



# Jointly modelling multiple transplant outcomes by a competing risk model via functional principal component analysis

Jianghu (James) Dong, Haolun Shi, Liangliang Wang, Ying Zhang & Jiguo Cao

To cite this article: Jianghu (James) Dong, Haolun Shi, Liangliang Wang, Ying Zhang & Jiguo Cao (2021): Jointly modelling multiple transplant outcomes by a competing risk model via functional principal component analysis, Journal of Applied Statistics, DOI: [10.1080/02664763.2021.1981256](https://doi.org/10.1080/02664763.2021.1981256)

To link to this article: <https://doi.org/10.1080/02664763.2021.1981256>



Published online: 22 Sep 2021.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)



# Jointly modelling multiple transplant outcomes by a competing risk model via functional principal component analysis

Jianghu (James) Dong <sup>a,b</sup>, Haolun Shi <sup>c</sup>, Liangliang Wang <sup>c</sup>, Ying Zhang <sup>a</sup> and Jiguo Cao <sup>c</sup>

<sup>a</sup>Department of Biostatistics, College of Public Health, University of Nebraska Medical Center, Omaha, Nebraska, USA; <sup>b</sup>Division of Nephrology, Department of Medicine, University of Nebraska Medical Center, Omaha, Nebraska, USA; <sup>c</sup>Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada

## ABSTRACT

In many clinical studies, longitudinal biomarkers are often used to monitor the progression of a disease. For example, in a kidney transplant study, the glomerular filtration rate (GFR) is used as a longitudinal biomarker to monitor the progression of the kidney function and the patient's state of survival that is characterized by multiple time-to-event outcomes, such as kidney transplant failure and death. It is known that the joint modelling of longitudinal and survival data leads to a more accurate and comprehensive estimation of the covariates' effect. While most joint models use the longitudinal outcome as a covariate for predicting survival, very few models consider the further decomposition of the variation within the longitudinal trajectories and its effect on survival. We develop a joint model that uses functional principal component analysis (FPCA) to extract useful features from the longitudinal trajectories and adopt the competing risk model to handle multiple time-to-event outcomes. The longitudinal trajectories and the multiple time-to-event outcomes are linked via the shared functional features. The application of our model on a real kidney transplant data set reveals the significance of these functional features, and a simulation study is carried out to validate the accurateness of the estimation method.

## ARTICLE HISTORY

Received 14 September 2020  
Accepted 9 September 2021

## KEYWORDS

Competing risks; functional principal component analysis; joint model; latent variables; kidney transplant

## 1. Background and introduction

Various studies, such as Levey *et al.* [11] and Wolfe *et al.* [17], have shown that kidney transplantation prolongs the survival of patients with end-stage renal disease. As patients may experience acute rejection or graft failure post-transplantation, how to extend the long-term survival of the kidney graft remains the main scientific question for transplant studies. If the rate of kidney graft failure can be reduced, the overall patient population would enjoy a longer survival time. Surrogate markers have been proposed to predict kidney graft

failure. For example, Marcen *et al.* [12] and Moranne *et al.* [13] proposed to use the slope of the GFR trajectories to predict graft failure using a Cox model.

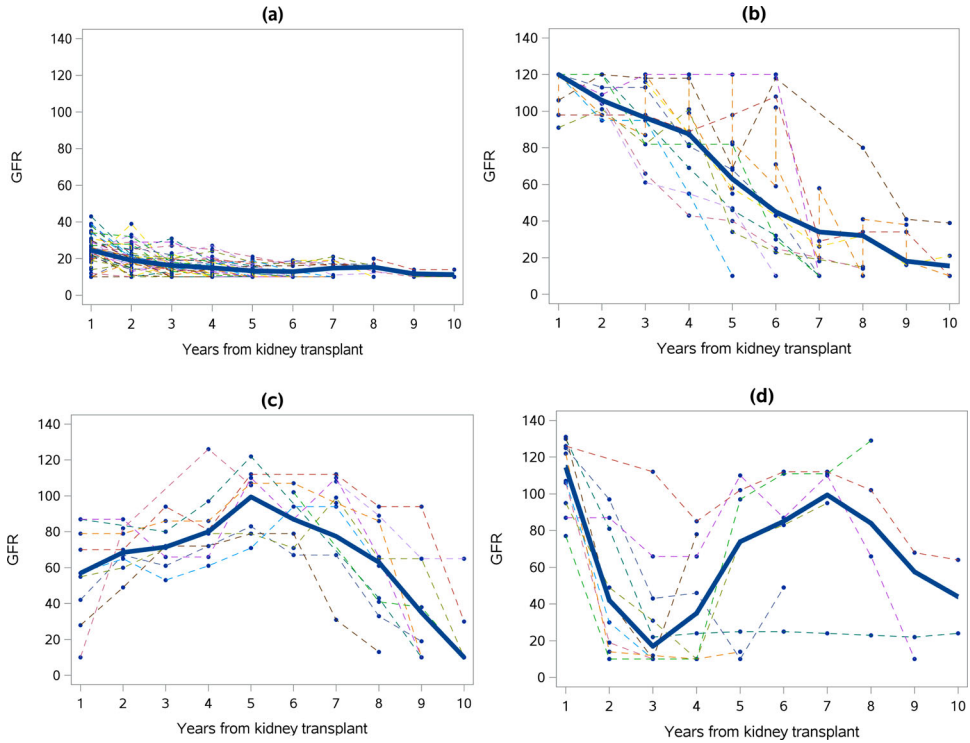
Three key questions pertain to understanding the interconnection between the GFR trajectories and the long-term transplant outcomes. The first question is how to fit a continuous trajectory from the repeated longitudinal GFR measurements. The second question is how to model the survival hazard of multiple outcomes/events simultaneously, as kidney recipients post-transplantation are subject to competing risks of transplant failure as well as death from causes other than transplant failure. The third question is how to identify the effects of the shape of the longitudinal trajectories on the prediction of multiple time-to-event outcomes.

To address the first question, several methods have been developed. For example, parametric models are commonly used to fit the GFR trajectories, e.g. Marcen *et al.* [12], Moranne *et al.* [13], and Dong *et al.* [5]. Another alternative is to use a nonparametric approach. For example, the functional principal component analysis is adopted by Dong *et al.* [4] to explore the major sources of variation among the GFR trajectories. The dimension of the GFR trajectories is effectively reduced and each curve can be represented by four functional principal components (FPCs). The top four FPCs account for 99.8% of the total variation.

To address the second question, various survival models have been proposed to handle competing events. In the competing risk framework, two popular competing risk models are used. One is the cause-specific hazard model proposed by Prentice *et al.* [14] and Putter *et al.* [15], and the other is the subdistribution hazards regression introduced by Fine and Gray [7]. We model the multiple time-to-event outcomes via the latter approach, which is based on a reweighted risk set for consistent estimation of the regression coefficients.

To address the third question, the longitudinal outcomes and the time-to-event outcomes are linked using the FPC scores as the shared latent features. As shown by Dong *et al.* [4], the four FPCs relate to four primary patterns of variation within the GFR trajectories. Figure 1 shows the GFR trajectories of four clusters of patients whose trajectories are dominated by their first, second, third, and fourth FPC scores, respectively. The risk of kidney transplant failure or death might be different when a patient's GFR trajectory is flat versus when the trajectory highly fluctuates. It is thus of interest to explore whether and how the progression of kidney function differs among these four clusters, and a natural approach is to use the FPC scores as the shared covariates between the longitudinal model and the survival model.

Several proposed joint models for longitudinal outcomes and time-to-event outcomes have been constructed for FPCA. Yao [18] developed a joint model where FPCA is used to fit the longitudinal trajectories and the longitudinal outcome is treated as a covariate in the Cox regression model. Ding and Wang [3] proposed a joint model that treats longitudinal outcomes as nonparametric multiplicative random effects within the Cox proportional hazard framework. However, the two joint models mentioned above can only accommodate a single time-to-event outcome. A number of joint models have been developed recently for longitudinal and competing risk data. For example, Hickey *et al.* [8] has given the summary for the published joint models with competing-risks event. However, none of these joint models for longitudinal and competing risk data were set up for FPCA. It is of interest to determine the relationship between the patient's progression of kidney function and the dominant variation patterns of the GFR trajectories in our clinical kidney



**Figure 1.** The 4 clusters of observable GFR curves by FPC scores from our preliminary analysis. The thick blue curve is the average of GFR curves in each panel.

transplantation data. Therefore, we propose a new joint model based on the FPC scores as the shared latent features between the longitudinal and survival components.

Our model uses functional principal component analysis for the modeling of the longitudinal measurements and a competing risk subdistribution hazard model for handling multiple time-to-event outcomes. The main highlight of this paper is that after reflecting on the three key clinical questions in the kidney transplant studies, we tailor a new joint model to adequately address them. To the best of our knowledge, the proposed model is the first to explore the relationship between the pattern of variation of the GFR trajectories and the patient's state of survival that is subject to multiple time-to-event outcomes, using FPC scores as the latent shared features between the longitudinal model and the survival model.

The rest of this article is organized as follows. The proposed joint model is introduced in Section 2. We present the estimation method for the proposed joint model in Section 3. Section 4 demonstrates the application of our joint model in the kidney transplant data. Section 5 presents a simulation study to investigate the finite sample performances of our joint model. Conclusions and discussion are given in Section 6.

## 2. Joint model

Let  $T_i$  and  $C_i$  respectively denote the event and censoring times for the  $i$ th subject, where  $i = 1, \dots, N$ . Let  $Y_i(t)$  denote the longitudinal outcome of the  $i$ th subject at time  $t$ ,  $t \in$

$\mathcal{T}$ , and  $Z_i$  the covariate vector of the  $i$ th subject. We observe  $X_i = \min(T_i, C_i)$  and the censoring indicator  $\Delta_i = I(T_i \leq C_i)$ . Each  $T_i$  may correspond to one of  $M$  different event types, and we denote  $m_i \in (1, \dots, M)$  as the index for the observed event types. We assume that the subjects are independent with each other.

## 2.1. Longitudinal model

The model for the longitudinal outcome  $Y_i(t)$  is based on functional principal component analysis, which decomposes the underlying random stochastic process into a linear combination of functional principal components. As mentioned in Dong *et al.* [4], for the analysis of GFR curves, the principal component analysis through the conditional expectation (PACE) approach is well-suited for conducting FPCA and handling longitudinal outcome with possibly missing values and measurement errors. The obtained first 4 leading FPCs account for a majority of the variation (99.8%). We adopt such an approach for fitting the longitudinal measurements.

Let  $Y_i^*(t) = Y_i(t) - \hat{\alpha}^\top Z_i$  be the longitudinal process adjusted for the effects of the covariates. We decompose  $Y_i^*(t)$  as

$$Y_i^*(t) = \mu(t) + \sum_{k=1}^K \xi_{ik} \phi_k(t) + \epsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where  $\epsilon_i$  are identically and independently distributed normal measurement error terms with mean 0 and variance  $\sigma^2$ . The function  $\phi_k(t)$  is the  $k$ th functional principal component, which satisfies  $\int_{\mathcal{T}} \phi_k(t) \phi_j(t) dt = \delta_{kj}$ , where  $\delta_{kj} = 1$  if  $k = j$  and 0 otherwise. The  $\xi_{ik}$  is the associated functional principal component score for the  $i$ th subject and the  $k$ th component, which is defined as

$$\xi_{ik} = \int_{\mathcal{T}} \{Y_i^*(t) - \mu(t)\} \phi_k(t) dt.$$

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

By the Mercer's theorem, the covariance function between any two time points  $s$  and  $t$  in the time-period  $\mathcal{T}$ , defined as  $G(s, t) = \text{Cov}(Y_i^*(s) - \mu(s), Y_i^*(t) - \mu(t))$ , can be expressed as

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

As it would be unrealistic to estimate an infinite number of  $\phi_k(t)$ , in reality,  $Y_i^*(t)$  is usually well approximated by retaining only the first  $K$  leading FPCs.

To estimate the FPCs, the first step is to establish smoothed estimates of the mean and covariance functions. The mean function is obtained by smoothing the data from all observations based on the one-dimensional local linear smoother [6], and the covariance function  $G(s, t)$  is estimated by a two-dimensional smoother [10,19]. Let  $\hat{\mu}(t)$  and  $\hat{G}(s, t)$  denote the estimated mean trajectory and the estimated smoothed covariance surface.

To obtain estimates for  $\phi_k(t)$  and  $\lambda_k$ , we solve the eigenequation

$$\int_{\mathcal{T}} \widehat{G}(s, t) \phi_k(s) \, ds = \lambda_k \phi_k(t), \quad (2)$$

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

Let  $n_i$  denote the number of timepoints on the trajectory of the  $i$ th subject. Let  $\widehat{\phi}_{ik}$  and  $\widehat{\mu}_i$  denote the vectors of values of  $\widehat{\phi}_k(t)$  and  $\widehat{\mu}(t)$  evaluated at time points of the  $i$ th subject, and let  $\widetilde{\mathbf{G}}_i$  denote the matrix of values of  $\widehat{G}(s, t)$  evaluated at the two-dimensional grid consisting of time points of the  $i$ th subject. The FPC score of the  $i$ th subject and the  $k$ th FPC is computed from the conditional expectation

$$\widehat{\xi}_{ik} = \widehat{\mathbf{E}}(\xi_{ik} \mid \mathbf{Y}_i^*) = \lambda_k \widehat{\phi}_k^\top (\widetilde{\mathbf{G}}_i + \widehat{\sigma}^2 \mathbf{I})^{-1} (\mathbf{Y}_i^* - \widehat{\mu}_i),$$

where  $\mathbf{Y}_i^* = (Y_{i1}^*, \dots, Y_{in_i}^*)^\top$  is the vector of covariate-adjusted longitudinal data points of the  $i$ th subject, and  $\mathbf{I}$  is an identity matrix of size  $n_i$ .

## 2.2. Competing risk model

The competing risk model is well-suited for our analysis of the time to multiple competing events of interest. Without loss of generality, we consider two types of events, one event of primary interest, and another event constituting a competing risk. If a competing event occurs before the event of primary interest, the primary event of interest would no longer be observable. For example, in the case of our application, kidney failure is the event of primary interest, whereas death from other causes constitutes the competing event. Censoring of the primary event (kidney failure) is not only due to loss of follow-up but also due to the competing event (death from other causes).

Due to the existence of a competing event, the usual Kaplan-Meier estimate of the survival function of the primary event would render biased results. This is because the usual assumption that any subject with a censored observation will eventually experience the primary event of interest no longer holds due to the presence of the competing event. If the competing event occurs prior to the primary event, it would be impossible for the subject to experience the event of primary interest, thus violating the assumption of the Kaplan-Meier method. Under the Kaplan-Meier method, the cumulative probability function will approach one given infinite time of follow-up; whereas under the presence of the competing risk, there remain a proportion of subjects who are affected by the competing event and will never experience the event of primary interest.

An alternative to the conventional cumulative probability function under the Kaplan-Meier method is the cumulative incidence function. Denote  $m = 1, 2$  as the index for the primary event and the competing event, respectively, and let  $h_m(t)$  denote the cause-specific hazard function for the  $m$ th event type. The cumulative incidence function of the  $m$ th event is

$$F_m(t) = \int_0^t S(s) h_m(s) \, ds,$$

where

$$S(t) = \exp \left\{ - \int_0^t (h_1(s) + h_2(s)) ds \right\}.$$

The survival function  $S(t)$  is the probability that the subject survives to time  $t$  without experiencing events 1 or 2. The cumulative incidence function can be regarded as the marginal subdistribution function of the time-to-event of a given type.

From the cumulative incidence function, the definition of the subdistribution hazard function is naturally derived as

$$\bar{h}_m(t) = - \frac{d}{dt} \log (1 - F_m(t)).$$

It is worth noting the  $\bar{h}_m(t)$  is a better characterization of the cumulative incidence function as opposed to the cause-specific hazard function  $h_m(t)$ . A transformation involving the integration of  $\bar{h}_m(t)$  would lead to the cumulative incidence function, i.e.  $F_m(t) = 1 - \exp\{-\int_0^t \bar{h}_m(s) ds\}$ , whereas conducting the same transformation on the  $h_m(t)$  would render an improper function that lacks meaningful interpretation.

To incorporate the effects of covariates into the subdistribution hazard function and thus direct effects on the cumulative incidence function, we may impose a proportional hazards structure on the hazard assumption as

$$\bar{h}_m(t | \xi_i, \mathbf{Z}_i) = \bar{h}_{m,0}(t) \exp\{\boldsymbol{\gamma}_m^\top \xi_i + \boldsymbol{\beta}_m^\top \mathbf{Z}_i\},$$

where  $\bar{h}_{m,0}(t)$  is the baseline subdistribution hazard function.

The estimate of the covariate effects in the subdistribution hazard function is not as straightforward as establishing separate estimations for different causes from the Cox model. Gray and Fine [7] proposed a reweighted approach on a modified risk set for consistent estimation of the regression coefficients. To be specific, to adjust for diminishing observability due to censoring, subjects who have experienced a competing event are retained in the risk set with weight tapering off as time progresses according to a function that depends on the censoring distribution.

### 2.3. Model setup

For the longitudinal process, in addition to extracting the major sources of variation via FPCA, we are interested in the association between the longitudinal outcome and the covariate vector  $\mathbf{Z}_i$ . For the survival process, it is of interest to model the cumulative incidence function  $F_m(t | \xi_i, \mathbf{Z}_i) = \Pr(T \leq t, m_i \Delta_i = m | \xi_i, \mathbf{Z}_i)$ , which is the probability that the event of type  $m$  occurs at or before time  $t$ . This is a conditional probability on the covariate vector as well as the functional principal component scores from the longitudinal process, denoted as  $\xi_i = (\xi_{1i}, \dots, \xi_{Ki})^\top$ , where  $K$  is the number of FPCs extracted.

We propose the following joint model for the longitudinal outcome  $Y_i(t)$  and the subdistribution hazard of multiple competing time-to-event outcomes,

$$\begin{cases} Y_i(t) = \mu(t) + \boldsymbol{\Phi}(t)^\top \xi_i + \boldsymbol{\alpha}^\top \mathbf{Z}_i + \epsilon_i, \\ \bar{h}_m(t | \xi_i, \mathbf{Z}_i) = \bar{h}_{m,0}(t) \exp\{\boldsymbol{\gamma}_m^\top \xi_i + \boldsymbol{\beta}_m^\top \mathbf{Z}_i\}. \end{cases} \quad (3)$$



The first equation represents the model for the longitudinal outcome  $Y_i(t)$ , where  $\mu(t)$  is the overall mean function of the longitudinal outcome,  $\Phi(t) = (\phi_1(t), \dots, \phi_K(t))^T$  is the vector of the functional principal component,  $\xi_i = (\xi_{i1}, \dots, \xi_{iK})^T$  the vector of FPC scores for the  $i$ th subject, i.e.  $\Phi(t)^T \xi_i = \sum_{k=1}^K \xi_{ik} \phi_k(t)$ , and  $\alpha = (\alpha_1, \dots, \alpha_p)^T$  is a vector of coefficients for the fixed effects of  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ip})^T$ . It should be noted that the usual Mercer expansion would have  $K \rightarrow \infty$  and we retain only the top  $K$  FPCs that explain the majority of variation among the longitudinal trajectories. The FPC scores  $\xi_i$  serve as the shared features in the survival model, and they link the longitudinal outcome and the competing time-to-event outcome in the joint model.

The second equation is a competing risk survival model. The subdistribution hazard function  $\bar{h}_m(t | \xi_i, \mathbf{Z}_i)$  is the hazard rate of the conditional cumulative incidence function  $F_m(t | \xi_i, \mathbf{Z}_i)$  for the  $m$ th competing risk,  $\gamma_m = (\gamma_{m1}, \dots, \gamma_{mK})^T$  is the vector of coefficients for the random effects of FPC scores  $\xi_i$ , and  $\beta_m = (\beta_{m1}, \dots, \beta_{mp})^T$  is the vector of coefficients for the fixed effects of  $\mathbf{Z}_i = [Z_{i1}, \dots, Z_{ip}]^T$ . We take  $m = 1$  as the primary event of interest corresponding to the event of kidney transplant failure.

There are several advantages of incorporating the FPC scores as the latent shared features in the new joint model. Contrary to the conventional models [18], which use the longitudinal outcome  $Y_i(t)$  as a covariate in the proportional hazard structure, the FPC scores, which represent the pattern of variation in the longitudinal curves, may offer a better characterization of the relationship between the longitudinal process and the survival outcomes. In particular, in the case of the kidney transplant data, prior studies [5] have identified notable variation in the GFR trajectories, e.g. one group of patients may have very flat GFR trajectories, whereas another group may exhibit huge variation in the trajectories in terms of shape or slope. It is thus of interest to study the effects of the curve patterns on the survival incidence function.

### 3. The estimation method

#### 3.1. The joint likelihood functions

This section gives the inference of the proposed joint model. Let  $(t_i, \Delta_i, m_i \Delta_i, \mathbf{Z}_i, Y_i(t))$  denote the observations of each subject in the data, where  $t_i$  is the observed survival time,  $m_i$  is the observed event type  $(1, \dots, M)$ ,  $\Delta_i$  is the censoring indicator of event,  $\mathbf{Z}_i$  is the observed covariate, and  $Y_i(t)$  is the longitudinal outcome. Let  $C_i$  be a potential censoring time, and  $T_i$  be the event time. We assume that  $X_i = \min(T_i, C_i)$ , and  $\Delta_i = I(T_i \leq C_i)$ . The parameters  $\Theta = (\gamma, \beta, h_0, \Lambda, \sigma^2)$  need to be estimated from data, where  $\gamma = (\gamma_1^T, \dots, \gamma_M^T)^T$ ,  $\beta = (\beta_1^T, \dots, \beta_M^T)^T$ , and  $\Lambda = (\lambda_1, \dots, \lambda_K)^T$  and  $\lambda_k$  is the variance for the FPC score  $\xi_{ik}$ .

In order to estimate the parameters, we need to construct the joint likelihood functions. As mentioned in Section 2.1, let  $\mathbf{Y}_i^*(t) = \mathbf{Y}_i(t) - \hat{\alpha}^T \mathbf{Z}_i$  be the longitudinal process after adjusted for the effects of the covariates. We assume that the longitudinal outcome  $Y_i(t)$  and the time-to-event  $T_i$  are conditionally independent given the latent FPC scores  $\xi_i$ . The longitudinal trajectories of  $Y_i^*(t)$  can be determined by the FPC score  $\xi_i = (\xi_{i1}, \dots, \xi_{iK})^T$  as shown in Section 2.1, so the joint probability density function of  $\mathbf{Y}_i^*(t)$  and the time-to-event  $T_i$  can be written as the factorization of the density distribution of the FPC score  $\xi_i$  and the conditional survival density distribution of  $T_i$  on the latent FPC score  $\xi_i$ . Also,



we assume that the subjects who are censored at time  $t$  should be representative of all the subjects in that subgroup who remained at risk at time  $t$  with respect to their survival experience. In other words, censoring is independently provided that it is random within any subgroup of interest. The full likelihood of the full set of parameters under independent censoring can be given by:

$$L(\Theta) = \prod_{i=1}^n \prod_{m_i=1}^M \left\{ \int f(T_i, \Delta_i, m_i \Delta_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta}) f(\mathbf{Y}_i^*(t) | \xi_i, \sigma) f(\xi_i | \Lambda) d\xi_i \right\}, \quad (4)$$

where the density survival function  $f(T_i, \Delta_i, m_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta})$  is given by

$$f(T_i, \Delta_i, m_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta}) = h_m(t_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta})^{I_{m_i \Delta_i = m}} S(t_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta})^{1 - I_{m_i \Delta_i = m}},$$

the longitudinal function

$$f(\mathbf{Y}_i^*(t) | \xi_i, \sigma) = (2\pi\sigma^2)^{-\frac{n_i}{2}} \exp \left\{ -\frac{1}{2\sigma^2} (\mathbf{Y}_i^*(t) - \boldsymbol{\mu}_i(t))^T (\mathbf{Y}_i^*(t) - \boldsymbol{\mu}_i(t)) \right\}$$

and  $\boldsymbol{\mu}_i(t) = \mu(t) + \sum_{k=1}^K \xi_{ik} \phi_k(t)$ , and the shared latent variable density function of the FPC score

$$f(\xi_i | \Lambda) = (2\pi|\Lambda|)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \xi_i^T \Lambda^{-1} \xi_i \right).$$

For example, we choose the sub-hazard density function for the sub-survival model in the application example as the following:

$$f(T_i, \Delta_i, m_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \left[ \frac{\exp \{ \boldsymbol{\gamma}_m^T \xi_i + \boldsymbol{\beta}_m^T \mathbf{Z}_i \}}{\sum_{j=1}^n R_j(T_i) w_j(T_i) \exp \{ \boldsymbol{\gamma}_m^T \xi_j + \boldsymbol{\beta}_m^T \mathbf{Z}_j \}} \right]^{I_{m_i \Delta_i = m}},$$

where

$$R_j(t) = \begin{cases} 1 & \text{if } m_i \neq m \\ I(T_i \geq t) & \text{if } m_i = m \end{cases}$$

and the weighting function

$$w_j(t) = \begin{cases} 1 & \text{if } t \leq T_i \\ 0 & \text{if } t > T_i \text{ and } \Delta_i = 0 \\ \widehat{G}(t)/\widehat{G}(T_i) & \text{if } t > T_i \text{ and } \Delta_i = 1, \end{cases}$$

where  $\widehat{G}(t)$  is the standard Kaplan-Meier estimator for the censoring distribution

$$\widehat{G}(t) = \prod_{t_i < t} \left\{ 1 - \frac{\sum_{j=1}^n I(t_j = t_i \cap \Delta_i = 0)}{\sum_{j=1}^n I(t_j \geq t_i)} \right\}.$$

According to Equation (4), the score function is found to be proportional to

$$\begin{aligned} S(\Theta) &\approx \frac{\partial (\sum_{i=1}^n \log \{ \prod_{m=1}^M \{ \int f(T_i, \Delta_i, m_i | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta}) f(\mathbf{Y}_i^*(t) | \xi_i, \sigma) f(\xi_i | \Lambda) d\xi_i \} \})}{\partial \Theta} \\ &= \sum_{i=1}^n \int \frac{\partial h(\Theta, \xi_i)}{\partial \Theta} f(\xi_i | T_i, \Delta_i, \mathbf{Y}_i^*(t), \Theta) d\xi_i, \end{aligned} \quad (5)$$

where

$$h(\Theta, \xi_i) = \log \left\{ \prod_{m=1}^M f(T_i, \Delta_{m_i} | \xi_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \boldsymbol{\beta}) f(Y_i^*(t) | \xi_i, \sigma) f(\xi_i | \Lambda) \right\}.$$

The observed data score vector in formula (5) is expressed as the expected value of the complete-data score vector with respect to the posterior distribution of the random effects of  $\xi$ . If the score equations in the formula (5) can be solved with respect to  $\Theta$ , with  $f(\xi_i | T_i, \Delta_i, Y_i^*(t), \Theta)$  fixed at the  $\Theta$  value of the previous iteration, then it is an EM algorithm. The proposed algorithm to estimate the parameters will be given in Section 3.2.

### 3.2. Parameter estimation

This section is focused on estimating the parameters in the joint-likelihood function. There are two main challenges to estimate parameters in the joint likelihood functions. One is the requirement for numerical integration of latent variables  $\xi_i$  when the dimension of random-effects increases. The other is to estimate the density function  $f(\xi_i | \Lambda)$  because we don't have a closed-form for FPC function  $\phi_k(t)$ . Therefore, we propose to use the modified two-stage algorithm to estimate the parameters, and the proposed algorithm is specified as follows:

#### (1) Stage I

- (a) Estimate all the parameters  $\hat{\mu}$ , the FPC  $\hat{\phi}_k$ , and the FPC score  $\hat{\xi}$ , and  $\hat{\sigma}^2$  as shown in Section 2.1. The vector of random effects  $\xi_i$  is shared between both longitudinal and survival sub-models. Thereby, we try to reduce the biases from the informative dropout problem for estimating the random effects parameters. Here the informative dropout means that there are missing measurements in the longitudinal trajectory  $Y_i^*(t)$ . These missing measurements can be recovered for all subjects as

$$\hat{Y}_i^*(t) = \hat{\mu}(t) + \sum_{k=1}^K \hat{\xi}_{ik} \hat{\phi}_k(t) + \epsilon_i, \quad (6)$$

where  $t$  can be any past or future time points before patient death. In this way, we can simulate the complete longitudinal measurements  $Y_i^*(t)$  in the next step.

- (b) Simulate complete measurements of the longitudinal data  $Y_i^*(t)$  from Equation (6), which is based on the estimated mean function  $\hat{\mu}(t)$ , the estimated score function  $\hat{\phi}_k(t)$ , the latent variables  $\hat{\xi}_i$  and  $\hat{\sigma}^2$ . The latent variables  $\hat{\xi}_i$  and  $\hat{\sigma}^2$  are simulated from normal distributions with means and variances estimated from the observed longitudinal data in the previous step. Using the complete longitudinal measurements, we can reestimate parameters  $\hat{\mu}(t)$ ,  $\hat{\phi}_k(t)$ ,  $\hat{\xi}_i$ , and  $\hat{\sigma}^2$  using the procedure in (a).. As  $\min(n_i) \rightarrow \infty$ , the estimated parameters in the submodel will convergence to the estimated parameters obtained from the joint model in probability as shown in Rizopoulos [16] and Huong *et al.* [9].

(2) *Stage II*

After the insertion of the fitted values from stage I, the proposed joint models become

$$\begin{cases} \hat{Y}_i^*(t) = \mu(\hat{t}) + \sum_{k=1}^K \hat{\xi}_{ik} \hat{\phi}_k(t) + \epsilon_i, \\ h_m(t | \hat{\xi}_i, \mathbf{Z}_i) = h_{m,0} \exp[\boldsymbol{\gamma}_m^T \hat{\xi}_i + \boldsymbol{\beta}_m^T \mathbf{Z}_i]. \end{cases}$$

- (a) Approximate the expected function of the complete data likelihood. After the fitted values from stage I, the full joint likelihood function is in the following form:

$$\begin{aligned} L(\Theta) = \prod_{i=1}^n \int \prod_{m=1}^M \left( \left\{ h_m[t_i | \hat{\xi}_i, \mathbf{Z}_i] \right\}^{I_{m_i \Delta_i = m}} \exp \left\{ - \int_0^\infty \sum_{i=1}^n R_i(u) \right. \right. \\ \left. \left. \times h_m(u | \hat{\xi}_i, \mathbf{Z}_i) du \right\}^{1 - I_{m_i \Delta_i = m}} \right) f(Y_i^*(t) | \hat{\xi}_i, \hat{\sigma}) f(\hat{\xi}_i | \hat{\Lambda}) d\hat{\xi}_i. \end{aligned}$$

As proven in the papers [9,16], the expected function of the complete data log-likelihood function can be approximated as in the following when  $\min(n_i) \rightarrow \infty$ ,

$$\begin{aligned} E(l(\Theta)) \approx \sum_{i=1}^n \log \left\{ \prod_{m=1}^M \left( \left\{ h_m[t_i | \hat{\xi}_i, \mathbf{Z}_i] \right\}^{I_{m_i \Delta_i = m}} \exp \left\{ - \int_0^\infty \sum_{i=1}^n R_i(u) \right. \right. \right. \\ \left. \left. \times h_m(u | \hat{\xi}_i, \mathbf{Z}_i) du \right\}^{1 - I_{m_i \Delta_i = m}} \right) \right\} + \log f(Y_i^*(t) | \hat{\xi}_i, \hat{\sigma}) \\ + \log f(\hat{\xi}_i | \hat{\Lambda}). \end{aligned} \quad (7)$$

- (b) Estimate the parameters  $\boldsymbol{\gamma}$  and  $\boldsymbol{\beta}$  for the survival model by maximizing the approximation of the expected function of the complete data log-likelihood as in the formula (7).

#### 4. Application: kidney transplant study

A total of 5654 kidney transplant recipients are included in the study from United Network for Organ Sharing (UNOS). Patients may experience kidney transplant failure, death or remain healthy until the end of their follow-up periods. Among the total 5654 patients, 1590 (28%) patients experience kidney transplant failure and 1735 (31%) patients eventually die. A total of 707 (44%) patients die after kidney failure, and 1028 (28%) patients die with a kidney still functioning. A competing risk model is adopted where the primary event of interest is the transplant failure and the competing event is death before transplant failure from other causes. We use the first four FPC scores computed from the functional principal component analysis on the patients' longitudinal GFR trajectories as covariates in the model. In addition, we included additional prognostic covariates such as age, sex,

**Table 1.** Summary of Kidney transplanted recipient characteristics in some kidney transplant data.

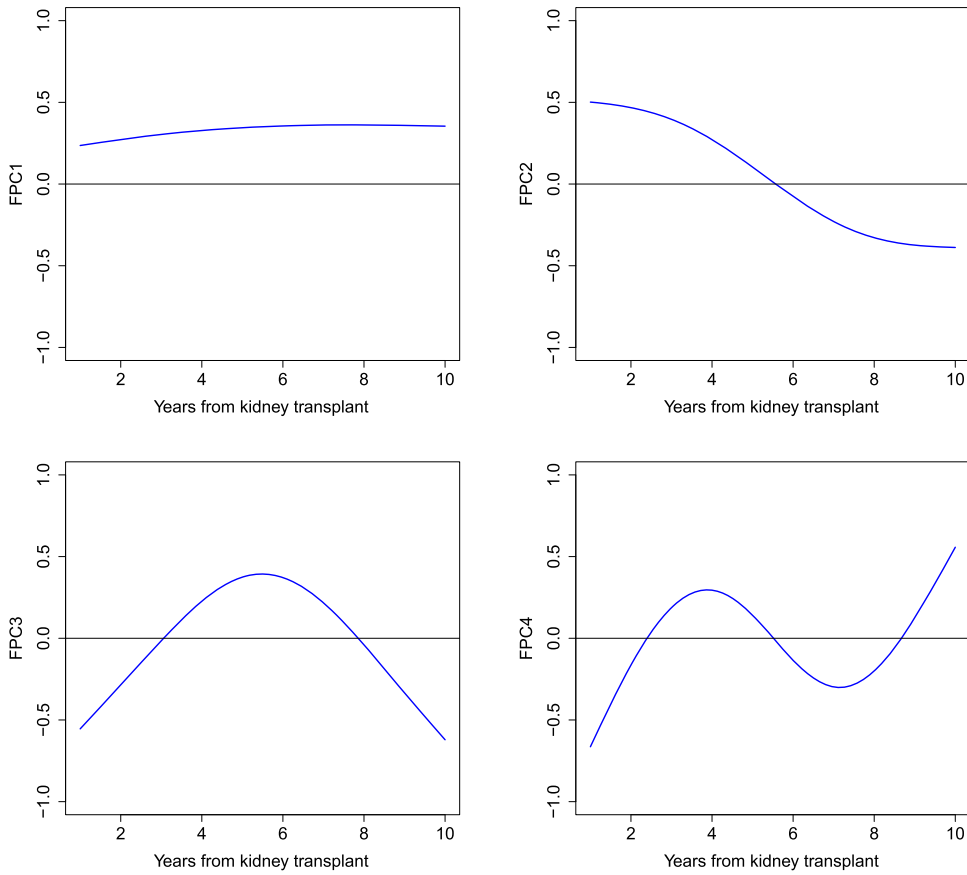
Patient characteristics variables	Percentage (%)
Age	
18–39	48
40–59	39
≥ 60	13
Sex	
Male	59
Female	41
Race	
White	61
Black	31
Other	8
Cause of ESRD	
Diabetes	32
Hypertension	21
Glomerular disease	29
Polycystic disease	9
Other	9
Kidney donor type	
Deceased	78
Living	22

race, cause of end-stage renal disease (ESRD) and kidney donor type. The detailed patient demographics are shown in Table 1.

#### 4.1. Results: longitudinal submodel

We choose  $K = 4$  based on the Akaike Information Criterion (AIC) that takes into account the joint likelihood of longitudinal and survival models to select the number of principal components to be extracted. From functional principal components analysis, the 4 leading functional principal components ( $K = 4$ ) account for 99.8% of the total variability of the GFR curves; the first, second, third, and fourth FPCs respectively account for 84.6%, 10.6%, 3.4%, and 1.1%. Although the first 2 FPCs account for 95.24% of the total variation, the third and fourth ones also contain important information that has potential predictive value.

The first four functional principal components are as shown in Figure 2. The patterns of the four leading functional principal component curves are similar to Figure 3 in Dong *et al.* [4], which are included in the Appendix. However, the new curves are more robust and smooth compared with Figure 3 in Dong *et al.* [4]. The reason is that the joint model can recover missing longitudinal GFR from survival data when modelling the longitudinal and survival data together. These functional principal component curves have a straightforward explanation. For example, the first FPC stays flat during the entire follow-up period, indicating that the largest GFR variation between subjects is the distance of a subject's GFR curve to the mean GFR curve. In other words, the degree of how far the curve is from the mean GFR curve captures the largest variation of the data, and the majority of patients have a relatively stable trajectory. The second FPC represents the ascent/decline in the GFR curve. The third and fourth FPCs represent a higher degree of fluctuation in the GFR curves.



**Figure 2.** The first four leading functional principal components that estimated from the observed longitudinal GFR data and the recovered GFR data by incorporating survival data information when jointly modelling the longitudinal and survival outcomes.

#### 4.2. Results: competing risk submodel

Table 2 show the hazard ratios of the covariates in the proposed joint model with  $m = 2$ . All of the 4 FPC scores are statistically significant for both death and kidney failure outcomes. The effects of these FPC scores on the hazard are different. For example, the first FPC score has a negative effect on the hazard while the second FPC score has a positive one. The clinical interpretations of the effects of FPC scores are as follows. The first FPC score depicts the main level of the GFR curve, i.e. how far it deviates from the mean. The level of GFR is related to the state of kidney function, as higher GFR values indicate better kidney health. A patient with a larger first FPC score has a higher GFR level and thus less likely to suffer from kidney failure, whereas the one with a smaller first FPC score (and thus lower GFR level) may be subject to a higher risk of kidney failure. Similarly, the second FPC score relates to the degree of decline in the GFR curve. A larger second FPC score indicates a steeper decline and is thus associated with a higher hazard of kidney failure. The third and fourth FPC scores are also significantly related to the event of interest; they represent the degree of abnormal fluctuation within the curve. The third and fourth FPC scores can be used as a guide to identifying those abnormal patients with highly fluctuating

**Table 2.** Estimated hazard ratios of kidney failure post kidney transplant ( $m = 1$ ) in the joint model with different survival submodels with 95% confidence interval given in brackets.

Covariates	Joint model			
	Competing risk submodel		Cox submodel	
	Hazard ratio	p-value	Hazard ratio	p-value
Age				
18–39	1.00		1.00	
40–59	0.68(0.61, 0.76)	< 0.001	1.61(1.35, 1.92)	0.001
≥ 60	0.47(0.38, 0.57)	< 0.001	1.49(1.15, 1.93)	0.001
Sex				
Male	1.00		1.00	
Female	0.89(0.80, 0.99)	0.048	0.78(0.66, 0.91)	0.038
Race				
White	1.00		1.00	
Black	1.39(1.23, 1.55)	< 0.001	1.37(1.16, 1.62)	< 0.001
Other	0.82(0.66, 1.02)	0.079	0.63(0.44, 0.89)	< 0.001
Cause of ESRD				
Diabetes	1.00		1.00	
Hypertension	0.89(0.74, 0.97)	0.026	0.74(0.61, 0.91)	< 0.001
Glomerular disease	0.99(0.85, 1.14)	0.834	0.55(0.44, 0.68)	< 0.001
Polycystic disease	0.76(0.60, 0.97)	0.026	0.44(0.31, 0.62)	< 0.001
Other	0.90(0.76, 1.07)	0.221	0.56(0.44, 0.71)	< 0.001
Kidney donor type				
Deceased	1.00		1.00	
Living	0.90(0.79, 0.99)	0.048	0.84(0.68, 1.02)	0.084
FPC score				
First FPC score	0.981(0.979, 0.982)	< 0.001	0.968(0.965, 0.970)	< 0.001
Second FPC score	1.009(1.005, 1.013)	< 0.001	1.003(1.000, 1.005)	0.001
Third FPC score	0.976(0.969, 0.983)	< 0.001	0.964(0.953, 0.975)	< 0.001
Fourth FPC score	0.993(0.974, 1.000)	0.050	0.972(0.946, 0.999)	0.048

GFR curves. The significance of the FPC scores indicates that the conventional proposal in the literature [12,13] to assume that the change of the GFR curves is primarily linear might result in an incomplete depiction of the variation within the GFR curves as well as its effects. Such a simplified assumption may lead to biased findings.

As a comparison, Table 2 also displays estimation results under the Cox model. The Cox model attempts to answer a different question from the competing risks model. Under a Cox model, all patients who die before kidney failure are regarded as censored. The effect identified in the Cox model can be interpreted as the association between a covariate and the hazard rate of kidney failure in patients who have not experienced either kidney failure or death. On the other hand, under a competing risks model, the covariate effect is interpreted as the degree of association to the instantaneous rate of kidney failure given that the patient has either been healthy and never experienced kidney failure, or has died and could never possibly have kidney failure. The issue addressed by the Cox model refers to the risks of kidney failure that could be expected if a patient lives long enough, instead of, as with the competing risks model, the actual risk of kidney failure. The Cox model might be more suitable for understanding the etiological association between the covariates and the event of kidney failure; whereas the competing risks model is more relevant for prediction and allocation of resources as it focuses on the actual risk of the event occurring.

Regarding the effect of age, we observe an interesting pattern. Specifically, under the Cox model, the older age group has a hazard ratio greater than 1, indicating that from

an etiological perspective, older patients are more likely to experience kidney failure. On the contrary, under the competing risks model, the hazard ratio of the older age group is smaller than 1, i.e. relative to the age group of 18 to 35, patients who are in age groups of 40 to 59, and 60 or above have hazard ratios equal to 0.68 and 0.47, respectively. This is possibly due to the fact that young patients are more likely to have kidney failure before death because they tend to live long enough until the occurrence of kidney failure, while old patients are less likely to experience a kidney failure because more often than not, death occurs before kidney failure.

The relationships of the time-to-event outcomes with the rest of the covariates have reasonable clinical interpretations. For example, female patients are less likely to experience a kidney failure event compared with male patients. Compared with patients whose donor is deceased, patients who have a living donor transplant are less likely to experience kidney failure.

## 5. Simulation study

We evaluate the estimation accuracy of our joint model under simulated scenarios. As the primary focus is the functional component of the joint model, for simplicity, we assume that there are no baseline covariates  $\mathbf{Z}_i$ . A FPCA model with no covariates is used to simulate the longitudinal trajectories,

$$Y_i(t) = \mu(t) + \sum_{k=1}^K \xi_{ik} \phi_k(t) + \epsilon_i,$$

where  $K = 4$  and the mean function  $\mu(t)$  and the eigenfunctions  $\phi_k(t)$  are given as the ones estimated from the kidney transplant data. The measurement error  $\epsilon_i$  are assumed to follow  $N(0, 0.85)$ , and the FPC scores follow the distributions  $\xi_{ik} \sim N(0, \sigma_k^2)$ , where  $k = 1, \dots, 4$ , and  $\sigma_1 = 16$ ,  $\sigma_2 = 8$ ,  $\sigma_3 = 4$ ,  $\sigma_4 = 1$ . For each subject, the scheduled repeated measurement times are set at a grid sequence of  $(1, \dots, T_i)$ , where  $T_i$  is the event time. Each cohort has a maximum follow-up time of 10. The time to the primary event is assumed to follow a log-normal distribution, and the time to the competing risk event is assumed to follow a Weibull distribution. The hazard for the time-to-event  $T_i$  is specified as

$$h_m(t | \xi_i) = h_m^0(t) \exp(\gamma_{m1}\xi_{i1} + \gamma_{m2}\xi_{i2} + \gamma_{m3}\xi_{i3} + \gamma_{m4}\xi_{i4}), \quad m = 1, 2.$$

The censoring times were simulated from a uniform distribution in the interval  $(0, 10)$ , the censoring rate about 40%, and the number of repeated longitudinal measurements per person is greater than 3. We generate data cohorts in two scenarios where  $(\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14})$  and  $(\gamma_{21}, \gamma_{22}, \gamma_{23}, \gamma_{24})$  take different values. As the estimation accuracies of the two event types are similar, only results for the estimation for the primary competing risk, i.e.  $(\gamma_{11}, \gamma_{12}, \gamma_{13}, \gamma_{14})$ , are presented. Table 3 displays the estimates, standard errors, and coverage probabilities for the two scenarios. It is evident that the estimation is quite accurate as the estimates are close to the true values. The empirical standard deviations of the estimates are also close to the average standard errors, corroborating the correctness in our estimation. The empirical coverage probabilities are slightly lower than the nominal levels of 95% and 99%; this could be due to the stage-wise nature of our estimation procedure. In



**Table 3.** Means, empirical standard deviations (SD), average standard errors (SE), and the 95% and 99% empirical coverages in two different scenarios. Each scenario has 100 simulation replicates and 100 subjects in each replicate.

Parameters	$\gamma_{11}$	$\gamma_{12}$	$\gamma_{13}$	$\gamma_{14}$
Scenario 1				
True value	1.000	1.000	1.000	1.000
Estimate	1.014	0.986	1.021	0.974
Empirical SD	0.127	0.212	0.131	0.132
Average SE	0.121	0.208	0.127	0.128
Empirical coverage(95%)	0.932	0.946	0.938	0.937
Empirical coverage(99%)	0.974	0.986	0.978	0.973
Scenario 2				
True value	−1.000	0.850	−0.750	0.500
Estimate	−0.985	0.844	−0.739	0.484
Empirical SD	0.131	0.112	0.196	0.114
Average SE	0.128	0.107	0.189	0.105
Empirical coverage(95%)	0.935	0.946	0.941	0.941
Empirical coverage(99%)	0.982	0.992	0.981	0.987

general, the model is proven to render satisfactory performance in terms of the estimation accuracy.

## 6. Conclusion and discussion

This paper proposed a joint model that includes a longitudinal FPCA submodel and a competing risk submodel, with shared latent functional features. The competing risk survival submodel can incorporate hazard ratios of multiple time-to-event outcomes. We have demonstrated the usefulness and applicability of the proposed joint model on a real kidney transplant data set. The main results from the application reveal meaningful clinical findings. The finite sample performance of the proposed method is verified in the simulation study.

One possible direction of future research relates to the stage-wise approach we adopted to estimate the coefficients. Compared with the regular approach based on seeking the maximizer of the complete data likelihood via the expectation-maximization algorithm, the stage-wise approach tremendously reduces the computational burden as it no longer requires computing integral via quadratures. Therefore, the stage-wise procedure is one of common approaches for joint models, and the improvements of stage-wise approaches for joint models have been developed in the literature [1,2,9,20,21]. Ye and Wu [20] and Huong *et al.* [9] have proved the asymptotic equivalence between the stage-wise estimates and the estimates obtained from the complete data likelihood and expectation-maximization algorithm. Possible improvement in parameter estimation might be achieved in the modified stage-wise EM approach in the future.

## Acknowledgements

We acknowledge the great support, professional advice, and constructive comments from the Editor-in-Chief Jie Chen, anonymous associate Editor, and three anonymous reviewers. We are also indebted to thank Rahul Unni and other persons for this great editorial assistance.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Jianghu (James) Dong  <http://orcid.org/0000-0001-6579-7442>

Haolun Shi  <http://orcid.org/0000-0002-3620-9717>

Liangliang Wang  <http://orcid.org/0000-0002-8509-7985>

Ying Zhang  <http://orcid.org/0000-0002-9680-8981>

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

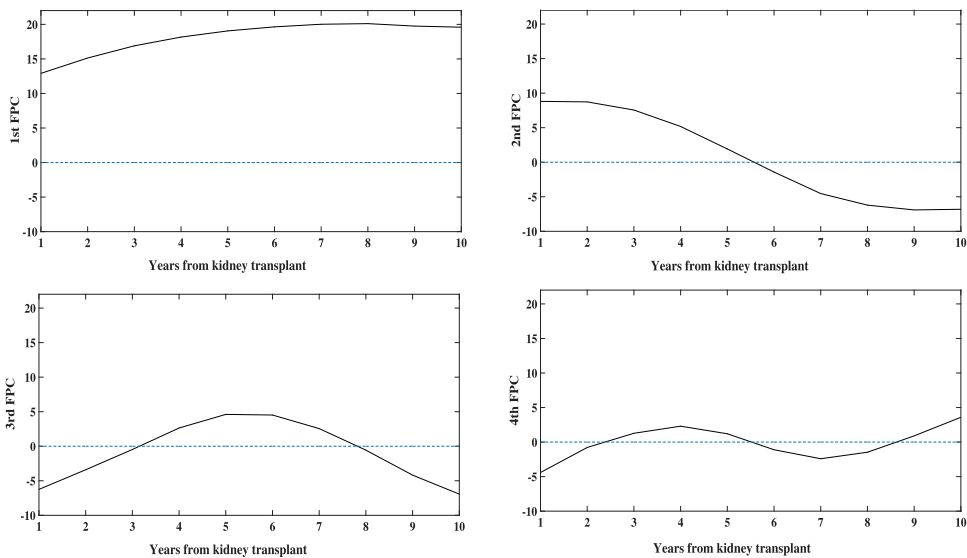
## References

- [1] P.S. Albert and J.H. Shih, *On estimating the relationship between longitudinal measurements and time-to-event data using a simple two-stage procedure*, *Biometrics* 63 (2008), pp. 983–987.
- [2] P.S. Albert and J.H. Shih, *An approach for jointly modeling multivariate longitudinal measurements and discrete time-to-event data*, *Ann. Appl. Statist.* 4 (2010), pp. 1517–1532.
- [3] J. Ding and J. Wang, *Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data*, *Biometrics* 64 (2007), pp. 546–556.
- [4] J. Dong, L. Wang, J. Gill, and J. Cao, *Functional principal component analysis of GFR curves after kidney transplant*, *Stat. Methods Med. Res.* 27 (2018), pp. 3785–3796.
- [5] J. Dong, S. Wang, L. Wang, J. Gill, and J. Cao, *Joint modelling for organ transplantation outcomes for patients with diabetes and the end-stage renal disease*, *Stat. Methods Med. Res.* 28 (2019), pp. 2724–2737.
- [6] J. Fan and I. Gijbels, *Local Polynomial Modelling and its Applications*, London: CRC Press, 1996.
- [7] J.P. Fine and R.J. Gray, *A proportional hazards model for the subdistribution of a competing risk*, *J. Am. Stat. Assoc.* 94 (1999), pp. 496–509.
- [8] G. Hickey, P. Philipson, A. Jorgensen, and R. Kolamunnage-Dona, *A comparison of joint models for longitudinal and competing risks data, with application to an epilepsy drug randomized controlled trial*, *J. R. Soc. A* 181 (2018), pp. 1105–1123.
- [9] P. Huong, D. Nur, H. Pham, and A. Branford, *A modified two-stage approach for joint modelling of longitudinal and time-to-event data*, *J. Stat. Comput. Simul.* 88 (2018), pp. 3379–3398.
- [10] G.M. James, T.J. Hastie, and C.A. Sugar, *Principal component models for sparse functional data*, *Biometrika* 87 (2010), pp. 587–602.
- [11] A.S. Levey, L.A. Stevens, C.H. Schmid, Y.L. Zhang, A.F. Castro, H.I. Feldman, J.W. Kusek, P. Eggers, F. Van Lente, T. Greene, and J. Coresh, *A new equation to estimate glomerular filtration rate*, *Ann. Intern. Med.* 150 (2009), pp. 604–612.
- [12] R. Marcen, J.M. Morales, A. Fernandez-Rodriguez, L. Capdevila, L. Pallardo, J.J. Plaza, J.J. Cubero, J.M. Puig, A. Sanchez-Fructuoso, M. Arias, G. Alperovich, and D. Seron, *Long-term graft function changes in kidney transplant recipients*, *Nephrol. Dial. Transplant.* 9 (2010), pp. ii2–ii8.
- [13] O. Moranne, N. Maillardb, C. Fafina, L. Thibaudinb, E. Alamartineb, and C. Mariatb, *Rate of renal graft function decline after one year is a strong predictor of all-cause mortality*, *Am. J. Transplant.* 13 (2013), pp. 695–706.
- [14] R. Prentice, J. Kalbfleisch, A. Peterson, N. Flournoy, V. Farewell, and N. Breslow, *The analysis of failure times in the presence of competing risks*, *Biometrika* 34 (1978), pp. 541–554.
- [15] H. Putter, M. Fiocco, and R.B. Geskus, *Tutorial in biostatistics: competing risks and multistate models*, *Stat. Med.* 26 (2007), pp. 2389–2430.
- [16] D. Rizopoulos, *Fast fitting of joint models for longitudinal and event time data using a pseudoadaptive Gaussian quadrature rule*, *Comput. Stat. Data Anal.* 56 (2011), pp. 2061–2077.
- [17] R.A. Wolfe, V.B. Ashby, E.L. Milford, A.O. Ojo, R.E. Ettenger, L.Y. Agodoa, P.J. Held, and E.K. Port, *Comparison of mortality in all patients on dialysis, patients on dialysis awaiting*

- transplantation, and recipients of first cadaveric transplant*, New Engl. J. Med. 341 (1999), pp. 1725–1730.
- [18] F. Yao, *Functional principal component analysis for longitudinal and survival data*, Stat. Sin. 17 (2007), pp. 965–983.
- [19] F. Yao, H.G. Muller, and J.L. Wang, *Functional data analysis for sparse longitudinal data*, J. Am. Stat. Assoc. 100 (2005), pp. 577–590.
- [20] Q. Ye and L. Wu, *Two-step and likelihood methods for joint models of longitudinal and survival data*, J. Stat. Comput. Simul. 46 (2017), pp. 0361–0918.
- [21] W. Ye, X. Lin, and J.M.G. Taylor, *Semiparametric modeling of longitudinal measurements and time-to-event data—A two-stage regression calibration approach*, Biometrics 64 (2008), pp. 1238–1246.

## Appendix

For comparison with the patterns of the four leading FPCs, Figure 3 in Dong *et al.* [4] is copied here for the convenience of the readers.



**Figure A1.** The first four leading functional principal components (FPCs) that estimated from functional principal components analysis in the paper [4].