_{L,i}} \\ &\quad \{S_u(L_i|\mathbf{x}_i) - S_u(R_i|\mathbf{x}_i)\}^{\delta_{I,i}} S_u(L_i|\mathbf{x}_i)^{\delta_{R,i}u_i}. \end{aligned}$$

By expanding the baseline hazard function in $S_u(L_i|\mathbf{x}_i)$ with the integrated spline basis functions as in (3), we have

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}|\mathbf{D}, \mathbf{u}) &= \prod_{i=1}^n \pi(\mathbf{z}_i)^{u_i} \{1 - \pi(\mathbf{z}_i)\}^{1-u_i} \times \left[1 - \exp\left\{-H_e(R_i)e^{\boldsymbol{\beta}'\mathbf{x}_i}\right\}\right]^{\delta_{L,i}} \\ &\quad \times \left[\exp\left\{-H_e(L_i)e^{\boldsymbol{\beta}'\mathbf{x}_i}\right\} - \exp\left\{-H_e(R_i)e^{\boldsymbol{\beta}'\mathbf{x}_i}\right\}\right]^{\delta_{I,i}} \\ &\quad \times \left[\exp\left\{-u_i H_e(L_i)e^{\boldsymbol{\beta}'\mathbf{x}_i}\right\}\right]^{\delta_{R,i}}, \end{aligned}$$

where $H_e(t) = \exp\{H(t)\}$.

Next, for each subject i , Next, following Zhou et al.¹⁰, for each subject i , we introduce two Poisson latent variables Y_i and W_i , which satisfy $Y_i|\phi_i \sim \text{Pois}(\lambda_i)$ and $W_i|\phi_i \sim \text{Pois}(\omega_i)$, where

$$\begin{aligned} \lambda_i &= e^{\boldsymbol{\beta}'\mathbf{x}_i} \{\delta_{L,i}H_e(R_i) + \delta_{I,i}H_e(L_i)\}, \\ \omega_i &= e^{\boldsymbol{\beta}'\mathbf{x}_i} [\delta_{I,i}\{H_e(R_i) - H_e(L_i)\} + \delta_{R,i}H_e(L_i)]. \end{aligned}$$

By introducing these latent variables, the problem of seeking the maximum likelihood estimate becomes tractable under the EM algorithm. Such a data augmentation procedure was first proposed by McMahan et al.²⁷ and is well-suited for flexible modeling of the interval censorship. Wang et al.⁹ showed that Y and W were associated with a latent Poisson counting process and the survival time was the time of the first jump from 0 to 1 of such a process. It can be shown that such a construction leads to three cases: $W_i > 0$ and $Y_i = 0$ corresponds to the interval censorship; $W_i = Y_i = 0$ corresponds to the right censorship; $W_i = 0$ and $Y_i > 0$ corresponds to the left censorship.

Based on the introduced latent variables $\mathbf{u}$, $\mathbf{Y} = (Y_1, \dots, Y_n)$, and $\mathbf{W} = (W_1, \dots, W_n)$, the complete data likelihood function is

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta}|\mathbf{D}, \mathbf{u}, \mathbf{Y}, \mathbf{W}) &= \prod_{i=1}^n \pi(\mathbf{z}_i)^{u_i} \{1 - \pi(\mathbf{z}_i)\}^{1-u_i} \times \left\{\frac{1}{Y_i!} (\lambda_i)^{Y_i} e^{-\lambda_i}\right\}^{\delta_{L,i}} \\ &\quad \times \left\{\frac{1}{W_i!} (\omega_i)^{W_i} e^{-(\lambda_i+\omega_i)}\right\}^{\delta_{I,i}} \times \exp(-\delta_{R,i}u_i\omega_i), \end{aligned}$$

subject to $Y_i > 0$ if $\delta_{L,i} = 1$ and $W_i > 0$ if $\delta_{I,i} = 1$.

We may further introduce another layer of Poisson latent variables by replacing Y_i and W_i with a series of k Poisson variables, where k is the number of integrated spline basis functions, i.e., $Y_{il} \sim \text{Pois}(\lambda_{il})$ and $W_{il} \sim \text{Pois}(\omega_{il})$ where

$$\begin{aligned}\lambda_{il} &= \gamma_l e^{\beta' \mathbf{x}_i} \{ \delta_{L,i} b_l(R_i) + \delta_{L,i} b_l(L_i) \}, \\ \omega_{il} &= \gamma_l e^{\beta' \mathbf{x}_i} [\delta_{L,i} \{ b_l(R_i) - b_l(L_i) \} + \delta_{R,i} b_l(L_i)].\end{aligned}$$

It can be verified that $\sum_{l=1}^k Y_{il} \sim \text{Pois}(\lambda_i)$ and $\sum_{l=1}^k W_{il} \sim \text{Pois}(\omega_i)$. The complete likelihood is thus

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta} | \mathbf{D}, \mathbf{u}, \mathbf{Y}, \mathbf{W}) &= \prod_{i=1}^n \pi(\mathbf{z}_i)^{u_i} \{1 - \pi(\mathbf{z}_i)\}^{1-u_i} \\ &\times \left[\prod_{l=1}^k \left\{ \frac{1}{Y_{il}!} \lambda_{il}^{Y_{il}} e^{-\lambda_{il}} \right\}^{\delta_{L,i}} \left\{ \frac{1}{W_{il}!} \omega_{il}^{W_{il}} e^{-(\lambda_{il} + \omega_{il})} \right\}^{\delta_{L,i}} \exp(-\delta_{R,i} u_i \omega_{il}) \right],\end{aligned}$$

subject to $\sum_{l=1}^k Y_{il} > 0$ if $\delta_{L,i} = 1$ and $\sum_{l=1}^k W_{il} > 0$ if $\delta_{L,i} = 1$.

To estimate the parameter $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\gamma})$, we adopt the EM algorithm which iterates between the E-step and M-step as follows:

E-step: The E-step computes the conditional expectation of the complete log-likelihood with respect to the latent variable, given all the observed data and the current estimate of all the parameters. Let c denote the current index of iteration. Following Zhou et al.¹⁰, the conditional expectation is

$$Q(\boldsymbol{\theta}; \boldsymbol{\theta}^{(c)}) = Q_1(\boldsymbol{\eta}; \boldsymbol{\eta}^{(c)}) + Q_2(\boldsymbol{\beta}, \boldsymbol{\gamma}; \boldsymbol{\beta}^{(c)}, \boldsymbol{\gamma}^{(c)}),$$

where

$$\begin{aligned}Q_1(\boldsymbol{\eta}; \boldsymbol{\eta}^{(c)}) &= \sum_{i=1}^n E(u_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) \log\{\pi(\mathbf{z}_i)\} + \{1 - E(u_i | \mathbf{D}, \boldsymbol{\theta}^{(c)})\} \log\{1 - \pi(\mathbf{z}_i)\}, \\ Q_2(\boldsymbol{\beta}, \boldsymbol{\gamma}; \boldsymbol{\beta}^{(c)}, \boldsymbol{\gamma}^{(c)}) &= \sum_{i=1}^n \sum_{l=1}^k \left[\{ \delta_{L,i} E(Y_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) + \delta_{L,i} E(W_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) \} (\log \gamma_l + \boldsymbol{\beta}' \mathbf{x}_i) \right. \\ &\quad \left. - \gamma_l e^{\beta' \mathbf{x}_i} E(u_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) \{ (1 - \delta_{R,i}) b_l(R_i) + \delta_{R,i} b_l(L_i) \} \right],\end{aligned}$$

and

$$\begin{aligned}E(Y_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) &= \frac{\delta_{L,i} \lambda_i^{(c)}}{1 - \exp(-\lambda_i^{(c)})}, \\ E(W_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) &= \frac{\delta_{L,i} \omega_i^{(c)}}{1 - \exp(-\omega_i^{(c)})}, \\ E(Y_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) &= \frac{\lambda_{il}^{(c)}}{\lambda_i^{(c)}} E(Y_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) = \frac{\delta_{L,i} \lambda_{il}^{(c)}}{1 - \exp(-\lambda_i^{(c)})}, \\ E(W_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) &= \frac{\omega_{il}^{(c)}}{\omega_i^{(c)}} E(W_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) = \frac{\delta_{L,i} \omega_{il}^{(c)}}{1 - \exp(-\omega_i^{(c)})}, \\ E(u_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) &= 1 - \delta_{R,i} + \frac{\delta_{R,i} \pi^{(c)}(\mathbf{z}_i) \exp(-\omega_i^{(c)})}{1 - \pi^{(c)}(\mathbf{z}_i) + \pi^{(c)}(\mathbf{z}_i) \exp(-\omega_i^{(c)})}.\end{aligned}$$

M-step: The M-step seeks to maximize the conditional expectation $Q_1(\boldsymbol{\eta}; \boldsymbol{\eta}^{(c)})$ and

$Q_2(\boldsymbol{\beta}, \boldsymbol{\gamma}; \boldsymbol{\beta}^{(c)}, \boldsymbol{\gamma}^{(c)})$ with respect to the parameters. The maximum likelihood estimate of $\boldsymbol{\eta}$ in $Q_1(\boldsymbol{\eta}; \boldsymbol{\eta}^{(c)})$ can be recast as a solution to the quasi-binomial logistic regression between $E(u_i | \mathbf{D}, \boldsymbol{\theta}^{(c)})$ and the covariate $\mathbf{z}$. Following Zhou et al.¹⁰, the

maximum likelihood estimate for $\gamma = (\gamma_1, \dots, \gamma_k)$ is updated as

$$\gamma_l^{(d+1)} = \frac{\sum_{i=1}^n \delta_{L,i} E(Y_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)}) + \delta_{I,i} E(W_{il} | \mathbf{D}, \boldsymbol{\theta}^{(c)})}{\sum_{i=1}^n e^{\boldsymbol{\theta}^{(d)\top} \mathbf{x}_i} E(u_i \phi_i | \mathbf{D}, \boldsymbol{\theta}^{(c)}) \{ (1 - \delta_{R,i}) b_l(R_i) + \delta_{R,i} b_l(L_i) \}}, \quad l = 1, \dots, k.$$

The maximum likelihood estimate for β is updated by maximizing $Q_2(\beta, \gamma(\beta); \beta^{(c)}, \gamma^{(c)})$ with respect to β via the Newton-Raphson type gradient descent algorithm.

Simulation study

Model setup

Simulation studies are carried out to assess the empirical performance of the proposed method. We assume that the survival times follow the proportional hazards structure where the baseline survival time follows an exponential distribution with rate 1. Moreover, the survival time T is assumed to be subject to both right censoring and interval censoring. We generate the right censoring time $C = \min(5, 15 \times U)$, where U is drawn from a Uniform(0, 1) distribution. If the time T is larger than C , the observation would be right-censored at time C , otherwise it would be interval-censored. For interval censoring, we assume that the individuals are followed up periodically, the time of which happens irregularly and varies by individual. The first follow-up time is at $t = 0$ and the duration between each follow-up time is drawn from a Uniform(0.1, 0.5) distribution.

For the covariates, the non-time-varying covariate vector for survival outcome $\mathbf{x}_i^0$ consists of a binary variable and a continuous variable. The binary variable is drawn from a Bernoulli distribution with a probability of 0.5, and the continuous variable from a standard normal distribution. For the non-time-varying covariate vector for cure probability $\mathbf{z}_i^0$, we assume that it consists of an intercept term and $\mathbf{x}_i^0$.

For the longitudinal trajectories $X_i(s)$ and $Z_i(s)$, we assume the two are the same for each individual. We simulate the underlying true trajectories as

$$X_i(s) = \mu(s) + \alpha_{i1}\psi_1(s) + \alpha_{i2}\psi_2(s), \quad i = 1, \dots, n.$$

The mean function $\mu(s)$ is assumed to be constant 0, and the component functions $\psi_1(s)$ and $\psi_2(s)$ are $\psi_1(s) = \sqrt{2}\sin(2\pi s)$ and $\psi_2(s) = \sqrt{2}\cos(2\pi s)$, which fulfill the constraints $\|\psi_j\|^2 = 1$ and $\langle \psi_j, \psi_k \rangle = 1$ if $j = k$, and 0 otherwise. The scores α_{i1} and α_{i2} are independently sampled from normal distributions with mean 0 and decreasing standard deviations of 2 and 1, respectively,

$$\alpha_{i1} \sim N(0, 2^2),$$

$$\alpha_{i2} \sim N(0, 1^2).$$

Finally, the coefficient functions are assumed to be $\beta(s) = \boldsymbol{\beta}_1^\top \boldsymbol{\psi}(s)$ and $\eta(s) = \boldsymbol{\eta}_1^\top \boldsymbol{\psi}(s)$, where $\boldsymbol{\psi}(s) = (\psi_1(s), \psi_2(s))^\top$. We experiment with various values of β_0, β_1, η_0 and η_1 .

Result analysis

We apply the FPCA to the simulated trajectories. In terms of the tuning parameter, the number of FPCs is chosen such that the proportion of variance explained is no less than 99%. The true stochastic process is generated by the linear combination of the two FPCs, and the proposed FPCA procedure is able to select an adequate number of FPCs.

We are primarily interested in the estimation of the coefficients of the functional and scalar covariates. Measuring the accurateness in the estimation of these coefficients allows us to confirm the unbiasedness of the model and understand the operating characteristics in the parameter estimation. We consider several metrics. First, we compare the true coefficient values versus the averaged estimated ones to check for the unbiasedness of the model. Moreover, we compute the average standard deviation of the coefficient estimates to understand its variability. Finally, we compute the 95% coverage probability, i.e., the proportion of times that the 95% confidence interval covers the true value; this metric is expected to be close to 95% to confirm the stability of the model estimation.

We consider both dense and sparse functional data. In the dense case, the observed data points are simulated from a uniform grid $\mathcal{S} = [0, 1]$, with the spacing between two consecutive design points being 0.01.

$$y_{ij} = X_i(s_{ij}) + \epsilon_{ij}, \quad i = 1, \dots, n,$$

where the error terms ϵ_{ij} are randomly drawn from the normal distribution with mean 0 and standard deviation 0.001. In the sparse case, we first simulate the full series of data points from a uniform grid $\mathcal{S} = [0, 1]$ with a spacing of 0.01, as in the dense design. Subsequently, only 25% of the grid points are randomly selected from the trajectory and treated as observed sample points. A total of 1000 trajectories are generated.

For the dense case, Table 1 shows the results averaged across 500 replications for various scenarios. It is evident that the estimation is quite accurate as the estimates are close to the true values. Moreover, the empirical coverage probability of the 95% confidence interval is also very close to the nominal level of 95%, suggesting a sound and stable numerical performance. In general, the model is proven to render satisfactory operating characteristics in terms of estimation accuracy. For the sparse case, Table 2 summarizes the results under the sparse design. It is evident that under the sparse setting, the proposed model still has a good performance in terms of estimation accuracy.

Finally, we gauge the estimation precision of the baseline survival function. Figure 1 shows the averaged estimate of the baseline survival function versus the true one in the simulation study. It can be seen that the estimation is quite accurate because the estimated survival curve matches closely with the true survival function.

Application

Data description

A total of 7656 T1-weighted MPRAGE brain magnetic resonance imaging (MRI) images are obtained from the ADNI database, which contains longitudinally scanned images from 1727 subjects. The subjects' demographic information, along with the diagnoses at all the scanning timepoints are available. Each patient is diagnosed as one of the three categories: CN, MCI or AD. The brain MRI are subdivided into 87 laterally separated anatomical structures using FreeSurfer processing framework (version 5.3). The following pipeline of processing is adopted: affine registration to the MNI205 probability atlas; B1 bias field correction; non-rigid registration to adjust for the local morphological variance; EM-based structural labeling and division^{28,29}.

We choose the hippocampus as the AD-related region of interest. The values of the hippocampus structural volume are first corrected and normalized to adjust for the effects of field strength, sex, and total intracranial volume using a generalized linear model (GLM). Each subject's hippocampus structural volumes are measured around every 6 months in the first 2 years and subsequently every 12 months until the end of their follow-up periods. The actual times of follow-up and diagnosis are not exactly spaced 12 months apart and may differ by the nominal time by several months. Figure 2 shows trajectories of the GLM-corrected hippocampus structural volume, as well as the smoothed mean function.

We construct the data set consisting of trajectories of the volumes of the hippocampus, as well as the time to disease progression to AD. Table 3 shows the demographic information (age, sex, MCI status, number of data points) of the

Table 1. Estimates, empirical standard deviations (SDs), and the 95% empirical coverage probabilities in different scenarios for dense design, averaged over 500 simulation replicates.

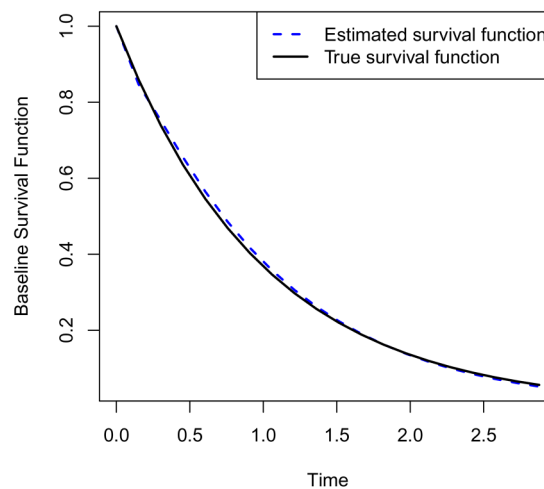
	η_0			η_1		β_0		β_1	
	η_{01}	η_{02}	η_{03}	η_{11}	η_{12}	β_{01}	β_{02}	β_{11}	β_{12}
Scenario 1									
True value	0.1	0.1	-0.5	-0.8	0.2	0.2	0.2	0.5	1
Estimate	0.103	0.089	-0.501	-0.809	0.203	0.205	0.201	0.508	1.009
Average SD	0.131	0.174	0.092	0.069	0.096	0.110	0.057	0.039	0.071
Coverage (95%)	0.930	0.956	0.936	0.930	0.954	0.958	0.932	0.894	0.936
Scenario 2									
True value	0.1	0.2	-0.7	-1	0.2	0.3	0.2	0.7	1
Estimate	0.110	0.202	-0.705	-1.019	0.196	0.304	0.204	0.712	1.013
Average SD	0.143	0.187	0.103	0.085	0.102	0.113	0.059	0.047	0.072
Coverage (95%)	0.924	0.954	0.942	0.952	0.942	0.940	0.960	0.912	0.938
Scenario 3									
True value	0.05	0.1	-1	-1	0.1	0.2	0.3	0.5	2
Estimate	0.040	0.102	-1.005	-1.006	0.111	0.204	0.300	0.501	2.005
Average SD	0.155	0.198	0.118	0.084	0.124	0.123	0.068	0.045	0.117
Coverage (95%)	0.922	0.940	0.924	0.930	0.936	0.962	0.944	0.826	0.940

Table 2. Estimates, empirical standard deviations (SDs), and the 95% empirical coverage probabilities in different scenarios for sparse design, averaged over 500 simulation replicates.

	η_0			η_1		β_0		β_1	
	η_{01}	η_{02}	η_{03}	η_{11}	η_{12}	β_{01}	β_{02}	β_{11}	β_{12}
Scenario 1									
True value	0.1	0.1	-0.5	-0.8	0.2	0.2	0.2	0.5	1
Estimate	0.096	0.113	-0.509	-0.815	0.205	0.196	0.203	0.508	1.018
Average SD	0.133	0.175	0.093	0.071	0.098	0.111	0.057	0.039	0.072
Coverage (95%)	0.936	0.952	0.944	0.938	0.918	0.952	0.948	0.884	0.948
Scenario 2									
True value	0.1	0.2	-0.7	-1	0.2	0.3	0.2	0.7	1
Estimate	0.103	0.206	-0.706	-1.012	0.206	0.300	0.205	0.711	1.018
Average SD	0.142	0.187	0.103	0.085	0.102	0.113	0.059	0.046	0.073
Coverage (95%)	0.906	0.938	0.936	0.948	0.948	0.954	0.940	0.920	0.928
Scenario 3									
True value	0.05	0.1	-1	-1	0.1	0.2	0.3	0.5	2
Estimate	0.058	0.097	-1.015	-1.024	0.090	0.198	0.301	0.506	2.026
Average SD	0.155	0.199	0.119	0.085	0.125	0.123	0.067	0.045	0.117
Coverage (95%)	0.938	0.956	0.952	0.936	0.944	0.944	0.940	0.844	0.944

data set. The goal is to use the normalized volume trajectories of the hippocampus region in the first 2 years as functional data to predict the time from the end of the second year to disease progression into AD. In addition to the functional data, the patient's MCI status, sex and age at enrollment are used as the non-time-varying covariates in the model. A total of 800 non-AD patients (447 MCI and 353 CN) at the end of year 2 are used in the analysis.

As the functional data are sparsely and irregularly sampled, we check for the sparsity of the data points by consolidating all pairs of (s_{ij}, s_{ik}) into a single plot³⁰. If the points are overly sparse, as noted by Yao et al.⁸, the estimation for the covariance function might be infeasible. Figure 3 shows all pairs of measurement time (s_{ij}, s_{ik}) for all subjects' trajectories in the ADNI data set. We observe several dense centers within the plot. We apply the FPCA on the trajectories. Figure 4 shows the fraction of variance explained by the leading FPCs in the ADNI data set. It is worth noting that the FPCs that explain the majority of variation are not necessarily the ones significantly associated with the survival hazards or the cure probability. It is possible that the higher-degree FPCs may have statistical significance. Thus, in terms of the choice of K , the number of FPCs, we repeatedly fit the model for increasing value of K , until no more significant FPCs can be revealed. Based on such a data-driven approach, we choose K to be 5.

**Figure 1.** The point-wise average of the estimated baseline survival function in the simulation study. The solid line is the true baseline survival function.

Analysis of modeling results

Not all the patients in the CN and MCI groups would eventually convert to AD. At the end of the follow-up period, around 70% and 20% of the CN and MCI patients remain in the non-AD group. It is reasonable to assume a cure fraction among the population (particularly from the CN group). We thus apply the proposed functional cure rate model to the data set. Table 4 shows the results of the regression coefficients, standard errors and p -values under our model, where p -values less than or close to 0.05 are highlighted in bold. The model produces estimates for coefficient and the Hessian matrix, and detailed formulae of the estimated Hessian matrix can be found in Wang et al.⁹ and Zhou et al.¹⁰. The standard Wald-type inference can be conducted based on the coefficient estimates and the estimated standard deviations. For example, the p -values are computed from the test statistics that equal the coefficient estimates divided by the estimated standard deviations. For the survival submodel, we observe that the first, third and fifth FPCs are significantly associated with the hazard, with p -values close to 0.01. The MCI status and age are mildly significant with p -values between 0.05 and 0.1. For the cure submodel, the MCI status is significantly associated with the cure fraction. Moreover, the first and third FPCs also have a strong association with the cure fraction.

Figure 5 displays the estimated FPCs. The interpretation for the significant FPCs (first, third and fifth) is as follows. The first FPC is flat above zero, indicating that the level of the horizontal deviation above or below the mean function is associated with the survival hazards and the cure fraction. The coefficients of the first FPC score are negative in both the survival model and the cure model, which imply that the smaller the positive horizontal deviation of the trajectory from the mean value, the larger the survival hazards and the degree of susceptibility to AD. The third FPC is positive before month 18 and negative afterwards, indicating that the contrast between the first 1.5 years and the last 0.5 years has significant effects. The coefficients of the third FPC score are positive in both the survival model and the cure model, which imply that the sharper the contrast, the larger the survival hazards the more likely that the subject would be susceptible to AD. The fifth FPC is positive between months 3 and 16 and negative in the other two intervals, indicating the contrast between the time period from months 3 to 16 and the rest of the time periods is associated with the survival hazards. The coefficient of the fifth FPC score is negative in the survival model, which implies that the sharper the contrast, the smaller the survival hazards.

Discussion

We have proposed a functional mixture cure rate model that incorporates both functional and scalar covariates for interval censoring and sparsely sampled functional data. To the best of our knowledge, our method is the first to apply the cure rate survival model with interval censoring on the functional data. The estimation accuracy of our proposed model is demonstrated via simulation studies. For the real application, the model is well-suited for the ADNI study and our results reveal the significance of the first FPC in both survival and cure submodels.

The EM algorithm based on Poisson data augmentation is a highly effective approach to estimating the parameters in survival models subject to interval censoring. A further extension to generalized odds rate model based on gamma-Poisson

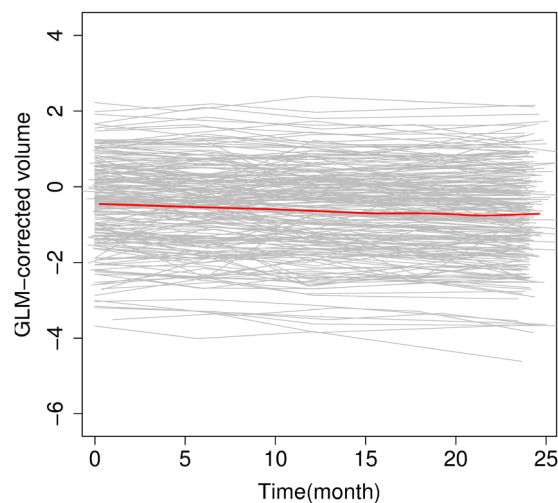


Figure 2. Individual trajectories of the hippocampus volume (gray lines) and the smooth estimate of the mean function (red line).

Table 3. Demographic information including age, sex, MCI status, number of data points for the ADNI data set

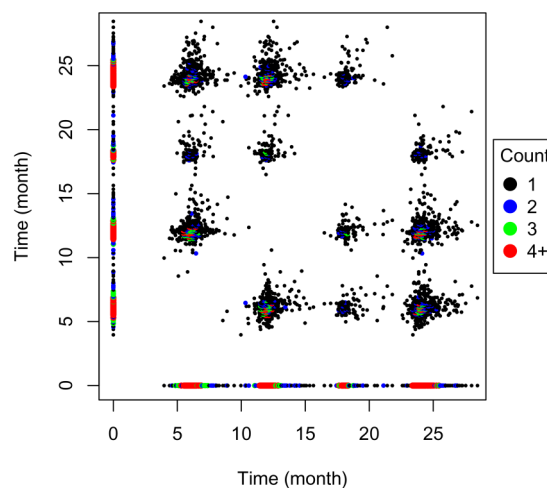
Group	Number of subjects
<i>Age group</i>	
Less than 60	24
Between 60 and 70	209
Between 70 and 80	421
Greater than 80	146
<i>Sex</i>	
Male	353
Female	447
<i>MCI status</i>	
MCI	447
CN	353
<i>Number of data points per subject</i>	
3	47
4	614
5	139

MCI: mild cognitive impairment; ADNI: Alzheimer's Disease Neuroimaging Initiative; CN:cognitive normal.

data augmentation is adopted by Zhou et al.¹⁰. It is feasible to extend our approach to the generalized model proposed by Zhou et al.¹⁰.

One direction of future research is to establish the asymptotical properties of the proposed algorithm. As confirmed by the numerical studies, the method appears to yield consistent estimation of the coefficients in the survival and cure models. Deriving the convergence of the scalar and functional coefficients would be of interest in such a model. There have been some recent developments on the theoretical of the functional proportional hazards Cox model. Kong et al.⁴ considered the consistency and convergence rates of the maximum partial likelihood estimator for the coefficients in the Cox model with dense functional data. Qu et al.⁷ tackled a similar problem and established the asymptotical normality of the scalar coefficient through reproducing kernel Hilbert space. Hao et al.³³ established the asymptotic distribution of the maximum partial likelihood estimator of the slope function using a joint Bahadur representation. The literature on the theoretical prosperities of the cure models is relatively limited; Tsodikov³⁴ discussed the asymptotical efficiency of the cure rate coefficient estimator of the proportional hazard cure model. A potential direction of research would be on the asymptotics of the cure rate coefficient estimator.

The proposed algorithm may suffer from an identifiability issue, which has been well studied in the literature of mixture cure models. For example, Hanin and Huang³¹ discussed the condition for establishing identifiability for mixture cure model. Second, when the sample size is very small (e.g., <100) or when the censoring percentage is very high, the

**Figure 3.** All pairs of measurement time (s_{ij} , s_{ik}) for all subjects' trajectories in Alzheimer's Disease Neuroimaging Initiative (ADNI) data set.

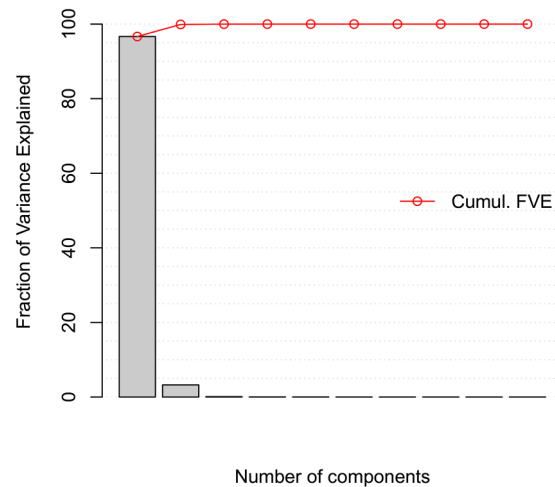


Figure 4. The fraction of variance explained by the number of leading functional principal components in Alzheimer's Disease Neuroimaging Initiative (ADNI) data set.

Table 4. Comparison of the regression coefficients, SE, and *p*-values under the functional cure rate model and functional Cox regression model in the ADNI data set.

	Estimate	SE	<i>p</i> -value
<i>Coefficients: Survival model</i>			
Hippocampus FPC1	−0.467	0.137	0.001
Hippocampus FPC2	−0.030	0.227	0.894
Hippocampus FPC3	0.872	0.369	0.018
Hippocampus FPC4	0.166	0.258	0.521
Hippocampus FPC5	−0.632	0.285	0.026
MCI status	0.919	0.497	0.064
Sex	−0.236	0.238	0.322
Age	−0.245	0.148	0.099
<i>Coefficients: Cure model</i>			
Intercept	1.646	0.825	0.046
Hippocampus FPC1	−1.764	0.589	0.003
Hippocampus FPC2	−1.067	0.593	0.072
Hippocampus FPC3	−3.194	1.655	0.054
Hippocampus FPC4	−1.374	1.034	0.184
Hippocampus FPC5	2.361	1.509	0.118
MCI status	2.288	0.819	0.005
Sex	−0.099	0.720	0.891
Age	0.036	0.547	0.947

SE: standard error; ADNI: Alzheimer's Disease Neuroimaging Initiative; FPC: functional principal component; MCI: mild cognitive impairment.

algorithm might suffer from the problem of non-convergence, e.g., due to the estimated covariance matrix being singular. In these cases, either the EM algorithm may diverge or the Newton–Raphson step would fail to proceed.

In the analysis of the longitudinal trajectories of the GLM—corrected hippocampus structural volume, a few trajectories appear to be outliers and may skew the estimation of the FPC. It is possible to improve the proposed functional cure rate model by incorporating the FPCs estimated from a robust FPCA procedure, see, e.g., Pietrosanu et al.³⁵, and Bali et al.³⁶. Furthermore, Zhang et al.³² proposed the quantile function-on-scalar regression and Yu et al.¹² proposed a partial functional quantile regression for the analysis of neuroimaging data. It would be of interest to explore the use of conditional quantiles to improve the robust estimation procedure.

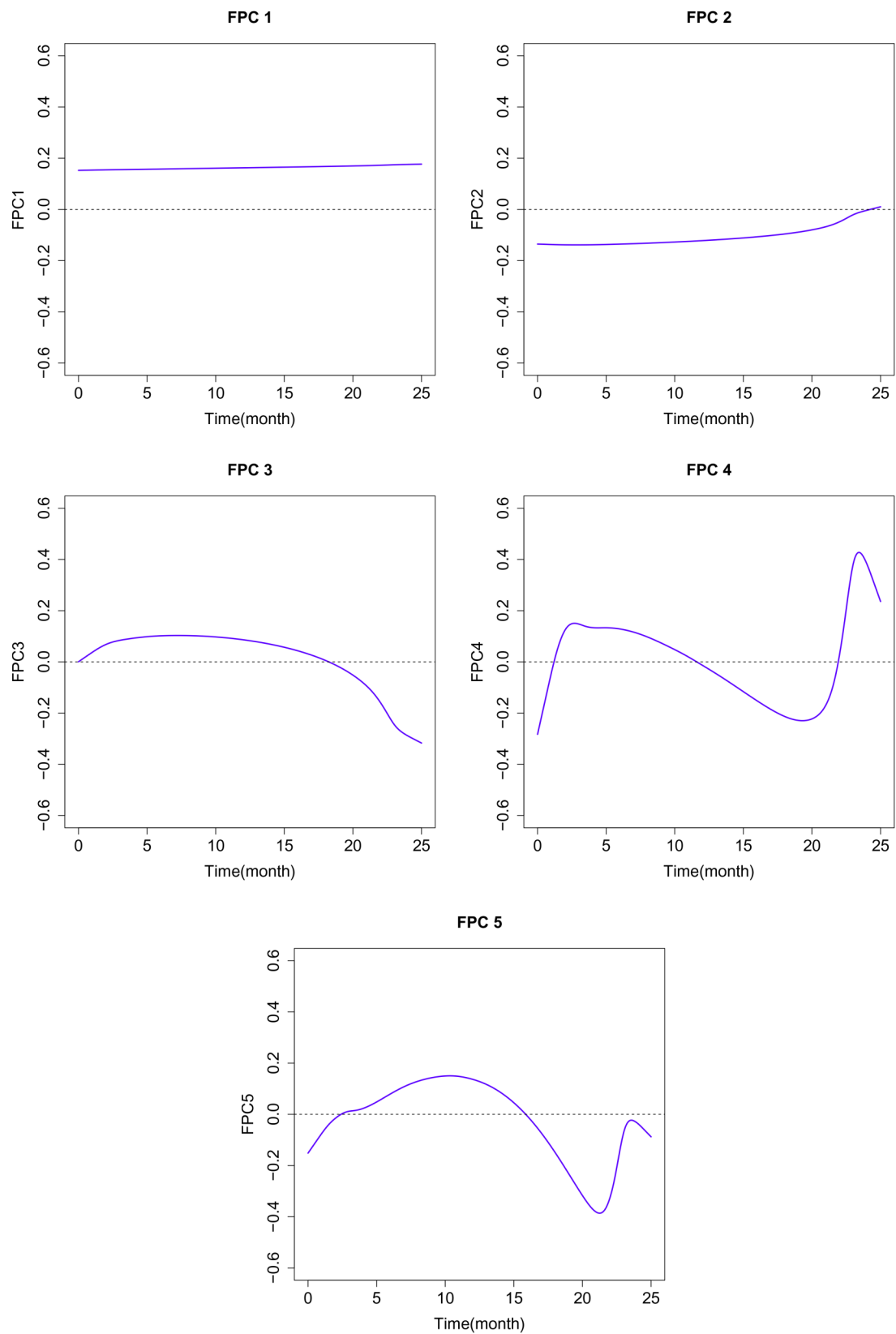


Figure 5. The estimated five FPCs for the normalized hippocampus volume trajectories in the ADNI data set. FPC: functional principal component; ADNI: Alzheimer’s Disease Neuroimaging Initiative.

Acknowledgment

The authors would like to thank the Editor, the Associate Editor and the referees for the constructive comments and suggestions, which have led to a much better exposition of our work.

Declaration of conflicting interests

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

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Part of data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Alzheimer's Association; Alzheimer's Drug Discovery Foundation; BioClinica Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Medpace Inc.; Merck & Co. Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (<http://www.fnih.org>). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the AD Cooperative Study at the University of California, San Diego.

ORCID iD

Jiguo Cao  <https://orcid.org/0000-0001-7417-6330>

References

- Chincarini A, Bosco P, Gemme G, et al. Automatic temporal lobe atrophy assessment in prodromal AD: Data from the DESCRIPA study. *Alzheimer's and Dementia* 2014; **10**: 456–467.
- Elahi S, Bachman AH, Lee SH, et al. Corpus callosum atrophy rate in mild cognitive impairment and prodromal alzheimer's disease. *J Alzheimer's Disease* 2015; **45**: 921–931.
- Khan W, Westman E, Jones N, et al. Automated hippocampal subfield measures as predictors of conversion from mild cognitive impairment to alzheimer's disease in two independent cohorts. *Brain Topogr* 2015; **28**: 746–759.
- Kong D, Ibrahim JG, Lee E, et al. FLCRM: Functional linear cox regression model. *Biometrics* 2018; **74**: 109–117.
- Lee E, Zhu H, Kong D, et al. BFLCRM: A Bayesian functional linear cox regression model for predicting time to conversion to alzheimer's disease. *Ann of Applied Stat* 2015; **9**: 2153–2178.
- Gellar JE, Colantuoni E, Needham DM, et al. Cox regression models with functional covariates for survival data. *Stat Modelling* 2015; **15**: 256–278.
- Qu S, Wang J and Wang X. Optimal estimation for the functional cox model. *Ann of Stat* 2016; **44**: 1708–1738.
- Yao F, Müller HG and Wang JL. Functional data analysis for sparse longitudinal data. *J American Stat Asso* 2005; **100**: 577–590.
- Wang L, McMahan CS, Hudgens M, et al. A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data. *Biometrics* 2016; **72**: 222–231.
- Zhou J, Zhang J and Lu W. Computationally efficient estimation for the generalized odds rate mixture cure model with interval-censored data. *J Computational and Graphical Stat* 2018; **27**: 48–58.
- Hall P, Müller HG and Wang JL. Properties of principal component methods for functional and longitudinal data analysis. *Ann of Stat* 2006; **34**: 1493–1517.
- Yu D, Kong L and Mizera I. Partial functional linear quantile regression for neuroimaging data analysis. *Neurocomputing* 2016; **195**: 74–87.
- Yao F, Müller HG and Wang JL. Functional linear regression analysis for longitudinal data. *Ann of Stat* 2005; **33**: 2873–2903.
- Hall P and Horowitz JL. Methodology and convergence rates for functional linear regression. *Ann of Stat* 2007; **35**: 70–91.
- Morris JS. Functional regression. *Ann Review Stat Appl* 2015; **2**: 321–359.
- Escabias M, Aguilera AM and Valderrama MJ. Principal component estimation of functional logistic regression: discussion of two different approaches. *J Nonparametric Stat* 2004; **16**: 365–384.
- Cardot H and Sarda P. Estimation in generalized linear models for functional data via penalized likelihood. *J Multivariate Analysis* 2005; **92**: 24–41.
- Zhu H, Li R and Kong L. Multivariate varying coefficient model for functional responses. *Ann Stat* 2012; **40**: 2634–2666.
- Yao F and Müller HG. Functional quadratic regression. *Biometrika* 2010; **97**: 49–64.
- Berkson J and Gage RP. Survival curve for cancer patients following treatment. *J American Stat Asso* 1952; **47**: 501–515.

21. Kuk AYC and Chen CH. A mixture model combining logistic regression with proportional hazards regression. *Biometrika* 1992; **79**: 531–541.
22. Peng Y and Dear KBG. A nonparametric mixture model for cure rate estimation. *Biometrics* 2000; **56**: 237–243.
23. Kim YJ and Jhun M. Cure rate model with interval censored data. *Stat in Medicine* 2008; **27**: 3–14.
24. Lam KF, Wong KY and Zhou F. A semiparametric cure model for interval-censored data. *Biometrical J* 2013; **55**: 771–788.
25. Lam KF and Wong KY. Semiparametric analysis of clustered interval-censored survival data with a cure fraction. *Comp Stat & Data Anal* 2014; **79**: 165–174.
26. Ramsay JO. Monotone regression splines in action. *Stat science* 1998; **3**: 425–441.
27. McMahan CS, Wang L and Tebbs JM. Regression analysis for current status data using the EN algorithm. *Stat in Medicine* 2013; **25**: 4452–4466.
28. Fischl B. Automatically parcellating the human cerebral cortex. *Cereb Cortex* 2004; **14**: 11–22.
29. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 2006; **31**: 968–980.
30. Che M, Kong L, Bell R, et al. Trajectory modeling of gestational weight: a functional principal component analysis approach. *PLoS ONE* 2017; **12**: e0186761.
31. Hanin L and Huang LS. Identifiability of cure models revisited. *J Multivariate Ana* 2014; **130**: 261–274.
32. Zhang Z, Wang X, Kong L, et al. High-dimensional spatial quantile function-on-scalar regression. *J of Amer Stat Asso* 2021; DOI: 10.1080/01621459.2020.1870984.
33. Hao M, Liu K, Xu W, et al. Semiparametric inference for the functional Cox model. *J Am Stat Assoc* 2021; **116**(535): 1319–1329.
34. Tsodikov A. Asymptotic efficiency of a proportional hazards model with cure. *Stat & Prob Letters* 1998; **39**: 237–244.
35. Pietrosanu M, Gao J and Kong L. Advanced algorithms for penalized quantile and composite quantile regression. *Comput Stat* 2021; **36**: 333–346.
36. Bali JL, Boente G, Tyler DE, et al. Robust functional principal components: a projection-pursuit approach. *Ann Stat* 2011; **39**: 2852–2882.