A Censored Quantile Transformation Model for Alzheimer's Disease Data with Multiple Functional Covariates

Shaopei Ma

School of Statistics, University of International Business and Economics, Beijing, China

Man-lai Tang

Centre of Data Innovation Research, Department of Physics, Astronomy and Mathematics,School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, United Kingdom

Keming Yu

Mathematical Sciences, Brunel University London, Uxbridge, UB8 3PH, London, United Kingdom

Wolfgang Karl Härdle

Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Berlin, Germany

Zhihao Wang

Institute of Statistics and Data Science, Xinjiang University of Finance and Economics, Urumqi, China

Wei Xiong

School of Statistics, University of International Business and Economics, Beijing, China

Xueliang Zhang, Kai Wang, Liping Zhang

Department of Medical Engineering and Technology, Xinjiang Medical University, Urumqi, China.

Maozai Tian*

Center for Applied Statistics, School of Statistics, Renmin University of China, Beijing, China

Department of Medical Engineering and Technology, Xinjiang Medical University, Urumqi, China

Summary.
Alzheimer's disease (AD) is a progressive disease that starts from mild cognitive impairment and may eventually lead to irreversible memory loss. It is imperative to

* The corresponding author is Maozai TIAN: mztian@ruc.edu.cn.

explore the risk factors associated with the conversion time to AD that is usually right-censored. Classical statistical models like mean regression and Cox models fail to quantify the impact of risk factors across different quantiles of a response distribution, and previous research has primarily focused on modeling a single functional covariate, possibly overlooking the interdependence among multiple functional covariates and other crucial features of the distribution. To address these issues, this paper proposes a multivariate functional censored quantile regression model based on dynamic power transformations, which relaxes the global linear assumption and provides more robustness and flexibility. Uniform consistency and weak convergence of the quantile process are established. Simulation studies suggest that the proposed method outperforms the existing approaches. Real data analysis shows the importance of both left and right hippocampal radial distance curves for predicting the conversion time to AD at different quantile levels.

Keywords: ADNI study; Censored quantile regression; Multivariate functional data; Transformation model.

1. Introduction

Alzheimer's disease (AD) is recognized as one of the major healthcare challenges. It is characterized by a progressive decline in cognitive abilities in individuals (LaFerla et al., 2007; Mattson, 2004; Rabin et al., 2019). Globally, over 50 million people are suffering from dementia, with AD being the leading cause. Timely detection of AD is crucial because the disease can cause irreversible brain disorder, and therapeutic intervention is generally more effective during the early stage of the disease. Mild cognitive impairment (MCI) is considered as a transitional stage between normal aging and dementia state (Petersen et al., 1999). Therefore, it is imperative to explore early markers for the diagnosis of AD and predict the progression from MCI to AD for targeted treatment.

Extensive clinical data have been collected by scientists to investigate significant risk factors associated with the progression of AD. Our research is motivated by the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, which provides a comprehensive set of clinical measurements on the participants. These include neuropsychological assessments, biomedical images, and genetic data (Lee et al., 2015; Jiang et al., 2021). Among the imaging techniques used, magnetic resonance imaging (MRI) is widely employed for visualizing brain structure. Numerous studies have demonstrated that the atrophy of medial temporal lobe (MTL) structures, particularly the hippocampus, is closely linked to an increased risk of developing clinical dementia (Reiss and Ogden, 2010; Young et al., 2013). Figure 1 compares the T1-weighted horizontal (left panel) and sagittal (middle and right panels) brain images from a healthy individual (top panels) and an AD patient (bottom panels) in the ADNI dataset. The left and right hippocampus are marked in green and red. Clearly, the AD patient's hippocampus shows significant atrophy, reinforcing the strong connection between hippocampal shape and cognitive function.

In this study, the AD diagnosis is regarded as the survival event of interest and patients are diagnosed as having AD if they meet the specific inclusion criteria. Thus, the conversion time from MCI to AD is a right-censored response. Previous research has extensively explored the prediction of time-to-event outcomes using functional data analysis

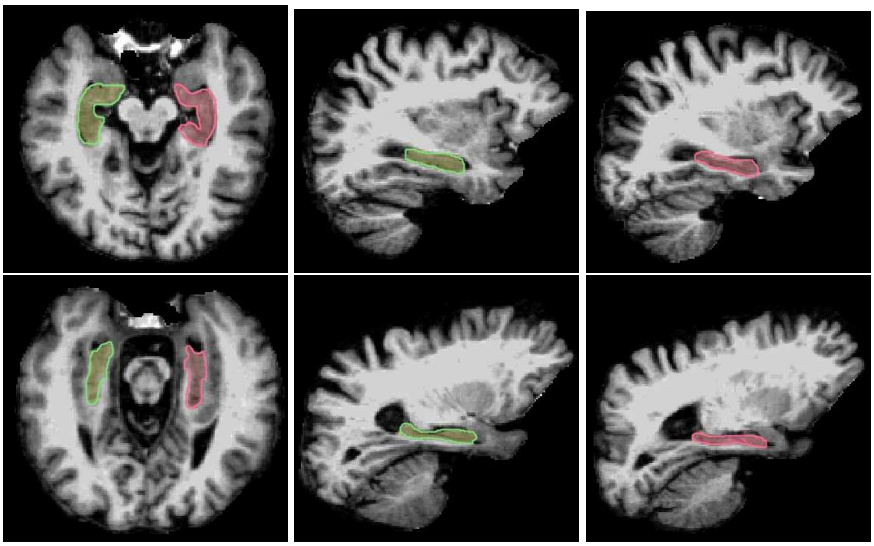

Fig. 1. T1-weighted horizontal (left panels) and sagittal (middle and right panels) brain images from a normal person (top panels) and an AD patient (bottom panels) in the ADNI dataset. The left and right hippocampus are marked in green and red, respectively.

(Yan et al., 2017, 2018). Li and Luo (2017) utilized the structural shape of hippocampus as the functional biomarker of interest. However, their model only incorporated one functional predictor, limiting the utilization of available data and potentially affecting the predictive power. Li and Luo (2019) proposed to use multiple longitudinal biomarkers to predict the disease progression, but their method relied on the assumption of proportional hazards, which may not hold in practice. Both Lee et al. (2015) and Kong et al. (2018) developed the functional linear Cox regression models, highlighting the predictive value of hippocampus surface data in MCI conversion. Nonetheless, all of these previous studies exclusively adopted Cox models, which are inadequate for capturing the higher or lower quantiles of the survival time.

The aim of this paper is to study the association between the conversion time and various clinical, genetic and hippocampus surface variables at baseline within the framework of quantile regression. However, the ADNI dataset presents specific challenges due to (i) correlations among multiple functional covariates (left and right hippocampus surface data) and (ii) the potential non-linear and complex relationship between the covariates and the time-to-event outcome across different quantiles. Consequently, a crucial question arises: how can these data be effectively utilized to enhance the estimation of covariate effects and identify the most effective early markers of AD conversion?

In this paper, we developed a novel multivariate functional quantile regression model to explore the predictor effects on the survival time in the presence of random censoring. Censored quantile regression has gained significant attention for exploring covariate effects across different quantiles in survival data analysis (Koenker and Geling, 2001; Bang and Tsiatis, 2002; Chernozhukov and Hong, 2002). Existing methods, such as the recursively

reweighted estimation procedure by Portnoy (2003) for the Kaplan-Meier estimator and the martingale-based estimation method by Peng and Huang (2008), are well-established for classical censored quantile regression. In most cases, quantile estimation is performed individually for each quantile level without any additional parametric structure assumed for the quantile coefficient functions. Frumento and Bottai (2016, 2017) suggested using a parametric model to describe quantile regression coefficient functions through a finite-dimensional parameter vector. However, these methodologies primarily concentrate on scalar predictors and do not incorporate functional covariates such as curves or images. To the best of our knowledge, the existing research in this area is limited to Jiang et al. (2020) and Wu et al. (2023), who investigated partially linear functional censored quantile regression models. However, these methods are designed specifically for the case of a single functional covariate and are not suitable for our multiple regression setting.

To address this issue, this paper develops a new functional censored quantile regression model to account for multiple correlated functional predictors. These functional covariates can be defined on different domains, such as curves and images, which may differ in dimension. Considering the commonly observed correlations among these covariates, we jointly model them as a multivariate functional predictor and employ multivariate functional principal component (mFPC) analysis to extract useful information (Happ and Greven, 2018). The scalar predictors and the estimated mFPC scores are then incorporated into the survival model. Estimation is based on the martingale procedure described in Peng and Huang (2008). However, it should be noted that the theoretical development poses non-trivial challenges due to two main reasons: (i) the covariates in the model include estimated mFPC scores, which are subject to contamination from estimation errors; (ii) the number of mFPCs is treated as a tuning parameter that diverges with the sample size. In addition, the use of logarithmic transformation on survival time is widely accepted in quantile regression. However, a common limitation is that the log transformation is typically implemented uniformly across all considered quantile levels, which may not accurately reflect the real data and can lead to incorrect inferences. In contrast, the Box-Cox transformation (Box and Cox, 1964) is more suitable to take into account different model structures in censored quantile regression (Yin et al., 2008; Leng and Tong, 2014) and the log transformation can be obtained as a special case. Therefore, in this paper, we introduce the Box-Cox transformation into our proposed model to relax the traditional global assumption of logarithmic linearity. Specifically, the unknown transformation parameter is allowed to dynamically vary with the quantile values. This extension eliminates the potentially restrictive global linearity assumption required in Peng and Huang (2008) and provides flexible non-linear structures at different quantile levels (Chu et al., 2021).

The main contributions of this paper are as follows. First, the proposed model provides a theoretically robust and computationally efficient approach for incorporating multivariate functional covariates in the presence of random censoring within the framework of quantile regression. Such a procedure successfully accommodates dynamic transformations on the response and functional predictors from diverse domains, and is generally applicable in scenarios where predictors exhibit non-negligible correlations, which are commonly encountered in practical applications. To address the issue of multicollinearity in the regression analysis, the informative signals are extracted using mFPC. Second, our procedure is rigorously justified to enjoy favorable sampling properties, such as uniform consistency

and weak convergence of the quantile process. Notably, as the sample size increases, the number of principal components in our model tends to infinity, which presents challenges to the theoretical analysis. Third, our procedure does not rely on the restrictive global linearity assumption, which makes our procedure more applicable in real scenarios.

The rest of the paper is organized as follows. In Section 2, we describe the motivating ADNI study and the data structure. Section 3 establishes the transformed censored quantile regression model for the multivariate functional predictor. The estimation procedure is constructed in Section 4. The large-sample properties of the proposed estimators are discussed in Section 5. Extensive numerical studies in Section 6 show the performance of the proposed method in various settings, especially in the case where curves and images are simultaneously collected as functional covariates. The analysis of the ADNI data is conducted in Section 7. Concluding remarks are given in Section 8. Technical proofs and extra numerical results are deferred to the Supplementary materials.

2. The Motivating ADNI Dataset

The development of the proposed method is motivated by the Alzheimer's Disease Neuroimaging Initiative (ADNI) study, which collected various neuropsychological assessments, brain images and clinical measurements of the participants. The ADNI study was first launched in year 2004 as a global public-private partnership led by Dr. Michael W. Weiner, a professor of radiology, medicine and neurology at the University of California San Francisco (UCSF). The primary goal of ADNI is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), neuropsychological assessments, and other biological markers can be combined to detect Alzheimer's disease as early as possible and identify ways to track the disease progression with specific biomarkers. This research study aims to build a supportive environment for researchers to develop new treatments that could slow or stop the progression of AD and reduce the cost of clinical trials. ADNI recruits participants between age 55 and 90 at 57 sites in the United States and Canada. After obtaining informed consent, participants undergo a series of tests, including a neuropsychological evaluation, genetic testing, and MRI scans, to collect various clinical, cognitive, genetic and imaging data. The first phase of ADNI (ADNI-1) enrolled 800 adults, including approximately 200 people for elderly controls, 400 people with MCI, and 200 people with early AD. Subsequently, the three extensions, ADNI-GO, ADNI-2, and ADNI-3, further recruited new participants into the cohort. All data generated by the ADNI study are entered into the data repository hosted at the Laboratory of Neuroimaging (LONI) at the University of Southern California. Qualified researchers worldwide can submit an online data access request and generally begin using ADNI data within a few days of request submission. For up-to-date information about ADNI, see http://www.adni-info.org.

In this paper, we selected 373 patients who had been diagnosed with MCI from phase 1 of the ADNI study (ADNI-1) to assess the progression of MCI and predict the conversion time from MCI to AD. Among the 373 MCI patients, 161 individuals converted to AD during the study period while the remaining 212 subjects did not progress further before the end of the study. Therefore, the conversion time from MCI to AD is a right-censored response. In addition, each individual's clinical, genetic and hippocampus surface data at

Table 1. Summary statistics about ADNI data, including $n = 373$ participants' clinical and genetic variables.

Variable	Mean $\pm$ SD / No.(%)
Age	74.95 $\pm$ 7.33
Education length	15.65$\pm$ 3.03
ADAS-Cog score	11.58 $\pm$ 4.45
Gender	
Male	237(63.5)
Female	136(36.5)
Handedness	
Right	342(91.7)
Left	31(8.3)
Marital status	
Married	300(80.4)
Widowed	45(12.1)
Divorced	24(6.4)
Never married	4(1.1)
Retirement	
Yes	303(81.2)
No	70(18.8)
APOE-ϵ_4	
Carrier	169(45.3)
Non-carrier	204(54.7)

baseline are included to account for the potential risks. The clinical covariates are Gender (0 = Male; 1 = Female), Handedness (0 = Right; 1 = Left), Marital status (1 = Married; 2 = Widowed; 3 = Divorced; 4 = Never married), Length of education, Retirement (1 = Yes; 0 = No), Age and the ADAS-Cog score. Marital status is coded by 3 dummy variables: "Widowed", "Divorced", "Never married". The ADAS-Cog test has been widely used to assess the severity of dysfunction in adults, with a higher score indicating poorer cognitive function. The genetic covariates concern two Apolipoprotein E (APOE) SNPs which define a 3-allele haplotype, namely, the ϵ_2, ϵ_3 and ϵ_4 variants. Among these variants, APOE-ϵ_4 is known to be a risk factor for the early onset of AD (Petersen et al., 2005). In this paper, we specifically investigate the impact of APOE-ϵ_4 as the covariate of interest on disease progression. The variant presence for these patients is obtained from the ADNI database.

The demographic information summary about the data set is presented in Table 1. The average age of subjects was 74.95 years with a standard deviation of 7.33 years. The youngest person was 55 years old, while the oldest person was 90 years old. On average, the participants had 15.65 years of education with a standard deviation of 3.03 years. The minimum education length was 4 years and the maximum length was 20 years. The average ADAS-Cog score was 11.58 with a standard deviation of 4.45. Among all the individuals, 237 participants were male and 136 were female; 342 were right-handed and 31 were left-handed; 303 were retired and 70 were not. Regarding Marital status, 300 were married, 45 were widowed, 24 were divorced, and 4 were never married at baseline. As for the genetics information on the APOE, 169 subjects carry at least one APOE-ϵ_4 allele and 204 subjects are non-carriers.

To investigate atrophy of the hippocampus, many researchers have conducted analysis based on the volumes of brain structures (Cui et al., 2011; Li et al., 2013), while some other studies have shown that surface-based morphology analysis has more advantages since it provides pointwise effect estimation of the subregion atrophy for cognitive impairment (Li and Luo, 2017; Wang et al., 2020). Mapped onto the hippocampal surface, the subfields are illustrated in Figure 1 of the Supplement. Thus, in this paper, we considered the hippocampal radial distance (HRD) of the left and right hippocampal surface point (referred to as vertex), which measures the distance from its medial core to each vertex and represents the thickness of hippocampus. For hippocampus 3D image processing, we followed the procedure commonly used by Wang et al. (2011) and Luders et al. (2013). Given the 3D MRI scans, we used FreeSurfer to segment hippocampal substructures and automatically reconstruct hippocampal surfaces. Then the left and right hippocampal surfaces are conformally mapped to a two-dimensional (2D) rectangle plane to form a feature image of the surface. After registering each feature image to a common template and calculating HRD for all vertexes, the 2D image matrices are vectorized into one-dimensional (1D) radial distance curves defined on the interval [0,1] (Shi et al., 2013, 2014), which are incorporated as the functional predictors into the proposed model. For detailed information about the hippocampus image processing procedure, readers could refer to Colom et al. (2013).

Figure 2 displays the left and right hippocampal radial distance curves from 35 randomly selected MCI subjects, where the red curves come from the individuals progressed to AD prior to the end of study and the black curves come from the non-converters whose conversion times are censored. It is observed that the hippocampal radial distance curves of the converters are generally lower than the non-converters, i.e., the sizes of hippocampi may be smaller for the patients who progressed quicker to AD. Although the overall trends of the left and right hippocampal radial distance (HRD) curves show similarities, there are notable differences in terms of gradient and volatility. To explore the correlation between these curves, the study conducted a heatmap analysis on the 373 curve pairs shown in the left panel of Figure 3. The analysis revealed that regions spanning locations 0.2-0.8 of the hippocampus exhibit correlation coefficients around 0.6, indicating a significant relationship between the left and right HRD curves. Further, we separately conducted univariate functional principal component analysis (FPCA) for the left and right distance curves. The right panel of Figure 3 shows the empirical correlations of the first three univariate FPCA scores for the curves. Obviously, there is a nonnegligible correlation between the score pairs of the two functional predictors, which may lead to multicollinearity issues in the regression analysis if these correlated scores are incorporated in the model.

3. The Proposed Model

Let T denote the survival time, C denote the censoring time. Define $Y = T \wedge C$, $\delta = I(T \leq C)$, where $\wedge$ is the minimum operator and $I(\cdot)$ is the indicator function. The covariates include a p-dimensional vector $\mathbf{X} \in \mathbb{R}^p$ and a multivariate functional predictor $\mathbf{Z}(\boldsymbol{s}) = (Z_1(s_1), \ldots, Z_d(s_d))^\top \in \mathbb{R}^d$, where each functional element $Z_j(s_j) : \mathcal{S}_j \to \mathbb{R}$ is a stochastic process defined on the compact domain $\mathcal{S}_j$ and is assumed to be in $L^2(\mathcal{S}_j)$ with possibly different domains $\mathcal{S}_j$ for $j = 1, ..., d$. The argument $\boldsymbol{s} := (s_1, \ldots, s_d) \in \mathcal{S} := \mathcal{S}_1 \times \cdots \times \mathcal{S}_d$. This setting allows curves and images to be simultaneously included

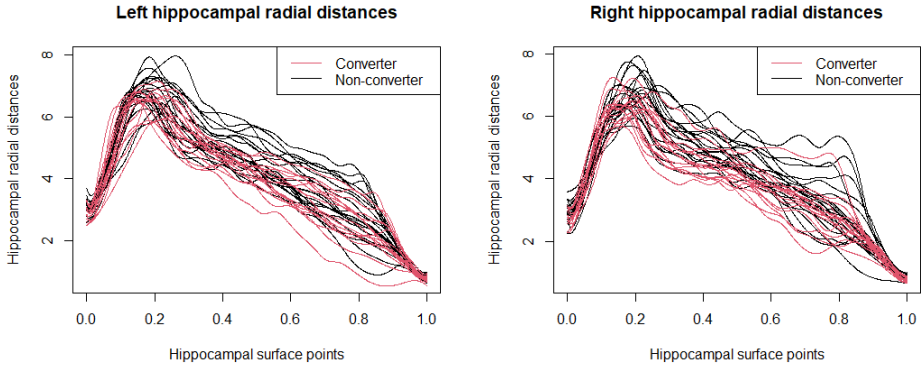

Fig. 2. The left and right hippocampal radial distances curves from 35 randomly selected MCI patients. The red curves are from the converters (individuals progressed to AD) and the black curves are from the non-converters. The horizontal axis represents the considered hippocampal surface points which are scaled to [0,1] for simplicity.

as functional covariates in the model, which greatly extends the application of the regression model. Let $\{Y_i, \delta_i, \mathbf{X}_i, \mathbf{Z}_i\}_{i=1}^n$ be independent and identically distributed copies of $\{Y, \delta, \mathbf{X}, \mathbf{Z}\}$. Conditional on the covariates $\{\mathbf{X}, \mathbf{Z}\}$, the censoring time C is assumed to be independent of the survival time T.

Given the covariates $\{\mathbf{X}, \mathbf{Z}\}$ and $\tau \in (0,1)$, the τth conditional quantile of the random variable T is defined as $Q_T(\tau|\mathbf{X}, \mathbf{Z}) = \inf\{t : P(T \leq t|\mathbf{X}, \mathbf{Z}) \geq \tau\}$. Denote $h_\gamma(\cdot)$ as a family of monotonic transformations with an unknown parameter γ, which includes the following Box-Cox transformation (Box and Cox, 1964) as a special case:

$$h_\gamma(T) = \begin{cases} \frac{T^\gamma - 1}{\gamma} & \text{if } \gamma \neq 0, \\ \log(T) & \text{if } \gamma = 0. \end{cases} \tag{1}$$

Then the proposed transformed quantile linear regression model is as follows:

$$Q_{h_{\gamma(\tau)}(T)}(\tau \mid \mathbf{X}, \mathbf{Z}; \boldsymbol{\beta}, \boldsymbol{\alpha}) = \boldsymbol{\beta}(\tau)^\top \mathbf{X} + \langle\langle \boldsymbol{\alpha}(\mathbf{s}, \tau), \mathbf{Z}(\mathbf{s}) \rangle\rangle, \tag{2}$$

where $\gamma(\tau)$ is the unknown transformation parameter for a given $\tau \in (0,1)$, $\boldsymbol{\beta}(\tau)$ is the p-dimensional regression coefficient vector, and $\boldsymbol{\alpha}(\mathbf{s}, \tau) = (\alpha_1(s_1, \tau), \ldots, \alpha_d(s_d, \tau))^\top$ is the d-dimensional vector of unknown square integrable coefficient functions. As defined in Happ and Greven (2018), the scalar product $\langle\langle \boldsymbol{\alpha}(\mathbf{s}, \tau), \mathbf{Z}(\mathbf{s}) \rangle\rangle = \sum_{j=1}^d \langle \alpha_j, Z_j \rangle_2 = \sum_{j=1}^d \int_{\mathcal{S}_j} \alpha_j(s_j, \tau) Z_j(s_j) ds_j$ and the norm induced by $\langle\langle \cdot \rangle\rangle$ is denoted by $|||\cdot|||$.

Based on the equivariance property of quantile regression to the monotone transformation $h_\gamma(\cdot)$, we can obtain

$$Q_{h_{\gamma(\tau)}(T)}(\tau \mid \mathbf{X}, \mathbf{Z}) = h_{\gamma(\tau)}\left(Q_T(\tau \mid \mathbf{X}, \mathbf{Z})\right), \quad \tau \in (0,1). \tag{3}$$

Then model (2) is equivalent to the following expression in terms of the quantile function of T:

$$Q_T(\tau \mid \mathbf{X}, \mathbf{Z}; \boldsymbol{\beta}, \boldsymbol{\alpha}, \gamma) = h_{\gamma(\tau)}^{-1}\left(\boldsymbol{\beta}(\tau)^\top \mathbf{X} + \langle\langle \boldsymbol{\alpha}(\boldsymbol{s}, \tau), \mathbf{Z}(\boldsymbol{s}) \rangle\rangle\right), \tau \in (0,1), \tag{4}$$

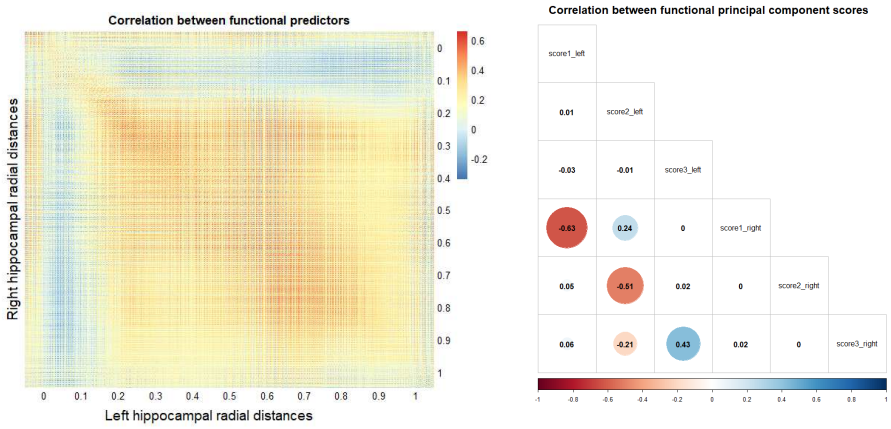

Fig. 3. Heatmaps of the correlation matrix between the 373 left and right hippocampal radial distances. Left Panel: Correlation between the left and right HRD curves; Right Panel: Correlation between the seperate FPC scores of the left and right distance curves.

where $h^{-1}_{\gamma(\tau)}$ is the inverse function of $h_{\gamma(\tau)}$. Therefore, the proposed model (4) relaxes the traditional global linear assumption by introducing the transformation $h_{\gamma(\tau)}$.

A large literature exists about functional quantile regression when the response T is completely observed (Cardot et al., 2005; Chen and Pouzo, 2012; Kato, 2012). However, when T is randomly censored, studies of functional quantile regression models are quite limited. For example, Jiang et al. (2020) and Wu et al. (2023) investigated the partially linear censored quantile regression model with only a single functional covariate. In practical applications, it is common to have multiple functional predictors that contain critical information which should be included in the model. Hence, the proposed method in this paper (Model 4) encompasses their model as a special case when the functional predictor $\mathbf{Z}(\boldsymbol{s})$ is univariate ($d = 1$) and the transformation parameter $\gamma(\tau)$ is zero (logarithmic transformation). In addition, existing approaches for multivariate functional regression are mainly restricted to functions observed on the same finite, one-dimensional interval (Ma et al., 2019). In contrast, our proposed model can handle multiple functional predictors, which are often correlated and defined on potentially distinct domains, such as curves and images. By introducing a class of dynamic transformations $h_{\gamma(\tau)}$ into the functional censored model, our framework allows for flexible coefficient estimations across various quantile levels. Uniform consistency and weak convergence are also shown in Section 5 for the parametric and functional estimators.

Denote the conditional distribution function of T by $F_T(t|\mathbf{X}, \mathbf{Z}) = \mathrm{P}(T \leq t|\mathbf{X}, \mathbf{Z})$. $\Lambda(t|\mathbf{X}, \mathbf{Z}) = -\log\{1 - F_T(t|\mathbf{X}, \mathbf{Z})\}$ is the cumulative hazard function of T conditional on $\mathbf{X}, \mathbf{Z}$. Define $N_i(t) = I(Y_i \leq t, \delta_i = 1)$ to be the counting process for $i = 1, \ldots, n$ and $H(x) = -\log(1-x)$. Following Peng and Huang (2008), we consider the martingale-based estimation method. Specifically, define the martingale

$$M_i(t) = N_i(t) - \Lambda\left(t \wedge Y_i|\mathbf{X}_i, \mathbf{Z}_i\right). \tag{5}$$

Denote $\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0$ and γ_0 as the true values of $\boldsymbol{\beta}, \boldsymbol{\alpha}$ and γ, respectively. Based on the condi-

tional independent censoring assumption (Fleming and Harrington, 2011), we have

$$\begin{aligned} & E\left[M_i\left\{Q_T\left(\tau \mid \mathbf{X}_i, \mathbf{Z}_i; \boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \gamma_0\right)\right\}\right] \\ = \; & E\left[N_i(Q_T\left(\tau \mid \mathbf{X}_i, \mathbf{Z}_i; \boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \gamma_0\right)) - \Lambda\left(Q_T\left(\tau \mid \mathbf{X}_i, \mathbf{Z}_i; \boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \gamma_0\right) \wedge Y_i | \mathbf{X}_i, \mathbf{Z}_i\right)\right] \\ = \; & 0. \end{aligned} \tag{6}$$

In the martingale above, the hazard $\Lambda(\cdot|\mathbf{X}_i, \mathbf{Z}_i)(i = 1, \ldots, n)$ are unknown functions. However, due to the fact that $F_T\{Q_T\left(\tau \mid \mathbf{X}_i, \mathbf{Z}_i\right)|\mathbf{X}_i, \mathbf{Z}_i\} = \tau$, we have

$$\begin{aligned} & \Lambda\left\{Q_T(\tau \mid \mathbf{X}_i, \mathbf{Z}_i) \wedge Y_i \mid \mathbf{X}_i, \mathbf{Z}_i\right\} \\ = \; & H\left(\tau\right) \wedge H\left(F_T\{Y_i \mid \mathbf{X}_i, \mathbf{Z}_i\}\right) \\ = \; & \int_0^\tau I\left[Y_i \geq Q_T\{u \mid \mathbf{X}_i, \mathbf{Z}_i\}\right] dH(u). \end{aligned} \tag{7}$$

The stochastic integration in Λ has a common grid-based approximation as

$$\begin{aligned} \widetilde{\Lambda}\left[Q_T\left(\tau_j \mid \mathbf{X}_i, \mathbf{Z}_i\right) \wedge Y_i \mid \mathbf{X}_i, \mathbf{Z}_i\right] = \sum_{k=0}^{j-1} & I\left[Y_i \geq Q_T\left(\tau_k \mid \mathbf{X}_i, \mathbf{Z}_i\right)\right] \\ & \times \left\{H(\tau_{k+1}) - H(\tau_k)\right\}, \end{aligned} \tag{8}$$

where $\tau_j \in \mathcal{T}_L = \{0 = \tau_0 < \tau_1 < \cdots < \tau_L = \tau_U < 1\}$ with τ_U an upper bound for estimable quantiles to avoid the identifiability issues due to censoring (Peng and Huang, 2008). Certain theoretical constraints regarding τ_U are described in the asymptotic results, specifically in Condition B5 outlined in Section 3 of the Supplementary material. Denote by $\|\mathcal{T}_L\| = \sup_{1 \leq j \leq L} |\tau_j - \tau_{j-1}|$ the size of $\mathcal{T}_L$. Then the mean zero property of the martingales could be used to construct the estimating equation. The proposed estimators $\{\hat{\boldsymbol{\beta}}(\tau), \hat{\boldsymbol{\alpha}}(\boldsymbol{s}, \tau), \hat{\gamma}(\tau)\}, \tau \in (0, \tau_U]$ are right-continuous step functions which jump on the grids of $\mathcal{T}_L$.

4. Estimation and computation

4.1. Multivariate Functional Principal Component Analysis

Considering the possible correlations among multiple functional predictors $Z_j(s_j), j = 1, \ldots, d$, this paper incorporates the multivariate functional principal component analysis (MFPCA) proposed by Happ and Greven (2018) to extract useful information for quantile regression.

Define the mean function of $\mathbf{Z}(\boldsymbol{s})$ as $\boldsymbol{\mu}(\boldsymbol{s}) = E\{\mathbf{Z}(\boldsymbol{s})\} = \{\mu_1(s_1), \ldots, \mu_d(s_d)\}^\top$, where $\mu_j(s_j) = E\left\{Z_j(s_j)\right\}$. The $d \times d$ covariance matrix of $\mathbf{Z}(\boldsymbol{s})$ is

$$\mathcal{C}(\boldsymbol{s}, \boldsymbol{t}) = E\left[\{\mathbf{Z}(\boldsymbol{s}) - \boldsymbol{\mu}(\boldsymbol{s})\}\{\mathbf{Z}(\boldsymbol{t}) - \boldsymbol{\mu}(\boldsymbol{t})\}^\top\right]$$

with elements $\mathcal{C}_{jj'}(s_j, t_{j'}) = E\left[\{Z_j(s_j) - \mu_j(s_j)\}\left\{Z_{j'}(t_{j'}) - \mu_{j'}(t_{j'})\right\}\right]$, where $s_j \in \mathcal{S}_j, t_{j'} \in \mathcal{S}_{j'}$ and $j, j' = 1, \ldots, d$.

By the multivariate version of Mercer's Theorem (Happ and Greven, 2018), $\mathcal{C}(\boldsymbol{s}, \boldsymbol{t})$ admits the spectral expansion $\mathcal{C}(\boldsymbol{s}, \boldsymbol{t}) = \sum_{k=1}^{\infty} \lambda_k \boldsymbol{\psi}_k(\boldsymbol{s}) \boldsymbol{\psi}_k(\boldsymbol{t})^\top$, where $\lambda_1 \geq \lambda_2 \geq \cdots > 0$

are ordered eigenvalues and $\boldsymbol{\psi}_k = (\psi_{k1}, \ldots, \psi_{kd})^\top$ are the corresponding eigenfunctions such that $\langle\langle \boldsymbol{\psi}_k, \boldsymbol{\psi}_{k'} \rangle\rangle = \int_{\mathcal{S}} \boldsymbol{\psi}_k(\boldsymbol{s})^\top \boldsymbol{\psi}_{k'}(\boldsymbol{s}) d\boldsymbol{s} = I(k = k')$. Based on the multivariate Karhunen-Loeve Theorem, we can obtain $\mathbf{Z}(\boldsymbol{s}) = \boldsymbol{\mu}(\boldsymbol{s}) + \sum_{k=1}^{\infty} \rho_k \boldsymbol{\psi}_k(\boldsymbol{s}), \boldsymbol{s} \in \mathcal{S}$, with zero mean random variables $\rho_k = \langle\langle \mathbf{Z} - \boldsymbol{\mu}, \boldsymbol{\psi}_k \rangle\rangle$ and $\operatorname{cov}(\rho_k, \rho_{k'}) = \lambda_k I(k = k')$. Since $\boldsymbol{\psi}_k(\boldsymbol{s}), k \in \mathbb{N}$ is a sequence of orthonormal basis, the functional coefficient $\boldsymbol{\alpha}(\boldsymbol{s}, \tau)$ can be expressed as $\boldsymbol{\alpha}(\boldsymbol{s}, \tau) = \sum_{k=1}^{\infty} a_{\tau k} \boldsymbol{\psi}_k(\boldsymbol{s})$. Suppose $\boldsymbol{\mu}(\boldsymbol{s}) = \mathbf{0}$ for simplicity. Then the functional component in model (2) can be calculated as $\langle\langle \boldsymbol{\alpha}(\boldsymbol{s}, \tau), \mathbf{Z}(\boldsymbol{s}) \rangle\rangle = \langle\langle \sum_{j=1}^{\infty} a_{\tau j} \boldsymbol{\psi}_j(\boldsymbol{s}), \sum_{k=1}^{\infty} \rho_k \boldsymbol{\psi}_k(\boldsymbol{s}) \rangle\rangle = \sum_{k=1}^{\infty} a_{\tau k} \rho_k$.

Although the number of principal components is infinite, it is commonly assumed that the relevant information regarding the response variable is primarily captured by the first $m = m_n$ FPC scores. The truncated number satisfies $1 \leq m_n \leq n - 1$ and $m_n \to \infty$ as $n \to \infty$. Then we can obtain $\langle\langle \boldsymbol{\alpha}(\boldsymbol{s}, \tau), \mathbf{Z}(\boldsymbol{s}) \rangle\rangle \approx \sum_{k=1}^{m} a_{\tau k} \rho_k := \boldsymbol{a}(\tau)^\top \boldsymbol{\rho}$, where $\boldsymbol{a}(\tau) = (a_{\tau 1}, \ldots, a_{\tau m})^\top, \boldsymbol{\rho} = (\rho_1, \ldots, \rho_m)^\top$. The FPC scores $\{\rho_k\}_{k=1}^{m}$ are unknown and need to be estimated.

In practice, the functional predictors $z_{ij}(s_j)$ may be observed on discrete grid points $\{s_{ijl} \in \mathcal{S}_j, 1 \leq l \leq L_{ij}\}$ with measurement errors. This paper mainly considers the dense observations for the functional predictors to construct consistent theory properties. In this case, many linear smoothers, such as local linear regression technique (Fan et al., 1996), local polynomial and regression splines etc., could be used to recover the true functions $z_{ij}(s_j)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, d$, leading to smoothed estimates $\hat{\boldsymbol{z}}_i(\boldsymbol{s}) = (\hat{z}_{i1}(s_1), \ldots, \hat{z}_{id}(s_d))^\top$.

Then, the mean and covariance of $\mathbf{Z}(\boldsymbol{s})$ could be estimated by

$$\begin{aligned} \widehat{\boldsymbol{\mu}}(\boldsymbol{s}) &= n^{-1} \sum_{i=1}^{n} \hat{\boldsymbol{z}}_i(\boldsymbol{s}), \\ \widehat{\mathcal{C}}(\boldsymbol{s}, \boldsymbol{t}) &= n^{-1} \sum_{i=1}^{n} \{\hat{\boldsymbol{z}}_i(\boldsymbol{s}) - \widehat{\boldsymbol{\mu}}(\boldsymbol{s})\} \{\hat{\boldsymbol{z}}_i(\boldsymbol{t}) - \widehat{\boldsymbol{\mu}}(\boldsymbol{t})\}^\top . \end{aligned} \tag{9}$$

The spectral decomposition of $\widehat{\mathcal{C}}(\boldsymbol{s}, \boldsymbol{t})$ is $\widehat{\mathcal{C}}(\boldsymbol{s}, \boldsymbol{t}) = \sum_{k=1}^{\infty} \widehat{\lambda}_k \widehat{\boldsymbol{\psi}}_k(\boldsymbol{s}) \widehat{\boldsymbol{\psi}}_k(\boldsymbol{t})^\top$, where $\{\widehat{\lambda}_k, k \geq 1\}$ are the sample eigenvalues in decreasing order and $\{\widehat{\boldsymbol{\psi}}_k(\boldsymbol{s}), k \geq 1\}$ are the corresponding eigenfunctions. The truncated mFPC score estimates are $\hat{\boldsymbol{\rho}}_i = (\widehat{\rho}_{i1}, \ldots, \widehat{\rho}_{im})^\top$ with $\widehat{\rho}_{ik} = \langle\langle \hat{\boldsymbol{z}}_i - \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\psi}}_k \rangle\rangle, k = 1, \ldots, m$. As stated in Li et al. (2010), the approximation error incurred by linear smoother is negligible in $\hat{\boldsymbol{\rho}}_i$ with dense functional observations. The choice of the truncation parameter m will be discussed in Section 4.3.

4.2. Estimation of Censored Transformation Model

Plugging the first m estimated mFPC scores into model (4), we obtain the transformed quantile regression model with $p + m$ predictors as follows:

$$Q_T(\tau \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma) = h^{-1}_{\gamma(\tau)} \left(\boldsymbol{\beta}(\tau)^\top \mathbf{X}_i + \boldsymbol{a}(\tau)^\top \hat{\boldsymbol{\rho}}_i \right), \tau \in (0, 1). \tag{10}$$

Note that, since our response variable is a time-to-event outcome, we can assume that $Q_T(0 \mid \mathbf{X}, \mathbf{Z}; \boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \gamma_0) = 0$. Based on the grid approximation method, for each τ_j

and any $\gamma(\tau_j), j = 1, \ldots, L$, we subsequently obtain $\{\hat{\boldsymbol{\beta}}(\gamma, \tau_j), \hat{\boldsymbol{a}}(\gamma, \tau_j)\}$ by solving the following monotone estimating equation:

$$\begin{aligned}\mathbf{S}_n(\gamma, \boldsymbol{\beta}, \boldsymbol{a}, \tau_j) = n^{-1} \sum_{i=1}^{n} (\mathbf{X}_i^\top, \hat{\boldsymbol{\rho}}_i^\top) \Big[N_i \left\{ Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma) \right\} \\ - \sum_{k=0}^{j-1} I \left\{ Y_i \geq Q_T(\tau_k \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\gamma}) \right\} \left\{ H\left(\tau_{k+1}\right) - H\left(\tau_k\right) \right\} \Big] = 0.\end{aligned} \tag{11}$$

Similar to Chu et al. (2021), solving (11) is transformed to minimizing a convex objective function of $\{\boldsymbol{\beta}\left(\tau_j\right), \boldsymbol{a}\left(\tau_j\right)\}$ for a given γ. That is,

$$\begin{aligned}\{\hat{\boldsymbol{\beta}}\left(\gamma, \tau_j\right), \hat{\boldsymbol{a}}\left(\gamma, \tau_j\right)\} &= \underset{\boldsymbol{\beta}(\tau_j), \boldsymbol{a}(\tau_j)}{\arg\min}\ \hat{L}(\gamma, \boldsymbol{\beta}, \boldsymbol{a}, \tau_j) \\ &= \underset{\boldsymbol{\beta}(\tau_j), \boldsymbol{a}(\tau_j)}{\arg\min}\ - \sum_{i=1}^{n} \left[h_{\gamma(\tau_j)}\left(Y_i\right) - h_{\gamma(\tau_j)}\left(Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma)\right) \right] \\ &\quad \times \Big[N_i \left\{ Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma) \right\} \\ &\quad - \sum_{k=0}^{j-1} I[Y_i \geq Q_T(\tau_k \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\gamma})] \left\{ H(\tau_{k+1}) - H(\tau_k) \right\} \Big].\end{aligned} \tag{12}$$

To estimate γ, this paper utilizes the standard grid search method. For a series of grid points in a given search interval of γ, we aim to find the minimizer $\hat{\gamma}$ of the following problem:

$$\begin{aligned}\hat{\gamma}\left(\tau_j\right) =& \underset{\gamma(\tau_j)}{\arg\min} \hat{R}(\gamma, \hat{\boldsymbol{\beta}}(\gamma), \hat{\boldsymbol{a}}(\gamma), \tau_j) \\ =& \underset{\gamma(\tau_j)}{\arg\min} - \sum_{i=1}^{n} \Big[Y_i - Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \gamma) \Big] \\ &\times \Big[N_i \Big\{ Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \gamma) \Big\} \\ &- \sum_{k=0}^{j-1} I[Y_i \geq Q_T(\tau_k \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\gamma})] \left\{ H(\tau_{k+1}) - H(\tau_k) \right\} \Big].\end{aligned} \tag{13}$$

The final estimates of $\{\boldsymbol{\beta}(\tau_j), \boldsymbol{a}(\tau_j)\}$ are obtained by substituting the optimal $\hat{\gamma}(\tau_j)$ into (12), i.e., $\hat{\boldsymbol{\beta}}(\tau_j) = \hat{\boldsymbol{\beta}}(\hat{\gamma}, \tau_j)$ and $\hat{\boldsymbol{a}}(\tau_j) = \hat{\boldsymbol{a}}(\hat{\gamma}, \tau_j)$. This two-stage estimation procedure has been previously adopted by Buchinsky (1995). The above minimization problems can be easily solved in the R *quantreg* package (Koenker, 2012) and R function *optim*.

4.3. Selection of the Cut-off Level m

The number of mFPC scores plays an important role in the regression. There are many criteria in the literature to choose the tuning parameter (Tang and Cheng, 2014; Kong

et al., 2018). Inspired by Jiang et al. (2020), this paper suggests the following generalized approximate cross-validation (GACV) objective function:

$$\begin{aligned}
&\mathrm{GACV}\left(\tau_j, m, \boldsymbol{\beta}, \boldsymbol{a}, \gamma\right) \\
&\quad = -\sum_{i=1}^{n}\left[h_{\gamma(\tau_j)}\left(Y_i\right) - \left\{\boldsymbol{\beta}(\tau_j)^{\top}\mathbf{X}_i + \boldsymbol{a}(\tau_j)^{\top}\hat{\boldsymbol{\rho}}_i\right\}\right] \\
&\qquad \times \left(N_i\left[h^{-1}_{\gamma(\tau_j)}\left\{\boldsymbol{\beta}\left(\tau_j\right)^{\top}\mathbf{X}_i + \boldsymbol{a}(\tau_j)^{\top}\hat{\boldsymbol{\rho}}_i\right\}\right]\right. \\
&\qquad - \sum_{k=0}^{j-1} I\left[Y_i \geq h^{-1}_{\hat{\gamma}(\tau_k)}\left\{\widehat{\boldsymbol{\beta}}\left(\tau_k\right)^{\top}\mathbf{X}_i + \hat{\boldsymbol{a}}(\tau_k)^{\top}\hat{\boldsymbol{\rho}}_i\right\}\right] \\
&\qquad \left.\times \left\{H\left(\tau_{k+1}\right) - H\left(\tau_k\right)\right\}\right) / \left\{n - C_n\left(m + p\right)\right\},
\end{aligned} \tag{14}$$

where $C_n = O_p\left[\{\log(n)\}^{1/2}\right]$. The cut-off level m is selected by minimizing the averaged GACV

$$\mathrm{AGACV} = \sum_{j=1}^{L} \mathrm{GACV}\left(\tau_j, m, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\gamma}\right) / L,$$

where L is the number of quantile grid points in $\mathcal{T}_L$ mentioned in Section 3.

One practical concern in the estimation is the choice of L or the grid size of $\mathcal{T}_L$. From the proof of weak convergence in Section 5, it is known that an order of $o(n^{-1/2})$ for $\|\mathcal{T}_L\|$ is enough to ensure convergence. In this paper, we utilize an equally spaced grid with a grid size of 0.05 in the simulation study, which yields satisfactory results for parameter estimation. Alternatively, as suggested by Peng and Huang (2008), a finer grid size of $\|\mathcal{T}_L\| = 0.01$ can be employed for more detailed analysis practically, albeit with slightly higher computational costs. The detailed estimation procedure for the proposed method is summarized in Algorithm 1.

5. Asymptotics

In this section, we establish the estimation consistency in Theorem 1, the weak convergence in Theorem 2, and the convergence of the functional estimator in Corollary 1.

First, define $F(t|\mathbf{X},\mathbf{Z}) = \mathrm{P}(Y \leq t|\mathbf{X},\mathbf{Z})$, $\bar{F}(t|\mathbf{X},\mathbf{Z}) = 1 - F(t|\mathbf{X},\mathbf{Z})$, $\widetilde{F}(t|\mathbf{X},\mathbf{Z}) = \mathrm{P}(Y \leq t, \delta = 1|\mathbf{X},\mathbf{Z})$, $f(t|\mathbf{X},\mathbf{Z}) = dF(t|\mathbf{X},\mathbf{Z})/dt$, $\bar{f}(t|\mathbf{X},\mathbf{Z}) = d\bar{F}(t|\mathbf{X},\mathbf{Z})/dt$, and $\widetilde{f}(t|\mathbf{X},\mathbf{Z}) = d\widetilde{F}(t|\mathbf{X},\mathbf{Z})/dt$. The detailed conditions and theoretical derivations are presented in Sections 1-2 of the Supplementary materials.

Theorem 1. Assuming that conditions A1-A4 and B1-B5 hold, if $\lim_{n\to\infty}\|\mathcal{T}_L\| = 0$, then we have $\sup_{\tau\in[v,\tau_U]}|\hat{\gamma}(\tau) - \gamma_0(\tau)| \to_p 0$, $\sup_{\tau\in[v,\tau_U]}\left\|\hat{\boldsymbol{\beta}}(\hat{\gamma},\tau) - \boldsymbol{\beta}_0(\tau)\right\| \to_p 0$, and $\sup_{\tau\in[v,\tau_U]}\|\hat{\boldsymbol{a}}(\hat{\gamma},\tau) - \boldsymbol{a}_0(\tau)\| \to_p 0$, where $0 < v < \tau_U$.

After obtaining the above estimators, the estimation for $\boldsymbol{\alpha}_0(\boldsymbol{s},\tau)$ can be calculated as $\hat{\boldsymbol{\alpha}}(\boldsymbol{s},\tau) = \sum_{k=1}^{m}\hat{a}_{\tau k}\hat{\boldsymbol{\psi}}_k$. The convergency of this functional coefficient estimate is shown in the following Corollary 1.

Corollary 1. Assuming that conditions A1-A4 and B1-B5 hold, if $\lim_{n\to\infty}\|\mathcal{T}_L\| = 0$, then $\sup_{\tau\in[v,\tau_U]}\|\|\hat{\boldsymbol{\alpha}}(\boldsymbol{s},\tau) - \boldsymbol{\alpha}_0(\boldsymbol{s},\tau)\|\| \to_p 0$, where $0 < v < \tau_U$.

Algorithm 1 Estimation procedure for the proposed method

Input: The response variable Y_i, the censoring indicator δ_i, the scalar and functional covariates $\{\mathbf{X}_i, \mathbf{Z}_i(\boldsymbol{s})\}, i = 1, \ldots, n$, the set of cut-off levels $\mathcal{M}$, the set of quantile values $\mathcal{T}_L$, the grid search interval $\mathcal{G}$.

Output: The coefficient estimates for $\boldsymbol{\beta}(\tau), \boldsymbol{a}(\tau), \gamma(\tau)$ and the optimal m.

for each $m \in \mathcal{M}$ **do**
  Estimate $\mathcal{C}(s,t), \lambda_k, \boldsymbol{\psi}_k(\boldsymbol{s})$ by Equation (9);
  Compute $\hat{\boldsymbol{\rho}}_i$ with its elements $\widehat{\rho}_{ik} = \langle\langle \hat{\boldsymbol{z}}_i - \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\psi}}_k \rangle\rangle, k = 1, \ldots, m$;
  for each $\tau_j \in \mathcal{T}_L$ **do**
    for each $\gamma \in \mathcal{G}$ **do**
      Find the solution $\hat{\boldsymbol{\beta}}(\gamma), \hat{\boldsymbol{a}}(\gamma)$ for the minimization problem (12);
      Compute $\hat{R}(\gamma, \hat{\boldsymbol{\beta}}(\gamma), \hat{\boldsymbol{a}}(\gamma), \tau_j)$;
    end for
    Find the minimizer $\hat{\gamma}(\tau_j) = \arg\min \hat{R}(\gamma, \hat{\boldsymbol{\beta}}(\gamma), \hat{\boldsymbol{a}}(\gamma), \tau_j)$;
    Find the minimizer $\{\hat{\boldsymbol{\beta}}(\tau_j), \hat{\boldsymbol{a}}(\tau_j)\} = \arg\min \hat{L}(\hat{\gamma}, \boldsymbol{\beta}, \boldsymbol{a}, \tau_j)$;
  end for
  Compute the criteria $\mathrm{AGACV} = \sum_{j=1}^{L} \mathrm{GACV}\left(\tau_j, m, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{a}}, \hat{\gamma}\right)/L$;
end for
Find the optimal m to minimize AGACV criteria.
return $\{\hat{\boldsymbol{\beta}}(\tau_j), \hat{\boldsymbol{a}}(\tau_j), \hat{\gamma}(\tau_j)\}_{j=1}^{L}, m$.

Theorem 2. Assuming that conditions A1-A4 and B1-B5 hold, if $\lim_{n\to\infty} n^{1/2} \|\mathcal{T}_L\| = 0$, then we have $n^{1/2}\left[\hat{\gamma}(\tau) - \gamma_0(\tau), \{\hat{\boldsymbol{\beta}}(\hat{\gamma}, \tau) - \boldsymbol{\beta}_0(\tau)\}^\top, \{\hat{\boldsymbol{a}}(\hat{\gamma}, \tau) - \boldsymbol{a}_0(\tau)\}^\top\right]^\top$ converges weakly to a Gaussian process for $\tau \in [v, \tau_U]$, where $0 < v < \tau_U$.

The proof of Theorem 2 reveals that the covariance matrix of the limiting process of the estimators consists of unknown conditional density functions $f(t|\mathbf{X}, \mathbf{Z})$ and $\tilde{f}(t|\mathbf{X}, \mathbf{Z})$. Estimating these functions typically requires numerical approximation methods, which may lack stability with small or moderate sample sizes (Peng and Huang, 2008). To address this issue, we propose to employ a simple resampling method for variance estimation by generalizing the minimand perturbing technique of Jin et al. (2001) and Qian and Peng (2010) to ensure computational stability.

The resampling procedure is described as follows. For a dataset with sample size n, we fix the data values and generate n i.i.d variates $\zeta_1, \ldots, \zeta_n$ from a known nonnegative distribution with mean 1 and variance 1, such as the standard exponential distribution. Then, $\gamma^*(\tau_j)$, $\boldsymbol{\beta}^*(\tau_j)$ and $\boldsymbol{a}^*(\tau_j)$ are sequentially obtained for each $j \in \{1, \ldots, L\}$ by

solving the following objective functions perturbed by $\{\zeta_1, \ldots, \zeta_n\}$:

$$\begin{aligned}\{\boldsymbol{\beta}^*(\gamma, \tau_j), \boldsymbol{a}^*(\gamma, \tau_j)\} &= \underset{\boldsymbol{\beta}(\tau_j), \boldsymbol{a}(\tau_j)}{\arg\min}\, L^*(\gamma, \boldsymbol{\beta}, \boldsymbol{a}, \tau_j) \\ &= \underset{\boldsymbol{\beta}(\tau_j), \boldsymbol{a}(\tau_j)}{\arg\min} -\sum_{i=1}^{n} \zeta_i \left[h_{\gamma(\tau_j)}(Y_i) - h_{\gamma(\tau_j)}(Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma))\right] \\ &\quad \times \Bigg[N_i \{Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}, \boldsymbol{a}, \gamma)\} \\ &\quad - \sum_{k=0}^{j-1} I[Y_i \geq Q_T(\tau_k \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}^*, \boldsymbol{a}^*, \gamma^*)] \{H(\tau_{k+1}) - H(\tau_k)\}\Bigg],\end{aligned}$$

$$\begin{aligned}\gamma^*(\tau_j) =& \underset{\gamma(\tau_j)}{\arg\min} R^*(\gamma, \boldsymbol{\beta}^*(\gamma), \boldsymbol{a}^*(\gamma), \tau_j) \\ =& \underset{\gamma(\tau_j)}{\arg\min} -\sum_{i=1}^{n} \zeta_i [Y_i - Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}^*, \boldsymbol{a}^*, \gamma)] \\ &\times [N_i \{Q_T(\tau_j \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}^*, \boldsymbol{a}^*, \gamma)\} \\ &- \sum_{k=0}^{j-1} I[Y_i \geq Q_T(\tau_k \mid \mathbf{X}_i, \hat{\boldsymbol{\rho}}_i; \boldsymbol{\beta}^*, \boldsymbol{a}^*, \gamma^*)] \{H(\tau_{k+1}) - H(\tau_k)\}\Bigg],\end{aligned}$$

where $Q_T(0 \mid \mathbf{X}, \hat{\boldsymbol{\rho}}; \boldsymbol{\beta}^*, \boldsymbol{a}^*, \gamma^*) = 0$, and $\boldsymbol{\beta}^*(\tau_j) = \boldsymbol{\beta}^*(\gamma^*, \tau_j)$, $\boldsymbol{a}^*(\tau_j) = \boldsymbol{a}^*(\gamma^*, \tau_j)$. The resampling estimators are also defined as right-continuous step functions that jump at $\tau_j, j = 1, \ldots, L$ (Chu et al., 2021). With the data fixed at the observed values, the above procedure is repeated for B times to obtain a large number of realizations, denoted by $\{\gamma_r^*(\tau), \boldsymbol{\beta}_r^*(\tau), \boldsymbol{a}_r^*(\tau)\}_{r=1}^B$.

In Section 3.3 of the Supplementary material, we show that the conditional distribution of $n^{1/2}[\gamma^*(\tau) - \hat{\gamma}(\tau)]$ given the observed data is asymptotically the same as the unconditional distribution of $n^{1/2}[\hat{\gamma}(\tau) - \gamma_0(\tau)]$ as a process of $\tau \in [v, \tau_U]$, where $0 < v < \tau_U$. Thus, the variance of $\hat{\gamma}(\tau)$ can be estimated by the sample variance of $\{\gamma_r^*(\tau)\}_{r=1}^B$. A confidence interval for $\gamma_0(\tau)$ can be constructed using a normal approximation. Similar properties and inferences can also be obtained for $\{\hat{\boldsymbol{\beta}}(\tau), \hat{\boldsymbol{a}}(\tau)\}$. Justification for the proposed resampling approach is provided in Section 3.3 of the Supplementary material.

6. Simulation Study

In this section, we evaluate the finite-sample performance of the proposed multivariate functional censored quantile regression model (MFCQR) through Monte Carlo simulations. In the first setting, we consider the Box-Cox transformation model with bivariate correlated functional data on a one-dimensional interval and compare the proposed method with several existing estimation methods. In the second setting, logarithmic transformation is used to examine the adaptability of the proposed model, where curves and images are simultaneously incorporated as functional covariates. Results are obtained based on 500 simulated datasets of sample size $n = 200$ and $n = 500$.

6.1. Functional Predictors on One-dimensional Intervals

In this setting, the event times are generated from the following Box-Cox transformed linear quantile regression model with heteroscedastic errors:

$$\begin{aligned} Q_T(\tau \mid \mathbf{X}, \mathbf{Z}) &= h_{\gamma_0(\tau)}^{-1}\left(\boldsymbol{\beta}_0(\tau)^\top \mathbf{X} + \langle\langle \boldsymbol{\alpha}_0(\boldsymbol{s}, \tau), \mathbf{Z}(\boldsymbol{s})\rangle\rangle\right) \\ &= h_{\gamma_0(\tau)}^{-1}\left(b_0 + b_1 X_1 + b_2 X_2 + \sum_{j=1}^{2} \int_0^1 \alpha_j(s) Z_j(s) ds + (1 + X_1) Q_\epsilon(\tau | \mathbf{X}, \mathbf{Z})\right), \end{aligned} \tag{15}$$

where $X_1 \sim N(0, 0.5^2)$, $X_2 \sim \text{Unif}[0, 1]$ and $\epsilon \sim N(0, 0.25^2)$. In this heteroscedastic model, the quantile regression coefficients of $\mathbf{X} = (1, X_1, X_2)^\top$ are $\boldsymbol{\beta}(\tau) = (\beta_0(\tau), \beta_1(\tau), \beta_2(\tau))^\top = (b_0 + Q_\epsilon(\tau), b_1 + Q_\epsilon(\tau), b_2)^\top$ with $b_0 = b_1 = b_2 = 1$, where the first two entries of $\boldsymbol{\beta}(\tau)$ are both τ-dependent.

The multivariate functional predictor is $\mathbf{Z}(\boldsymbol{s}) = \{Z_1(s), Z_2(s)\}^\top$ with sample observations $z_{i1}(s) = \sum_{k=1}^K \xi_{ik}^{(1)} \phi_k(s)$ and $z_{i2}(s) = \sum_{k=1}^K \xi_{ik}^{(2)} \phi_k(s)$, where $\xi_{ik}^{(1)} \sim N(0, \nu_{2k-1})$, $\xi_{ik}^{(2)} \sim N(0, \nu_{2k})$, $\nu_k = k^{-2}$, $\text{corr}\left(\xi_{ik}^{(j)}, \xi_{ik'}^{(j)}\right) = 0 \quad \text{for } k' \neq k$, $K = 20$, $\phi_1 = 1$ and $\phi_k(s) = \sqrt{2}\cos\{(k-1)\pi s\}, s \in \mathcal{S} = [0, 1]$ for $k > 1$. Following Wong et al. (2019), we further set $\text{corr}\left(\xi_{ik}^{(1)}, \xi_{ik'}^{(2)}\right) = \varrho$ for $k' = k$ and 0 otherwise, where $\varrho \in (0, 1)$ is a cross-correlation parameter controlling the strength of dependence between functional predictors. The coefficient functions are $\alpha_j(s) = \sum_{k=1}^K a_{jk}\phi_k(s)$ with $a_{j1} = 1, a_{j2} = 0.8, a_{j3} = 0.6, a_{j4} = 0.4$ and $a_{jk} = 8(k-2)^{-3}$ for $k = 5, \ldots, K$.

Similar to Chu et al. (2021), the transformation parameter $\gamma_0(\tau)$ is set to be 1 for $\tau \leq 0.4$ and 0.5 for $\tau > 0.4$ with $h_{\gamma_0(\tau)}^{-1}(x) = (\gamma_0(\tau)x + 1)^{1/\gamma_0(\tau)}$. The censoring time C is generated from a uniform distribution $\text{Unif}(0, V)$ with $V = \exp(c_0 + X_1 + X_2)$, where c_0 is taken to be 2.2 and 1.5 to yield censoring rates of 20% and 40%, respectively. The discrete observations of each functional predictor are obtained on 100 equally spaced grid points on $\mathcal{S} = [0, 1]$: $w_{ijl} = z_{ij}(s_{ijl}) + e_{ijl}, e_{ijl} \sim N(0, 0.25^2)$ for $i = 1, \ldots, n; j = 1, 2$ and $l = 1, \ldots, 100$. To obtain the standard errors of the estimators, we set the resampling size $B = 300$.

The proposed method (MFCQR) is also compared with some alternatives, namely the transformed quantile model with separate FPCA (SFPCA, Kong et al., 2018), the functional censored model with log transformation and B-splines approximation (Log-spline, Jiang et al., 2020), the kernel-based local Kaplan-Meier estimator (KM) proposed by Leng and Tong (2014) and the parametric modeling of quantile regression coefficients combined with MFPCA (FB, Frumento and Bottai, 2017). For comparison, the empirical bias (Bias) and the sample standard deviation (SD) are reported for each method. We also present the average of the estimated standard errors (SE) based on the resampling method and the coverage probabilities (CP) of the 95% confidence intervals based on a normal approximation for the proposed method.

Table 2 presents the estimation results with $n = 200$, considering 40% censoring rate at different quantile levels. For right-censored data, quantile functions in the upper tail may not be identifiable due to the censoring of event time information (Wu et al., 2015; Fei et al., 2023). In this case, we set the quantile interval to be (0, 0.7], which is wide enough to cover different quantile levels and, in the meantime, ensures the quantile parameters to be

estimable in the presence of censoring. The determination of the upper bound for estimable quantiles is discussed in Section 8. Compared with the alternatives, the proposed MFCQR provides smaller biases and standard errors, which are not affected by the high correlation of the functional predictors. Moreover, the resampling-based standard errors agree with the empirical standard deviations quite well and the coverage probabilities are close to the 95% nominal level. By contrast, both SFPCA and Log-spline methods incorporate new variables with strong correlations, resulting in poor performace under $\varrho = 0.6$. In particular, Log-spline employs the log-transformation directly, regardless of the dynamic parameter γ, and thus behaves badly even with $\varrho = 0.3$ in this setting. With slightly larger standard deviations, KM method performs well only in lower quantile cases, probably due to the unstability of the locally weighted kernel-based estimator for higher quantiles (Chu et al., 2021). By imposing smoothness conditions on the coefficient functions, FB method performs well in the situation where the basis functions are consistent with those in the true model (Frumento and Bottai, 2017). However, in our simulation settings, there is a dynamic transformation on the response variable, where the transformation parameter γ varies dynamically with the quantile values, making it difficult to select a proper set of basis functions for the parametric modeling of quantile regression coefficients.

The results for the case of 20% censoring rates are provided in Table 1 of the Supplementary material to save space. As shown in Table 2, the model performance experiences a decline to a certain extent as the censoring rate increases as expected. Estimation results for $n = 500$ are included in Tables 2-3 of the Supplementary material.

To evaluate the performance of the functional estimator $\hat{\alpha}_j(s)$, the integrated squared error (ISE) of $\hat{\alpha}_j(s)$, defined as $\text{ISE}\left(\widehat{\alpha}_j\right) = \int_0^1 \left\{\hat{\alpha}_j(s) - \alpha_j(s)\right\}^2 ds$, is used for comparison. Table 3 presents the averaged ISEs of $\hat{\alpha}_1(s)$ for $n = 200$. The results are similar for $\hat{\alpha}_2(s)$ and thus are omitted. Clearly, the proposed method demonstrates robustness across different settings and yields more accurate estimates compared to alternative methods.

6.2. Functional Covariates Including Curves and Images

In this setting, event times are generated from the log-transformed quantile regression model, that is, the true transformation parameter $\gamma_0(\tau) = 0$ for all $\tau \in (0, 1)$ in the proposed Box-Cox transformation method:

$$\begin{aligned} Q_T(\tau \mid \mathbf{X}, \mathbf{Z}) &= \exp\left(\boldsymbol{\beta}_0(\tau)^\top \mathbf{X} + \langle\langle \boldsymbol{\alpha}_0(\boldsymbol{s}, \tau), \mathbf{Z}(\boldsymbol{s})\rangle\rangle\right) \\ &= \exp\left(b_0 + b_1 X_1 + b_2 X_2 + \sum_{j=1}^{2} \int_{\mathcal{S}_j} \alpha_j(s_j) Z_j(s_j) ds_j \right. \\ &\qquad \left. + \left(1 + \int_{\mathcal{S}_1} Z_1(s_1) ds_1\right) Q_\epsilon(\tau|\mathbf{X}, \mathbf{Z})\right). \end{aligned} \tag{16}$$

The quantile regression coefficient of $\mathbf{X} = (1, X_1, X_2)^\top$ is $\boldsymbol{\beta}_0(\tau) = (b_0 + Q_\epsilon(\tau), b_1, b_2)^\top$ with $b_0 = b_1 = b_2 = 1$, $X_1 \sim N(0, 0.5^2)$, $X_2 \sim \text{Unif}[0, 1]$ and $\epsilon \sim N(0, 0.25^2)$. Notably, the bivariate functional data $\mathbf{Z}(\boldsymbol{s}) = (Z_1(s_1), Z_2(s_2))^\top$ contains functions and images simultaneously. To be specific, $\boldsymbol{z}_i(\boldsymbol{s}) = \sum_{m=1}^{25} \rho_{im} \boldsymbol{\psi}_m(\boldsymbol{s})$, where the principal component $\boldsymbol{\psi}_m(\boldsymbol{s}) = (\psi_{1m}(s_1), \psi_{2m}(s_2))^\top$ has the same structure as $\boldsymbol{z}_i(\boldsymbol{s})$. As in Happ and

Table 2. The estimation results for the scalar parameters with $n = 200$, in the setting of 40% censoring rate.

			$\varrho = 0.3$				$\varrho = 0.6$			
τ	Method		$\hat{\gamma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.1	MFCQR	Bias	-0.076	0.037	-0.035	0.027	0.072	-0.039	-0.032	-0.029
		SD	0.345	0.223	0.215	0.212	0.341	0.220	0.213	0.218
		SE	0.367	0.241	0.207	0.203	0.324	0.229	0.201	0.224
		CP	0.962	0.957	0.944	0.945	0.947	0.960	0.942	0.964
	SFPCA	Bias	0.093	0.044	-0.034	-0.035	0.374	-0.265	0.323	-0.306
		SD	0.357	0.228	0.213	0.217	0.379	0.286	0.299	0.267
	Log-spline	Bias	-	-0.487	0.514	0.499	-	-0.519	-0.588	0.549
		SD	-	0.165	0.178	0.192	-	0.248	0.231	0.285
	KM	Bias	-0.074	-0.041	0.036	0.038	0.078	-0.037	-0.044	0.039
		SD	0.344	0.229	0.225	0.230	0.375	0.241	0.234	0.245
	FB	Bias	-	0.361	-0.359	-0.383	-	0.365	0.353	-0.377
		SD	-	0.227	0.221	0.218	-	0.224	0.216	0.229
0.3	MFCQR	Bias	0.063	0.035	-0.032	-0.024	0.068	-0.034	-0.030	0.027
		SD	0.313	0.206	0.201	0.199	0.315	0.213	0.205	0.216
		SE	0.335	0.218	0.216	0.187	0.303	0.234	0.212	0.205
		CP	0.961	0.954	0.958	0.943	0.944	0.957	0.955	0.942
	SFPCA	Bias	0.091	-0.038	-0.031	0.029	-0.404	-0.253	0.305	0.309
		SD	0.346	0.224	0.193	0.205	0.371	0.278	0.284	0.265
	Log-spline	Bias	-	0.458	0.476	-0.493	-	0.504	-0.517	-0.536
		SD	-	0.162	0.173	0.190	-	0.243	0.227	0.267
	KM	Bias	-0.072	-0.035	0.031	0.033	0.073	-0.043	0.036	0.035
		SD	0.324	0.226	0.218	0.221	0.372	0.232	0.227	0.235
	FB	Bias	-	-0.343	0.337	-0.355	-	0.349	-0.345	0.351
		SD	-	0.211	0.204	0.206	-	0.215	0.209	0.222
0.5	MFCQR	Bias	0.060	-0.033	0.026	-0.021	-0.059	0.029	-0.027	0.024
		SD	0.304	0.198	0.191	0.185	0.312	0.209	0.193	0.207
		SE	0.319	0.193	0.205	0.173	0.301	0.217	0.182	0.194
		CP	0.957	0.948	0.955	0.944	0.946	0.954	0.945	0.943
	SFPCA	Bias	-0.088	-0.035	0.032	-0.025	0.407	-0.259	0.327	0.319
		SD	0.314	0.219	0.192	0.198	0.376	0.296	0.281	0.268
	Log-spline	Bias	-	0.451	-0.467	0.484	-	0.478	0.514	-0.515
		SD	-	0.160	0.172	0.185	-	0.245	0.223	0.261
	KM	Bias	-0.069	0.038	-0.029	0.027	-0.071	0.041	0.038	-0.036
		SD	0.311	0.215	0.212	0.214	0.367	0.224	0.213	0.231
	FB	Bias	-	0.322	-0.315	-0.326	-	0.331	0.329	0.338
		SD	-	0.205	0.197	0.193	-	0.218	0.202	0.214
0.7	MFCQR	Bias	0.074	-0.038	0.031	-0.032	-0.078	0.034	0.035	0.037
		SD	0.349	0.231	0.219	0.217	0.346	0.236	0.224	0.220
		SE	0.360	0.245	0.211	0.208	0.359	0.247	0.208	0.203
		CP	0.958	0.961	0.946	0.942	0.963	0.962	0.943	0.945
	SFPCA	Bias	-0.107	-0.042	0.047	-0.039	-0.418	-0.317	0.332	-0.348
		SD	0.364	0.248	0.235	0.246	0.388	0.301	0.296	0.277
	Log-spline	Bias	-	-0.474	0.528	0.533	-	0.495	0.558	-0.553
		SD	-	0.193	0.201	0.195	-	0.296	0.324	0.315
	KM	Bias	0.101	-0.051	-0.057	-0.043	0.106	-0.079	-0.083	0.079
		SD	0.398	0.245	0.242	0.249	0.407	0.275	0.258	0.262
	FB	Bias	-	-0.351	0.343	0.358	-	-0.354	0.347	-0.350
		SD	-	0.237	0.232	0.226	-	0.243	0.230	0.235

Table 3. The integrated squared errors (ISE) of $\hat{\alpha}_1(s)$ with $n = 200$ in the setting of 20% and 40% censoring rates.

τ	MFCQR	SFPCA	Log-spline	KM	FB	MFCQR	SFPCA	Log-spline	KM	FB
	$\varrho = 0.3$, censoring rate=20%					$\varrho = 0.6$, censoring rate=20%				
0.1	0.113	0.121	0.342	0.127	0.236	0.115	0.187	0.476	0.125	0.245
0.3	0.105	0.118	0.329	0.124	0.221	0.109	0.181	0.468	0.121	0.228
0.5	0.101	0.116	0.311	0.120	0.214	0.104	0.176	0.462	0.116	0.217
0.7	0.110	0.119	0.332	0.132	0.227	0.112	0.183	0.469	0.128	0.232
	$\varrho = 0.3$, censoring rate=40%					$\varrho = 0.6$, censoring rate=40%				
0.1	0.122	0.135	0.351	0.133	0.254	0.124	0.194	0.485	0.138	0.267
0.3	0.120	0.129	0.337	0.128	0.248	0.116	0.189	0.479	0.133	0.255
0.5	0.117	0.124	0.325	0.125	0.235	0.113	0.183	0.474	0.127	0.244
0.7	0.121	0.131	0.353	0.139	0.251	0.120	0.191	0.482	0.141	0.257

Greven (2018), the first function element $\psi_{1m}(s_1)$ is generated by Legendre polynomials on $\mathcal{S}_1 = [-1, 1]$ and the second image element $\psi_{2m}(s_2)$ is formed by tensor products of Fourier basis functions on $\mathcal{S}_2 = [0, 1] \times [0, 1]$. The scores ρ_{im} are independently generated from $N(0, \nu_m)$ with exponentially decreasing eigenvalues $\nu_m = \exp(-(m + 1)/2)$. The corresponding functional coefficient vector $\boldsymbol{\alpha}_0(\boldsymbol{s}, \tau) = (\alpha_1(s_1) + Q_\epsilon(\tau), \alpha_2(s_2))^\top$, where $\alpha_j(s_j) = \sum_{m=1}^{25} a_{jm}\psi_{jm}$ with $a_{j1} = 1, a_{j2} = 0.6, a_{j3} = -0.4, a_{j4} = -0.1, a_{j5} = 0.05$, $a_{jm} = 2(-1)^{m+1}m^{-3}, m \geq 6$.

The censoring time $C \sim \text{Unif}(0, c)$ with $c = \exp(c_0 + X_1 + X_2)$, where c_0 is taken to be 3 and 2 to yield censoring rates of 20% and 40%, respectively. The functional observations are discretized based on a grid of 200 equidistant points for the function elements $\boldsymbol{z}_{i1}$ and 100×100 equidistant points for the image elements $\boldsymbol{z}_{i2}$.

Table 4 shows the performance of the proposed method for the estimators $\hat{\gamma}$ and $\hat{\boldsymbol{\beta}}$ at $\tau = 0.25, 0.50, 0.75$ with $n = 200$, in the setting of 20% and 40% censoring rates. It is seen that the estimation results are satisfactory across different quantile levels, indicating the good adaptability of the proposed method to the logarithmic transformation model. The performance is further improved by larger sample size $n = 500$, which is displayed in Table 5 of the Supplementary material.

The estimation results for the functional coefficient vector $\boldsymbol{\alpha}_0(\boldsymbol{s}, \tau)$ are evaluated in Figures 4 and 5 with $n = 200$ under the censoring rates of 20% and 40%. Figure 4 displays the estimates and the empirical 95% confidence intervals for the first functional coefficient element $\alpha_1(s_1, \tau)$. The true curves are well recovered and the confidence intervals have stable performance under different settings. Figure 5 shows the estimates for the second image coefficient $\alpha_2(s_2, \tau)$, all of which have a nice approxiation to the true ones (the left panels). Though the estimation errors increase with larger censoring rate, the overall performation is satisfactory. The results with $n = 500$ is displayed in Figures 2-3 of the Supplementary material.

Table 4. The estimation results for the scalar parameters with $n = 200$ in the setting of 20% and 40% censoring rates.

		censoring rate=20%				censoring rate=40%			
τ		$\hat{\gamma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}$	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$
0.25	Bias	-0.023	0.015	-0.018	-0.013	0.036	-0.024	0.029	-0.027
	SD	0.182	0.131	0.135	0.137	0.212	0.165	0.157	0.161
	SE	0.171	0.146	0.127	0.149	0.201	0.178	0.160	0.167
	CP	0.942	0.959	0.945	0.956	0.944	0.958	0.949	0.953
0.50	Bias	-0.019	-0.013	0.015	-0.012	-0.028	-0.020	0.024	0.023
	SD	0.152	0.122	0.131	0.125	0.188	0.151	0.153	0.150
	SE	0.144	0.137	0.140	0.132	0.195	0.163	0.145	0.146
	CP	0.946	0.958	0.955	0.952	0.956	0.960	0.944	0.948
0.75	Bias	-0.021	0.017	0.013	0.014	0.032	-0.028	0.022	-0.025
	SD	0.176	0.139	0.135	0.130	0.207	0.173	0.167	0.162
	SE	0.185	0.131	0.148	0.137	0.189	0.179	0.164	0.171
	CP	0.962	0.945	0.960	0.956	0.943	0.954	0.948	0.955

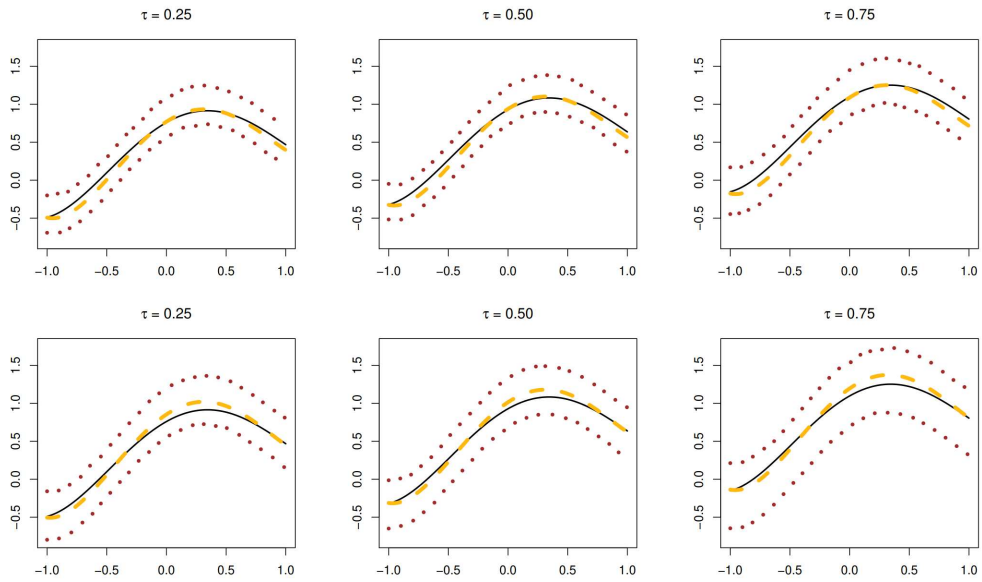

Fig. 4. Estimation results for the first coefficient function $\alpha_1(s_1, \tau)$ at quantiles $\tau = 0.25, 0.50, 0.75$ with $n = 200$ under censoring rates of 20% (top panels) and 40% (bottom panels), respectively. The black solid lines are the true coefficient functions, the yellow dashed lines represent the estimated functions and the brown dotted lines are the empirical confidence intervals.

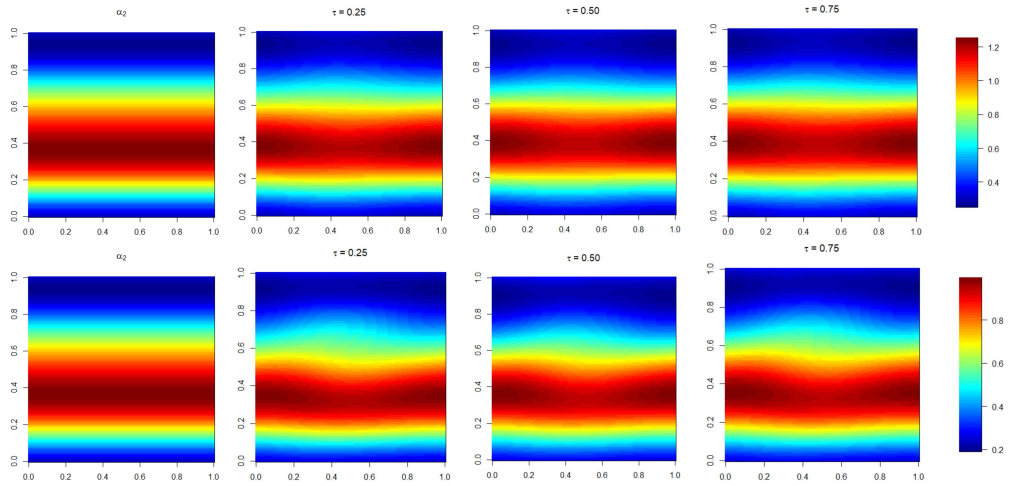

Fig. 5. Estimation results for the second coefficient function $\alpha_2(s_2, \tau)$ at quantiles $\tau = 0.25, 0.50, 0.75$ with $n = 200$ under censoring rates of 20% (top panels) and 40% (bottom panels), respectively.The left panels are the true image coefficients and the others are estimates under different quantiles.

7. Real Data Analysis

We applied the proposed Box-Cox transformed quantile regression model to the ADNI data, using the IGACV criterion in Section 4.3 to select the number of mFPCs. The paper finally selects $m = 4$ at which the AGACV reaches its minimum value. The resampling size is set as 500 to obtain the variance estimation. Table 5 presents the estimation results for the transformation parameter $\gamma(\tau)$ and the regression coefficients $\boldsymbol{\beta}(\tau)$ of 10 scalar covariates: (Gender, Handedness, Widowed, Divorced, Never married, Length of education, Retirement, Age, APOE-ϵ4 carrier, ADAS-cog Score) with 95% bootstrap confidence intervals at different quantiles. Since shorter conversion time represents faster disease progression, which is the main focus of researchers, we considered the estimation results at the lower levels of quantiles $\tau = 0.15, 0.30, 0.45$. It is seen that the transformation parameter estimates are significantly different from zero, indicating that it is necessary to employ data-driven model structures at different quantile levels. The 95% confidence intervals for $\beta_6, \beta_8, \beta_9, \beta_{10}$ do not contain 0 for all considered quantiles, showing that the corresponding covariates, Length of education, Age, APOE-ϵ4 carrirer and ADAS-cog score, are important predictors for AD. To be specific, higher education and older age at baseline are associated with longer conversion time from MCI to AD. Individuals carrying the APOE-ϵ4 allele have a higher propensity to transition from MCI to AD compared to non-carriers suggesting the presence of ϵ4 allele in APOE may increase the risk of developing AD and attaining a definitive AD diagnosis. These findings also coincide with the outcomes reported in Corder et al. (1993) and Da et al. (2014). In addition, patients with larger ADAS-Cog scores are expected to experience a more accelerated transition to AD diagnosis.

Table 5. Estimates and 95% bootstrap confidence intervals (CI) of the scalar parameters at different quantile values for the ADNI data.

τ		γ	β_1	β_2	β_3	β_4
	estimate	0.345	0.264	0.119	-0.146	0.290
	95% CI	(0.237,0.426)	(0.087,0.415)	(-0.283,0.511)	(-0.414,0.149)	(-0.268,0.883)
		β_5	β_6	β_7	β_8	β_9
0.15	estimate	0.201	0.108	-0.355	0.042	-0.371
	95% CI	(-0.052,0.414)	(0.045,0.162)	(-0.713,0.039)	(0.011,0.074)	(-0.668,-0.105)
		β_{10}				
	estimate	-0.117				
	95% CI	(-0.209,-0.056)				
		γ	β_1	β_2	β_3	β_4
	estimate	0.482	0.176	0.147	-0.139	0.225
	95% CI	(0.426,0.541)	(-0.128,0.462)	(-0.142,0.471)	(-0.308,0.032)	(-0.134,0.607)
		β_5	β_6	β_7	β_8	β_9
0.30	estimate	0.286	0.279	-0.313	0.074	-0.396
	95% CI	(-0.098,0.665)	(0.142,0.438)	(-0.641,0.022)	(0.026,0.113)	(-0.718,-0.119)
		β_{10}				
	estimate	-0.149				
	95% CI	(-0.283,-0.033)				
		γ	β_1	β_2	β_3	β_4
	estimate	0.621	0.153	0.130	-0.181	0.175
	95% CI	(0.515,0.709)	(-0.087,0.431)	(-0.198,0.455)	(-0.412,0.047)	(-0.203,0.572)
		β_5	β_6	β_7	β_8	β_9
0.45	estimate	0.107	0.296	-0.294	0.035	-0.423
	95% CI	(-0.351,0.552)	(0.118,0.512)	(-0.655,0.087)	(0.012,0.069)	(-0.812,-0.045)
		β_{10}				
	estimate	-0.263				
	95% CI	(-0.479,-0.048)				

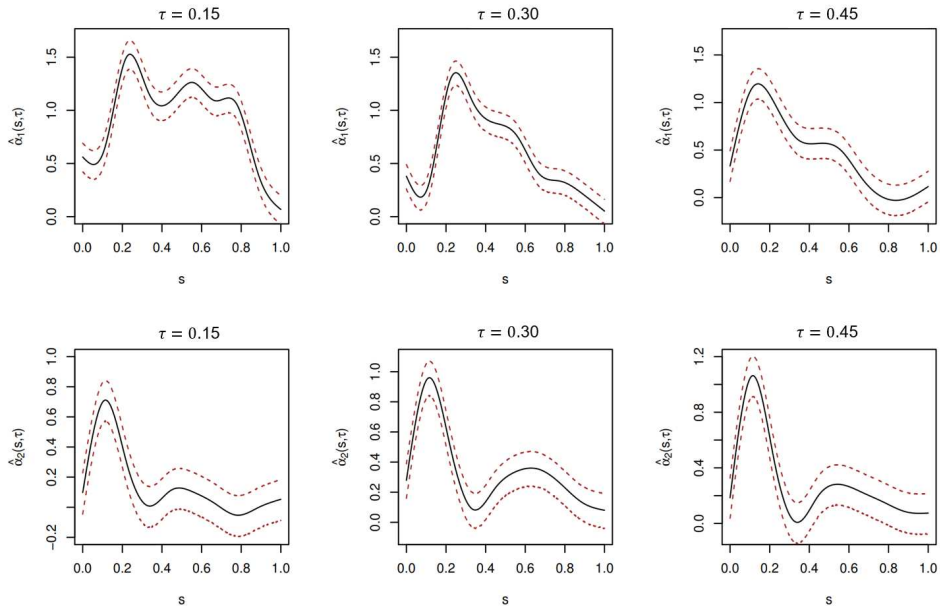

Fig. 6. The estimated functional effects $\hat{\alpha}_j(s,\tau), j = 1,2$ and the corresponding 95% bootstrap confidence intervals at different quantile values. Top panels: effects of left hippocampal radial distances. Bottom panels: effects of right hippocampal radial distances.

Figure 6 displays the estimated functional effects $\hat{\alpha}_j(s,\tau), j = 1,2$ and the corresponding 95% bootstrap confidence intervals at different quantile values. The results imply that both left and right HRDs are significant predictors across the quantiles and have positive effects on the conversion time to AD. Therefore, the atrophy of hippocampal morphology will accelerate the MCI progress to AD, i.e., we may expect a higher risk of AD conversion for MCI patients with thinner hippocampal morphology. This result is consistent with the finding that hippocampal radial distance is a good measure of deteriorating cognitive functions and serves as an important functional predictor of AD conversion time for MCI patients (Li and Luo, 2019). Comparing the scales of $\hat{\alpha}_1$ and $\hat{\alpha}_2$, we observe that the effects of the left and right HRDs are not identical. For the left hippocampus, regions spanning 0-0.8 have a strong impact at the lower quantile $\tau = 0.15$ and the effects reduce slightly with contributing regions mainly lying in $s \in [0, 0.6]$ at the quantile $\tau = 0.45$. For the right hippocampus, the locations between 0 and 0.2 play the dominant role with varying covariate effects at different quantile levels. Thus, it is necessary to pay attention to both left and right hippocampal radial atrophy and keep in mind the regional effects differences at the same time. Based on the hippocampus image processing procedure, the location in [0.1,0.4] is mainly concentrated at the CA1 and subiculum (Sub) subfields of the hippocampus, which indicates that the thinner these areas on the hippocampus are, the shorter the progression time is to AD. The map of the hippocampal subfields is illustrated in Figure 1 of the Supplementary material.

Since the paper employs a Box-Cox transformation in the model, the marginal effects

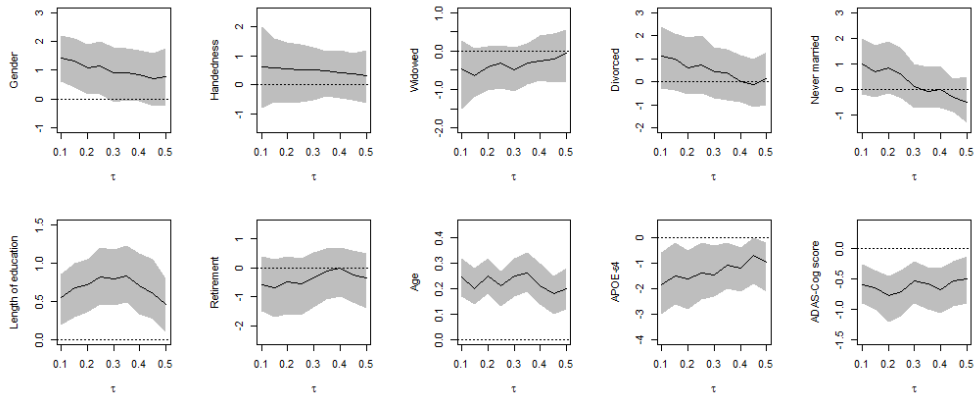

Fig. 7. Estimates and pointwise 95% confidence intervals for the marginal covariate effects for an individual in the ADNI data, with the continuous or categorical covariates fixed at the mean values or reference levels.

of the covariates are indeed functions of both $\gamma(\tau)$ and $\boldsymbol{\beta}(\tau)$. To examine the covariate effects on the original scale of the survival time, we may consider their marginal effects. Following Mu and He (2007) and Yin et al. (2008), if the jth covariate X_j is continuous, then its marginal effect is given by $\frac{\partial}{\partial X_j} Q_T(\tau \mid \mathbf{X}, \mathbf{Z}) = \beta_j(\tau) \cdot Q_T(\tau \mid \mathbf{X}, \mathbf{Z})^{1-\gamma(\tau)}$, where $\beta_j(\tau)$ is the jth component of $\boldsymbol{\beta}(\tau)$. For a discrete covariate taking values 0 and 1, its marginal effect is defined as $Q_T(\tau \mid \mathbf{X}, \mathbf{Z})|_{X_j=1} - Q_T(\tau \mid \mathbf{X}, \mathbf{Z})|_{X_j=0}$.

The marginal effect of X_j at a specified set of covariates $\{\boldsymbol{x}_0, \boldsymbol{z}_0\}$ is estimated by plugging in $(\hat{\gamma}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}})$. For illustration, the continuous covariates are taken to be the mean values and the categorical covariates are fixed at reference levels. Figure 7 presents the estimates and the pointwise 95% bootstrap confidence bands for the marginal covariate effects. It is seen that the marginal effects vary with the quantile values and the confidence intervals are relatively stable across different quantiles. In addition to the four significant covariates identified from Table 5, we find that Gender may also play a role in the AD progression. The positive effects of Gender indicate that females have relatively longer conversion time than males, though the impacts decrease with larger quantile values. Other covariates, such as Handedness and Marital status, seem not much important with the confidence intervals containing 0 at most quantiles. The estimated marginal coefficient of the significant allele, APOE-ϵ4, displays an increase from from -1.83 ($\tau = 0.1$) to -0.95 ($\tau = 0.5$), indicating that the presence of allele ϵ4 may have a greater detrimental effect at lower quantiles. This finding, which remains undetectable using conventional mean-based models, highlights the importance of our proposed method. Compared with the Cox proportional hazard model, the proposed quantile model in this paper presents a global view of the association between the AD conversion time and the covariates of interest, which could provide more valuable reference for the disease research.

The proposed method is also compared with three competing models based on the predictive performance. Model 1 employs the Log-spline method mentioned in Section 5.2

to analyze the same ADNI data with both left and right HRDs. Model 2 only incorporates the left HRDs and uses the univariant FPCA to handle this functional predictor. Model 3 replaces the HRD curve data with the hippocampal size, which is a scalar covariate. To predict the conversion time, the full data set is randomly split into a training set with 200 subjects and a test set with 173 subjects. The interquartile range is used as the prediction interval.This random split is repeated 100 times to obtain the average of coverage frequencies and interval widths for all the considered models. The coverage frequency is calculated as the proportion of the interquartile prediction intervals that cover the observed conversion time. It is found that the proposed model achieves the best predictive performance with a value of 0.521 for the coverage frequency. The deviation of Model 1 is significantly larger, leading to a frequency of 0.437. This indicates that the traditional log-transformation is inadequate to fit the data. Model 2 performs much better with a frequency of 0.462, and Model 3 achieves 0.441, both of which are inferior to the proposed model. This shows clearly that combining the left and right HRDs could provide substantially more information than only the left ones or the scalar covariates and the technique of multivariate FPCA successfully helps to extract important information from the correlated functional predictors.

8. Discussion

Alzheimer's disease is an irreversible brain disorder and understanding of its progression is quite beneficial for early intervention. In this study, we proposed a censored quantile transformation model with a multivariate functional predictor and applied it to ADNI data to investigate the risk factors of the MCI progression and predict the conversion time to onset of AD. The Box-Cox transformation is rather flexible and includes a broad class of model structures as special cases, thus relaxing the global linear assumption in the previous studies. To get rid of overlapping information in the correlated functional predictors, we carried out dimension reduction by means of multivariate functional principal component analysis. The resulting independent mFPC scores effectively avoid the multicollinearity in the regression. Uniform consistency and weak convergence of the quantile process are developed for the proposed martingale-based estimators. Simulation studies suggest that the proposed method outperforms the existing approaches, especially the broad applicability in the cases where curves and images are simultaneously collected as functional covariates.

Identifiability is an inherent and subtle issue in the censored quantile regression. Due to the loss of event information in the upper tail, regression quantiles with τ close to 1 may not be identifiable. In this paper, our focus is restricted to $\tau \in (0, \tau_U]$ with $\tau_U < 1$. Theoretically, selection of τ_U involves satisfying certain intrinsic identifiability conditions as outlined in the Supplementary material, while practically, τ_U is initially chosen based on the range of quantiles of interest and then adaptively adjusted (Wu et al., 2015). For example, if the optimization problem in the estimation becomes infeasible for some regression quantile at $\tau_j \in (0, \tau_U]$, it suggests that τ_U may exceed the upper bound of identifiability conditions and should be reset to a smaller value. Otherwise, no adjustment of τ_U is needed or we can increment τ_U by a small fraction to gradually push it to the largest estimable quantile level.

The application to the ADNI study reveals that Length of education, Age, APOE-ϵ4

carrirer and ADAS-cog score are important predictors for AD. These results are basically consistent with those in the previous mean-based or Cox models, but the proposed quantile method apparently provides a more global view for the parameter estimation at different quantile levels, especially for lower quantiles, which are of primary interest to researchers. Moreover, the data-driven transformation parameter $\gamma(\tau)$ is dynamically estimated to reveal the non-linear and varying relationships between the risk factors and disease progression. It is also shown that including multiple functional predictors (both left and right HRD curves) is necessary to improve the prediction accuracy for the MCI-AD conversion, although some previous studies only demonstrated that the left hippocampal volume was associated with delayed verbal memory. Based on the pointwise processing of the hippocampus scans, the significant CA1 and subicular subfields are identified, which may facilitate case studies for evaluating clinical efficacy in slowing AD progression. Besides, the proposed model can readily incorporate multiple brain region profiles as functional predictors to assess their association with Alzheimer's disease. It is worth mentioning that patients in the ADNI study undergo periodic examinations, which may result in interval-censored conversion time. However, since the study was conducted for a long period, the uncertainty caused by interval-censoring has little effect on the analysis. Nevertheless, this aspect represents an intriguing area for exploration, with potential applicability to other research domains, warranting further investigation in future studies.

According to the estimation procedure in Frumento and Bottai (2017), the coefficient functions can be obtained on the quantile process by imposing smoothness conditions on the coefficients, which may facilitate interpretation of the results. In practice, the selection of basis functions is critical to balance the parsimony and flexibility. In the absence of prior knowledge, one may employ polynomials, known quantile functions, trigonometric functions, splines, and combinations of the above. However, important criteria for model selection are usually needed to compare numerous alternative models. In some complex cases, it is not easy to select a proper set of basis functions for the parametric modeling of quantile regression coefficients. The performance of the parameterized estimators merits further investigations, which represents an important subject for quantile regression.

ACKNOWLEDGEMENTS

The authors thank the Editor, AE, and reviewers for their constructive comments, which have greatly improved the early version of this paper. Data used in preparation of this paper were obtained from the ADNI database. As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this article. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`.

FUNDING

The work of S. P. Ma was supported by the Fundamental Research Funds for the Central Universities in UIBE (23QD03). The work of M. L. Tang was partially supported by the Research Matching Grant (project: 700006 Applications of SAS Viya in Big Data

Analytics) and FDS Grant (UGC/FDS14/P05/20) from the Research Grants Council of the Hong Kong Special Administration Region, and the Big Data Intelligence Centre in The Hang Seng University of Hong Kong. The work of Z. H. Wang was supported by Natural Science Foundation of Xinjiang Uygur Autonomous Region(2023D01A74).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are provided by the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`https://adni.loni.usc.edu`) under licence. Data will be shared on reasonable request to the corresponding author with permission of ADNI.

SUPPLEMENTARY MATERIAL

Additional supporting information, including the technical details, simulation results and R codes, are provided in the Supplementary material, which may be found in the online version of the article at the publisher's website.

References

Bang, H. and A. A. Tsiatis (2002). Median regression with censored cost data. *Biometrics 58*(3), 643–649.

Box, G. E. and D. R. Cox (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological) 26*(2), 211–243.

Buchinsky, M. (1995). Quantile regression, box-cox transformation model, and the us wage structure, 1963–1987. *Journal of econometrics 65*(1), 109–154.

Cardot, H., C. Crambes, and P. Sarda (2005). Quantile regression when the covariates are functions. *Nonparametric Statistics 17*(7), 841–856.

Chen, X. and D. Pouzo (2012). Estimation of nonparametric conditional moment models with possibly nonsmooth generalized residuals. *Econometrica 80*(1), 277–321.

Chernozhukov, V. and H. Hong (2002). Three-step censored quantile regression and extramarital affairs. *Journal of the American statistical Association 97*(459), 872–882.

Chu, C. W., T. Sit, and G. Xu (2021). Transformed dynamic quantile regression on censored data. *Journal of the American Statistical Association 116*(534), 874–886.

Colom, R., J. L. Stein, P. Rajagopalan, K. Martínez, D. Hermel, Y. Wang, J. Álvarez-Linera, M. Burgaleta, M. Á. Quiroga, P. C. Shih, et al. (2013). Hippocampal structure and human cognition: Key role of spatial processing and evidence supporting the efficiency hypothesis in females. *Intelligence 41*(2), 129–140.

Corder, E. H., A. M. Saunders, W. J. Strittmatter, D. E. Schmechel, P. C. Gaskell, G. Small, A. Roses, J. Haines, and M. A. Pericak-Vance (1993). Gene dose of apolipoprotein e type 4 allele and the risk of alzheimer's disease in late onset families. *Science 261*(5123), 921–923.

Cui, Y., B. Liu, S. Luo, X. Zhen, M. Fan, T. Liu, W. Zhu, M. Park, T. Jiang, J. S. Jin, et al. (2011). Identification of conversion from mild cognitive impairment to alzheimer's disease using multivariate predictors. *PloS one 6*(7), e21896.

Da, X., J. B. Toledo, J. Zee, D. A. Wolk, S. X. Xie, Y. Ou, A. Shacklett, P. Parmpi, L. Shaw, J. Q. Trojanowski, et al. (2014). Integration and relative value of biomarkers for prediction of mci to ad progression: spatial patterns of brain atrophy, cognitive scores, apoe genotype and csf biomarkers. *NeuroImage: Clinical 4*, 164–173.

Fan, J., I. Gijbels, T.-C. Hu, and L.-S. Huang (1996). A study of variable bandwidth selection for local polynomial regression. *Statistica Sinica*, 113–127.

Fei, Z., Q. Zheng, H. G. Hong, and Y. Li (2023). Inference for high-dimensional censored quantile regression. *Journal of the American Statistical Association 118*(542), 898–912.

Fleming, T. R. and D. P. Harrington (2011). *Counting processes and survival analysis*. John Wiley & Sons.

Frumento, P. and M. Bottai (2016). Parametric modeling of quantile regression coefficient functions. *Biometrics 72*(1), 74–84.

Frumento, P. and M. Bottai (2017). Parametric modeling of quantile regression coefficient functions with censored and truncated data. *Biometrics 73*(4), 1179–1188.

Happ, C. and S. Greven (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *Journal of the American Statistical Association 113*(522), 649–659.

Jiang, F., Q. Cheng, G. Yin, and H. Shen (2020). Functional censored quantile regression. *Journal of the American Statistical Association 115*(530), 931–944.

Jiang, S., Y. Xie, and G. A. Colditz (2021). Functional ensemble survival tree: Dynamic prediction of alzheimer's disease progression accommodating multiple timevarying covariates. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 70*(1), 66 – 79.

Jin, Z., Z. Ying, and L. Wei (2001). A simple resampling method by perturbing the minimand. *Biometrika 88*(2), 381–390.

Kato, K. (2012). Estimation in functional linear quantile regression. *The Annals of Statistics 40*(6), 3108–3136.

Koenker, R. and O. Geling (2001). Reappraising medfly longevity: a quantile regression survival analysis. *Journal of the American Statistical Association 96*(454), 458–468.

Kong, D., J. G. Ibrahim, E. Lee, and H. Zhu (2018). Flcrm: Functional linear cox regression model. *Biometrics 74*(1), 109–117.

LaFerla, F. M., K. N. Green, and S. Oddo (2007). Intracellular amyloid-β in alzheimer's disease. *Nature Reviews Neuroscience 8*(7), 499–509.

Lee, E., H. Zhu, D. Kong, Y. Wang, K. Sullivan Giovanello, J. Ibrahim, N. Initiative, et al. (2015). Bflcrm: A bayesian functional linear cox regression model for predicting time to conversion to alzheimer's disease. *Ann. Appl. Stat 9*, 2153–2178.

Leng, C. and X. Tong (2014). Censored quantile regression via box-cox transformation under conditional independence. *Statistica Sinica 24*(1), 221–249.

Li, K. and S. Luo (2017). Functional joint model for longitudinal and time-to-event data: an application to alzheimer's disease. *Statistics in medicine 36*(22), 3560–3572.

Li, K. and S. Luo (2019). Dynamic predictions in bayesian functional joint models for longitudinal and time-to-event data: An application to alzheimers disease. *Statistical methods in medical research 28*(2), 327–342.

Li, S., O. Okonkwo, M. Albert, and M.-C. Wang (2013). Variation in variables that predict progression from mci to ad dementia over duration of follow-up. *American journal of Alzheimer's disease (Columbia, Mo.) 2*(1), 12.

Li, Y., N. Wang, and R. J. Carroll (2010). Generalized functional linear models with semiparametric single-index interactions. *Journal of the American Statistical Association 105*(490), 621–633.

Luders, E., P. M. Thompson, F. Kurth, J.-Y. Hong, O. R. Phillips, Y. Wang, B. A. Gutman, Y.-Y. Chou, K. L. Narr, and A. W. Toga (2013). Global and regional alterations of hippocampal anatomy in long-term meditation practitioners. *Human brain mapping 34*(12), 3369–3375.

Ma, H., T. Li, H. Zhu, and Z. Zhu (2019). Quantile regression for functional partially linear model in ultra-high dimensions. *Computational Statistics & Data Analysis 129*, 135–147.

Mattson, M. P. (2004). Pathways towards and away from alzheimer's disease. *Nature 430*(7000), 631–639.

Mu, Y. and X. He (2007). Power transformation toward a linear regression quantile. *Journal of the American Statistical Association 102*(477), 269–279.

Peng, L. and Y. Huang (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association 103*(482), 637–649.

Petersen, R. C., G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of neurology 56*(3), 303–308.

Petersen, R. C., R. G. Thomas, M. Grundman, D. Bennett, R. Doody, S. Ferris, D. Galasko, S. Jin, J. Kaye, A. Levey, et al. (2005). Vitamin e and donepezil for the treatment of mild cognitive impairment. *New England Journal of Medicine 352*(23), 2379–2388.

Portnoy, S. (2003). Censored regression quantiles. *Journal of the American Statistical Association 98*(464), 1001–1012.

Qian, J. and L. Peng (2010). Censored quantile regression with partially functional effects. *Biometrika 97*(4), 839–850.

Rabin, J. S., H. Klein, D. R. Kirn, A. P. Schultz, H.-S. Yang, O. Hampton, S. Jiang, R. F. Buckley, A. Viswanathan, T. Hedden, et al. (2019). Associations of physical activity and β-amyloid with longitudinal cognition and neurodegeneration in clinically normal older adults. *JAMA neurology 76*(10), 1203–1210.

Reiss, P. T. and R. T. Ogden (2010). Functional generalized linear models with images as predictors. *Biometrics 66*(1), 61–69.

Shi, J., N. Lepore, B. A. Gutman, P. M. Thompson, L. C. Baxter, R. J. Caselli, Y. Wang, and A. D. N. Initiative (2014). Genetic influence of apolipoprotein e4 genotype on hippocampal morphometry: An n= 725 surface-based alzheimer's disease neuroimaging initiative study. *Human brain mapping 35*(8), 3903–3918.

Shi, J., P. M. Thompson, B. Gutman, Y. Wang, A. D. N. Initiative, et al. (2013). Surface fluid registration of conformal representation: Application to detect disease burden and genetic influence on hippocampus. *NeuroImage 78*, 111–134.

Tang, Q. and L. Cheng (2014). Partial functional linear quantile regression. *Science China Mathematics 57*(12), 2589–2608.

Wang, Y., J. G. Ibrahim, and H. Zhu (2020). Partial least squares for functional joint models with applications to the alzheimer's disease neuroimaging initiative study. *Biometrics 76*(4), 1109–1119.

Wang, Y., Y. Song, P. Rajagopalan, T. An, K. Liu, Y.-Y. Chou, B. Gutman, A. W. Toga, P. M. Thompson, A. D. N. Initiative, et al. (2011). Surface-based tbm boosts power to detect disease effects on the brain: an n= 804 adni study. *Neuroimage 56*(4), 1993–2010.

Wong, R. K., Y. Li, and Z. Zhu (2019). Partially linear functional additive models for multivariate functional data. *Journal of the American Statistical Association 114*(525), 406–418.

Wu, C., N. Ling, P. Vieu, and W. Liang (2023). Partially functional linear quantile regression model and variable selection with censoring indicators mar. *Journal of Multivariate Analysis 197*, 105189.

Wu, Y., Y. Ma, and G. Yin (2015). Smoothed and corrected score approach to censored quantile regression with measurement errors. *Journal of the American Statistical Association 110*(512), 1670–1683.

Yan, F., X. Lin, and X. Huang (2017). Dynamic prediction of disease progression for leukemia patients by functional principal component analysis of longitudinal expression levels of an oncogene. *The Annals of Applied Statistics 11*(3), 1649–1670.

Yan, F., X. Lin, R. Li, and X. Huang (2018). Functional principal components analysis on moving time windows of longitudinal data: dynamic prediction of times to event. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 67*(4), 961–978.

Yin, G., D. Zeng, and H. Li (2008). Power-transformed linear quantile regression with censored data. *Journal of the American Statistical Association 103*(483), 1214–1224.

Young, J., M. Modat, M. J. Cardoso, A. Mendelson, D. Cash, S. Ourselin, A. D. N. Initiative, et al. (2013). Accurate multimodal probabilistic prediction of conversion to alzheimer's disease in patients with mild cognitive impairment. *NeuroImage: Clinical 2*, 735–745.