Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
2019, Vol. 28(9) 2724–2737

DOI: 10.1177/0962280218786980
journals.sagepub.com/home/smm
SAGE

# Joint modelling for organ transplantation outcomes for patients with diabetes and the end-stage renal disease

**Jianghu (James) Dong,[1] Shijia Wang,[1] Liangliang Wang,[1] Jagbir Gill[2] and Jiguo Cao[1]**

**Abstract**

This article is motivated by jointly modelling longitudinal and time-to-event clinical data of patients with diabetes and end-stage renal disease. All patients are on the waiting list for the pancreas transplant after kidney transplant, and some of them have a pancreas transplant before kidney transplant failure or death. Scant literature has studied the dynamical joint relationship of the estimated glomerular filtration rates trajectory, the effect of pancreas transplant, and time-to-event outcomes, although it remains an important clinical question. In an attempt to describe the association in the multiple outcomes, we propose a new joint model with a longitudinal submodel and an accelerated failure time submodel, which are linked by some latent variables. The accelerated failure time submodel is used to determine the relationship of the time-to-event outcome with all predictors. In addition, the piecewise linear function in the survival submodel is used to calculate the dynamic hazard ratio curve of a time-dependent side event, because the effect of the side event on the time-to-event outcome is non-proportional. The model parameters are estimated with a Monte Carlo EM algorithm. The finite sample performance of the proposed method is investigated in simulation studies. Our method is demonstrated by fitting the joint model for the clinical data of 13,635 patients with diabetes and the end-stage renal disease.

**Keywords**

Jointly modelling, longitudinal data, survival analysis, organ transplantation, glomerular filtration rates

## 1 Introduction

This article is motivated by a longitudinal and time-to-event clinical dataset, where all patients have the end-stage renal disease (ESRD) and diabetes. All patients have already had a kidney transplantation from a living or decreased kidney donor, and all of them are on the waiting list for the pancreas transplant to treat the diabetes disease. Partial patients have a pancreas transplantation if a matched pancreas organ is available for them. It is known that the organ transplantation can prolong the survival of type 1 diabetic patients with ESRD.[1–7] However, how to extend the long-term kidney function still remains the main challenge for transplantation.

No studies have used a joint model to predict the long-term kidney function, which includes the longitudinal continuous outcome of glomerular filtration rate (GFR) and the time-to-event outcome of all-cause graft loss (ACGL). From our preliminary result as in Figure 1, the levels of the observed GFR trajectories for patients with ACGL are higher than those of patients without ACGL, and we find that the slopes of GFR trajectories for patients with ACGL are steeper in comparison with patients without ACGL. In addition, all patients may have a pancreas transplant at any time post kidney transplant to treat the diabetic disease, and they are in different statuses at different time points during the follow-up period. For example, they are in status 1 (alive without pancreas transplant) at the time of the admission to the waiting list for a pancreas. They move to status 2

[1]Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, Canada
[2]Department of Medicine, University of British Columbia, Vancouver, Canada

**Corresponding author:**
Jiguo Cao, Department of Statistics and Actuarial Science, Simon Fraser University, Room SC K10545, 8888 University Drive, Burnaby V5A 1S6, BC, Canada.
Email: jca76@sfu.ca

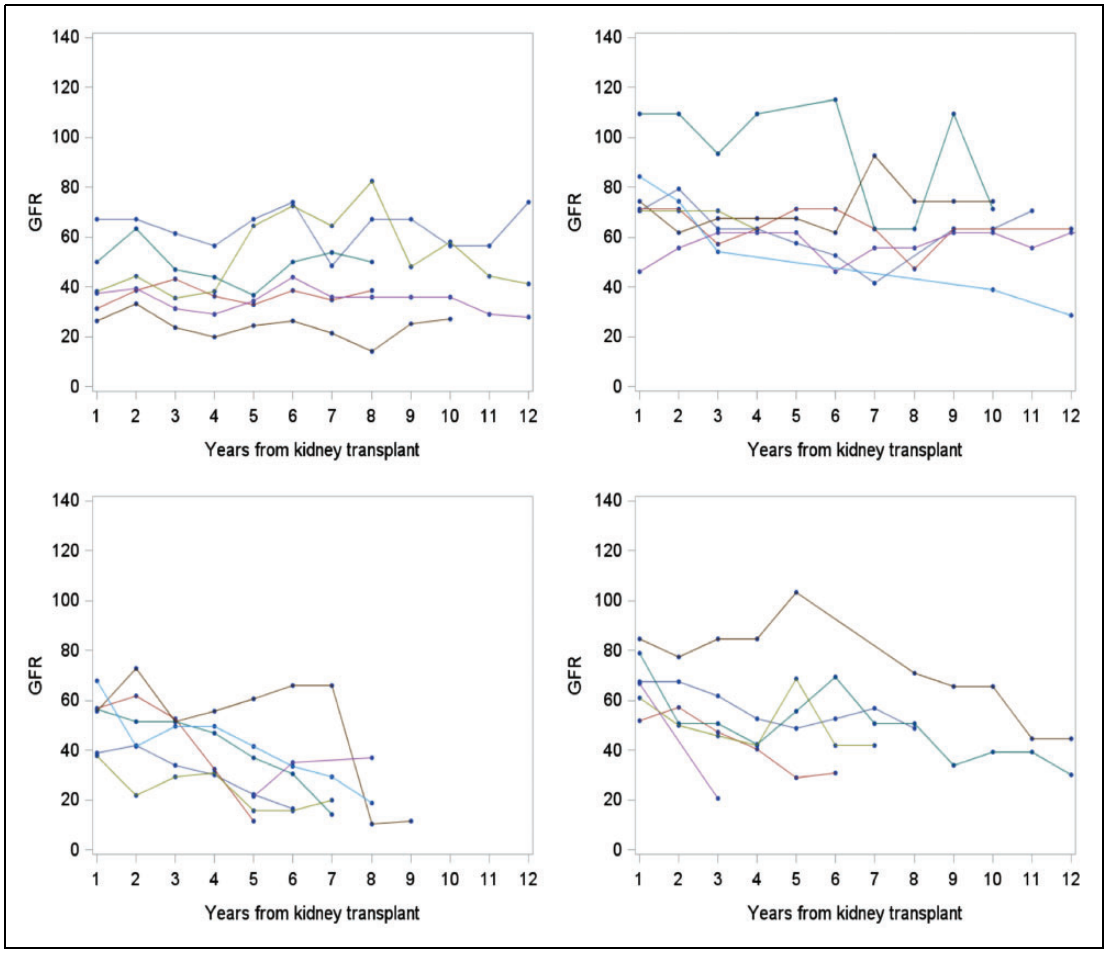


**Figure 1.** Observed individual GFR trajectory curves. The left two panels, from top to bottom, are GFR curves for patients with All-cause graft loss (ACGL) events or without ACGL events, respectively, when they do not have a pancreas transplantation. The right two panels, from top to bottom, are GFR curves for patients with ACGL event or without ACGL events, respectively, when they have a pancreas transplantation. Each color represents the individual patient in each panel.

(alive with pancreas transplant) if a matched pancreas organ becomes available for them before ACGL, or they directly move to status 3 with ACGL. In fact, the failure rates of the time-to-event outcome are different as patients change their status. As shown in Figure 1, it seems that patients who have a pancreas transplant are less likely to have ACGL. In short, the above scenarios motivate us to develop a new dynamical joint model to predict the long-term outcome, since there are at least two advantages in using a joint model. First, joint modelling multiple outcomes together can increase the power and decrease the Type I error.[8,9] Second, this joint model can estimate the parameters in the longitudinal component by incorporating the time-to-event information through censoring, and similarly, for the converse situation, the estimation of the time-to-event is incorporated by the longitudinal data information.

Several methods for estimating joint models of multiple outcomes have been developed. The main challenge in jointly modelling multiple outcomes is the lack of a suitable multivariate joint distribution. The first approach is a two-stage approach,[10] where a random components model is developed to describe repeated longitudinal measures in the first stage, and a Cox proportional hazards model is estimated in the second stage. However, this approach may cause bias when the observation of the longitudinal process is interrupted by the event. To address this problem, the second approach[11] directly specifies the joint distribution by factorizing it into the conditional distribution of one outcome and a marginal distribution of the other outcome. This approach was reviewed with some insightful comments.[12] The accelerated failure time model is considered in their joint model rather than the Cox proportional hazards model.[13,14] The third approach directly formulates a joint model for longitudinal repeated measurements and the time-to-event outcome. For instance, a copula is used to construct

the joint distribution.[15] Another challenge in jointly modelling multiple outcomes is the intensive computation due to the complex correlation structure of latent variables and measurement errors in covariates.[11,13] So the EM or the Monte Carlo EM algorithm is developed to estimate parameters in the joint models.[13,14] The Bayesian method is also developed to estimate the joint models.[16]

Most of above joint models are based on the Cox proportional hazard regression, and only a few joint models such as Tseng et al. [14] use the accelerated failure time regression. As shown in Figure 3, the assumption of Cox proportional hazard model fails because the cumulative survival lines cross with each other for patients with/without a pancreas transplant. Therefore, the accelerated failure time submodel is used in our proposed joint model. On the other hand, the proposed joint model is different from the joint models in Tseng et al.,[14] which treats the longitudinal component as a covariate in the survival analysis. In our proposed joint model, instead of using the whole longitudinal component as a covariate, we propose to use some latent features of the longitudinal component in the survival submodel. Finally, none of the above joint models has considered a method to obtain the dynamical non-proportional hazard ratio curve of a side event when hazard ratios are non-proportional during the followed-up time period.

The main contribution of this paper is that we review the clinical question in the transplantation data, and accordingly develop a new joint model to determine the relationships of multiple outcomes to account for correlations within/between subjects. To the best of our knowledge, it is the first time for the joint model to be applied to the organ transplantation research. Our proposed joint model has three advantages. First, the survival submodel shares a vector of latent variables with the longitudinal submodel. The advantages of this model are that unnecessary noise can be filtered, and the effects of other covariates can be adjusted in the longitudinal submodel. In addition, it is easy to interpret the coefficients from the model results. For example, the latent features are the baseline and the slope of GFR trajectories in the application example. The coefficients in the survival component represent their corresponding relationship with the time-to-event outcome. Second, the survival submodel shares the data information together with the longitudinal submodel. For example, our proposed joint model in the application example has considered that the occurrence of death or transplant failure may lead to the censoring of GFR, which overcomes the drawback of separate analyses for each outcome. Finally, our proposed joint model includes a piecewise linear function to display the dynamical non-proportional hazard ratios of the side event on the time-to-event outcome.

The rest of this article is organized as follows. Our proposed joint models are introduced in Section 2. We present our estimation method for the joint model in Section 3. Section 4 demonstrates the application of our joint model in the transplantation clinical data. Section 5 presents three simulation studies to investigate the finite sample performance of our joint model. Conclusions and discussion are given in Section 6.

## 2 The joint model

Let $Y_i(t_{ij})$ be a repeated continuous measured outcome at times $t_{ij}$ for the $i$-th subject, where $i = 1, \ldots, n$, $j = 1, \ldots, m_i$, and $m_i$ is the number of repeated measurements for the $i$-th subject. For example, the longitudinal outcome $Y_i(t_{ij})$ is the repeated measurements of GFR at different time points in the application example of transplant clinical data. Let $T_i$ be the $i$-th subject's survival time to the event of interest, $C_i$ be a possible censoring time, $\delta_i = 1_{\{T_i \le C_i\}}$ be the censoring indicator, $S_i = \min(T_i, C_i)$ be the observed survival time, and $\mathbf{Z}_i = [Z_{i1}, \ldots, Z_{iP}]^T$ be the observed covariates for the $i$-th subject. We propose the following joint model

$$\begin{cases} Y_i(t_{ij}) = \boldsymbol{\alpha}^T \mathbf{Z}_i + \boldsymbol{\beta}_i^T \boldsymbol{\xi}(t_{ij}) + \epsilon_i, \quad i = 1, \ldots, n, \\ \lambda(t|\mathbf{Z}_i, \boldsymbol{\beta}_i, \mathbf{w}_i(t)) = \lambda_0\{\int_0^t \phi(s, \mathbf{Z}_i, \mathbf{w}_i(s), \boldsymbol{\beta}_i, \boldsymbol{\gamma})\mathrm{d}s\}\phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) \end{cases} \tag{1}$$

The first equation in the joint model (1) is the longitudinal submodel for repeated measurement outcome $Y_i(t_{ij})$, where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_P)^T$ is a vector of coefficients for the fixed effects of $\mathbf{Z}_i = [Z_{i1}, \ldots, Z_{iP}]^T$, and $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{iL})^T$ is a vector of coefficients for the random effects of $\boldsymbol{\xi}(t) = (\xi_1(t), \ldots, \xi_L(t))^T$. Here, $\xi_\ell(t), \ell = 1, \ldots, L$ is a parametric function of $t$. For example, $\xi_1(t) = 1$ and $\xi_2(t) = t$ in our application example. We assume that $\boldsymbol{\beta}_i \sim \text{Normal}(\mathbf{b}, \mathbf{B})$. The vector of measurement errors $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{im_i})$ are assumed to be multivariate normal distributed with the mean $\mathbf{0}$ and the variance–covariance matrix $\sigma^2 \mathbf{I}_n$.

The second equation in the joint model (1) is the survival sub-model with the accelerated failure time hazard function, where $\phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) = \exp[\boldsymbol{\gamma}_1^T \mathbf{Z}_i + \boldsymbol{\gamma}_2^T \boldsymbol{\beta}_i + \mathbf{w}_i(t|\boldsymbol{\gamma}_3)]$, which represent the joint effects of covariates, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^T, \boldsymbol{\gamma}_2^T, \boldsymbol{\gamma}_3^T)$ are the coefficients in the survival model, and $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{iL})^T$ are the vector of latent variables, which are shared in the longitudinal sub-model and the survival sub-model. The time-dependent

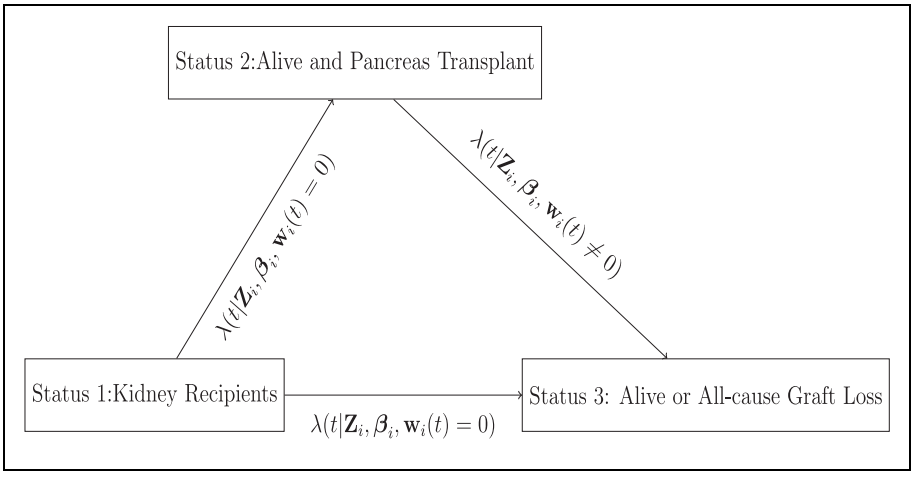


**Figure 2.** The three statuses of kidney transplant patients. All patients start from the date of the kidney transplant (Status 1), then they may move to Status 2 (pancreas transplantation) when a matched pancreas organ is available during the followed-up time period. If not, they directly move to Status 3 when the time-to-event outcome of all-cause graft loss happens, or they still are on the waiting-list for the pancreas transplant.

indicator function $\mathbf{w}_i(t|\gamma_3)$ captures the dynamic relative risk of the side event at different time points post the side event. Here, $\lambda_0\{\int_0^t \phi(s, \mathbf{Z}_i, \mathbf{w}_i(s), \boldsymbol{\beta}_i, \boldsymbol{\gamma})\mathrm{d}s\}$ is the baseline hazard function. This survival sub-model is justified with more details in section 2.1.

This proposed joint model has two major advantages. First, the mixed-effect submodel of the longitudinal outcome can adjust for other co-variables to filter some noise because we do not treat the longitudinal outcome as a covariate in the survival submodel. Second, the survival submodel shares the latent features $\boldsymbol{\beta}_i$ with the mixed-effect submodel. The estimates of the latent features in the joint model can offer an answer for specific clinical questions. For example, in our kidney transplant application, the latent variable $\beta_{i1}$ is the baseline of GFR, the latent variable $\beta_{i2}$ is the slope of GFR. Their corresponding coefficients ($\gamma_{21}$ and $\gamma_{22}$) in the survival model show the effect of the baseline and the slope of GFR to the time-to-event outcome.

## 2.1 The survival submodel

This section provides the justification for the survival submodel in the proposed joint model with more details. In the transplant clinical data, all subjects have a kidney transplant, and only part of them have a pancreas transplant at a certain time after kidney transplant before death.

From our preliminary analysis, the clinical transplant data has several aspects. First, as shown in Figure 1, patients with a pancreas transplant are less likely to have the time-to-event outcome in comparison with patients without a pancreas transplant. Second, patients have a dynamical status as shown in Figure 2. For instance, each individual is on the waiting-list program for the pancreas transplantation after kidney transplant (status 1). Then patients either move to status 2 (alive and pancreas transplant) when a matched pancreas organ is available, or they directly move to status 3 (ACGL or on the waiting). The hazard rates are different when moving from status 1 to status 3 in comparison with the other scenario when moving from status 2 to status 3. Third, the assumption of Cox proportional hazard model fails as shown in Figure 3, because the cumulative survival line of patients with a pancreas transplant cross with the cumulative survival line of patients who have no pancreas transplant. Therefore, we recommend the alternative hazard model rather than Cox proportional hazards model in this paper.

We propose to use the accelerated failure time (AFT) model, which was first introduced by Cox[17] to determine whether the effect of a covariate is to accelerate or decelerate the life course of a disease by some constant. Cox and Oakes[18] extended an AFT model with time-dependent covariates. James[19] provided a method to estimate the time-dependent AFT model in the presence of confounding factors. Reid[20] and Kay[21] mentioned that the AFT models were more appealing than the proportional hazard models because they could give direct physical interpretations. For example, as mentioned in the paper by Kay,[21] the AFT model can supply a more straightforward interpretation of the treatment effect on the time-to-event data because the coefficient of the treatment indicator can be estimated across various intervals defined by the cut time points from the date of treatment.

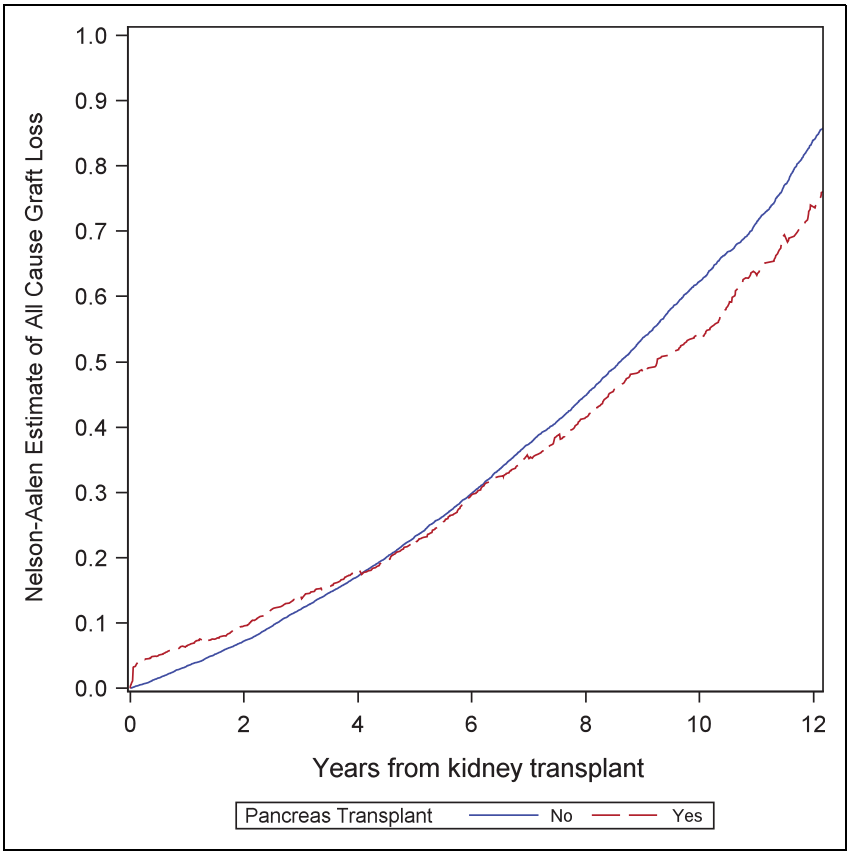


**Figure 3.** The cumulative Nelson–Aalen estimate of all-cause graft loss by patient status of pancreas transplantation. The red line is the cumulative Nelson–Aalen estimate of all-cause graft loss for patient with a pancreas transplant and the blue line is the cumulative Nelson–Aalen estimate of all-cause graft loss for patient without a pancreas transplant.

In this paper, the hazard function of the accelerated failure time submodel is specified as

$$\lambda(t|\mathbf{Z}_i, \boldsymbol{\beta}_i, \mathbf{w}_i(t)) = \lambda_0 \left\{ \int_0^t \phi(s, \mathbf{Z}_i, \mathbf{w}_i(s), \boldsymbol{\beta}_i, \gamma) \mathrm{d}s \right\} \phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \gamma) \tag{2}$$

where $\phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \gamma) = \exp[\boldsymbol{\gamma}_1^T \mathbf{Z}_i + \boldsymbol{\gamma}_2^T \boldsymbol{\beta}_i + \mathbf{w}_i(t)]$, $\mathbf{Z}_i = [Z_{i1}, \ldots, Z_{iP}]^T$ is a vector of time-independent covariates, $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{iL})^T$ is a vector of latent variables, which are random factors shared with the longitudinal sub-model, and $\mathbf{w}_i(t)$ is a time-dependent indicator function. We assume $\mathbf{w}_i(t)$ to be a piecewise linear function, which is used to capture the dynamic relative risk at different time points post the side event. We express $\mathbf{w}_i(t)$ as

$$\mathbf{w}_i(t|\boldsymbol{\gamma}_3) = \begin{cases} 0 & \text{if } t \le W_i, \\ \gamma_{30} & \text{if } W_i < t \le W_i + \frac{D_0}{365}, \\ \gamma_{3k}\left(t - W_i - \frac{D_{k-1}}{365}\right) & \text{if } W_i + \frac{D_{k-1}}{365} < t \le W_i + \frac{D_k}{365}, \quad k = 1, \ldots, K, \\ \gamma_{3(K+1)} & \text{if } t > W_i + \frac{D_K}{365} \end{cases} \tag{3}$$

where $W_i$ is the time of the side event for the $i$-th subject, $D_0, D_1, \ldots, D_K$ are denoted as the specified number of days post the side event, and $\boldsymbol{\gamma}_3 = (\gamma_{30}, \ldots, \gamma_{3(K+1)})^T$ are the coefficients in the piecewise function. In our application example of the clinical transplant data, $W_i$ is the time from the kidney transplant to the pancreas transplant for those patients who have pancreas transplant. For patients without pancreas transplant, $W_i$ is set to be larger than the end date of the study cohort minus the date of kidney transplant.

It is worth mentioning that the hazard function defined in equation (2) can be a strata hazard model because it can provide different hazard functions when patients are in different statuses. For example, the hazard function $\lambda(t|\mathbf{Z}_i, \boldsymbol{\beta}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3) = 0)$ is the hazard rate when patients move from status 1 to status 3 if $t \leq W_i$. The hazard function $\lambda(t|\mathbf{Z}_i, \boldsymbol{\beta}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3) \neq 0)$ is the hazard rate when patients move from status 2 to status 3 if $t > W_i$.

More importantly, the piecewise linear function (3) in the proposed joint model can be used to determine the time-dependent hazard ratios of the side event when the effect of the side event on the time-to-event outcome is non-proportional. For example, in our application example, this piecewise linear function can be used to calculate the dynamic relative risk of the pancreas transplant on all-cause graft loss at different time points ($D_k$) after a pancreas transplant. For instance, we set $D_k$ as $D_0 = 14$ days, $D_1 = 45$ days, $D_2 = 90$ days, $D_3 = 180$ days, $D_4 = 365$ days, and $D_5 = 730$ days from the date of pancreas transplant, and then we can obtain the relative hazard ratios at each time point from the joint model. These relative hazard ratios can supply some references to control the potential risk of the pancreas transplant in the clinical practice. We give the change curve of the relative hazard ratios over time $D_k$ and discuss it in more detail in Section 4.

## 3 Estimation method

We discuss the likelihood function in a general framework for this proposed joint model with latent variables. Let $\boldsymbol{\Theta} = (\boldsymbol{\gamma}^T, \boldsymbol{\alpha}^T, \boldsymbol{\lambda}^T, \mathbf{b}^T, \mathbf{B}^T, \sigma^2, \lambda_0)^T$ be the parameters to estimate. The overall likelihood function based on the observed information is given by

$$L(\Theta) = \prod_{i=1}^{n}\left[ f(t_i, \mathbf{w}_i(t_i), S_i, \delta_i | Y_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \lambda_0)\left\{\prod_{j=1}^{m_i} f(Y_{ij}|\mathbf{Z}_i, \boldsymbol{\beta}_i, t_i, \boldsymbol{\alpha}, \sigma^2)\right\} f(\boldsymbol{\beta}_i|\mathbf{b}, \mathbf{B})\right] \tag{4}$$

where

$$\begin{aligned} f(t_i, \mathbf{w}_i(t_i), \delta_i | Y_i, \mathbf{Z}_i, \boldsymbol{\gamma}, \lambda_0) = {} & \left\{\lambda_0\left\{\Phi(t, \mathbf{Z}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3), \boldsymbol{\beta}_i, \boldsymbol{\gamma}), \boldsymbol{\gamma}\right\}\Phi'(t, \mathbf{Z}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3), \boldsymbol{\beta}_i, \boldsymbol{\gamma})\right\}^{\delta_i} \\ & \times \exp\left[-\int_0^{\Phi(t_i, \mathbf{Z}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3), \boldsymbol{\beta}_i, \boldsymbol{\gamma})} \lambda_0(s)\mathrm{d}s\right] \end{aligned}$$

is the density function of the survival submodel of the proposed joint model, and

$$\begin{aligned} \Phi(t, \mathbf{Z}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) &= \int_0^t \phi(s, \mathbf{Z}_i, \mathbf{w}_i(t|\boldsymbol{\gamma}_3), \boldsymbol{\beta}_i, \boldsymbol{\gamma})\mathrm{d}s \\ &= \int_0^t \exp(\boldsymbol{\gamma}_1^T\mathbf{Z}_i + \boldsymbol{\gamma}_2^T\boldsymbol{\beta}_i + \mathbf{w}_i(t|\boldsymbol{\gamma}_3))\mathrm{d}s \end{aligned}$$

The function $f(Y_{ij}|\mathbf{Z}_i, \boldsymbol{\beta}_i, t_i, \boldsymbol{\alpha}, \sigma^2)$ is the density function of Normal$(\boldsymbol{\alpha}^T\mathbf{Z}_i + \boldsymbol{\beta}_i^T\xi(t_i), \sigma^2)$, and $f(\boldsymbol{\beta}_i|\mathbf{b}, \mathbf{B})$ is the density function of Normal$(\mathbf{b}, \mathbf{B})$.

We propose to estimate the parameters in the joint model (1) by using the Monte Carlo EM algorithm.[22] The EM-algorithm[23] is an iterative procedure with two steps: the expectation (E) step and the maximization (M) step. In the E-step, we compute the expectation of joint log-likelihood function over the latent variable $\boldsymbol{\beta}_i$ using the observations and parameter estimates obtained so far. In the M-step, we maximize the expected joint log-likelihood over the parameters.

### 3.1 E-step

At the $t$-th iteration of the E-step, the expectation of the log-likelihood function w.r.t the latent variable $\boldsymbol{\beta}_i$ can be expressed in the following form

$$\begin{aligned} Q(\Theta|\Theta^{(t)}) &= E_{\boldsymbol{\beta}}\left[\log L(\Theta|t, \mathbf{w}(t), S, \delta, \mathbf{Z}, Y(t))|\Theta^{(t)}\right] \\ &= \sum_{i=1}^{n}\int\Big[\log f(t_i, \mathbf{w}(t_i), S_i, \delta_i|\boldsymbol{\beta}_i, \boldsymbol{\gamma}, \lambda_0) + \sum_{j=1}^{m_i}\log f(Y_{ij}|\boldsymbol{\beta}_i, \boldsymbol{\alpha}, \sigma^2) \\ &\quad + \log f(\boldsymbol{\beta}_i|\mathbf{b}, \mathbf{B})\Big] f\left(\boldsymbol{\beta}_i|t, \mathbf{w}_i(t), S_i, \delta_i, \mathbf{Z}_i, Y_i(t), \Theta^{(t)}\right)\mathrm{d}\boldsymbol{\beta}_i \end{aligned} \tag{5}$$

where $f(\boldsymbol{\beta}_i|t_i, \mathbf{w}_i(t), S_i, \delta_i, \mathbf{Z}_i, Y_i(t), \Theta^{(t)}) = \frac{f(\boldsymbol{\beta}_i|\mathbf{Z}_i, Y_i(t), \Theta^{(t)}) f(t_i, \mathbf{w}_i(t), S_i, \delta_i|\boldsymbol{\beta}_i, \Theta^{(t)})}{f(t_i, \mathbf{w}_i(t), S_i, \delta_i|\mathbf{Z}_i, Y_i(t), \Theta^{(t)})}$,

$$f(\boldsymbol{\beta}_i|\mathbf{Z}_i, Y_i(t), \Theta^{(t)}) \sim MVN\left(\mathbf{A}_i\left[\frac{\boldsymbol{\xi}_i^T(t)(Y_i(t) - \mathbf{Z}_i^T\alpha)}{\sigma^2}\right], \mathbf{A}_i\right)$$

$\mathbf{A}_i = \left[\frac{\boldsymbol{\xi}_i^T(t)\boldsymbol{\xi}_i(t)}{\sigma^2} + \mathbf{B}^{-1}\right]^{-1}$. The integral in the above equation is intractable because of the intractability of normalizing constant $f(t_i, \mathbf{w}_i(t), S_i, \delta_i|\mathbf{Z}_i, Y_i(t), \Theta^{(t)})$. An alternative is to use the importance sampling to approximate the integral in E-step.

- Draw $N$ samples $\boldsymbol{\beta}_i^{(1)}, \ldots, \boldsymbol{\beta}_i^{(N)}$ from $f(\boldsymbol{\beta}_i|\mathbf{Z}_i, Y_i(t), \Theta^{(t)})$ based on the current parameter estimates $\Theta^{(t)}$, and compute the normalized weights $w_i^{(s)} \propto f(t_i, \mathbf{w}_i(t), S_i, \delta_i|\boldsymbol{\beta}_i^{(s)}, \Theta^{(t)})$.
- Calculate $\hat{Q}(\Theta|\Theta^{(t)}) = \sum_{i=1}^{n}\sum_{s=1}^{N} w_i^{(s)} \cdot l_i^{(s)}(\Theta|t_i, \mathbf{w}_i(t), S_i, \delta_i, \mathbf{Z}_i, Y_i(t), \Theta^{(t)})$, where $l_i^{(s)} = \log f(t_i, \mathbf{w}(t_i), S_i, \delta_i|\boldsymbol{\beta}_i^{(s)}, \Theta^{(t)}) + \sum_{j=1}^{m_i} \log f(Y_{ij}|\boldsymbol{\beta}_i^{(s)}, \Theta^{(t)}) + \log f(\boldsymbol{\beta}_i^{(s)}|\Theta^{(t)})$

## 3.2 M-step

After computing the expectation of the log-likelihood function in equation (5), in M-step we estimate each parameter of $\Theta$ by maximizing $\hat{Q}(\Theta|\Theta^{(t)})$. The MLEs of $\widehat{\mathbf{b}}$, $\widehat{\mathbf{B}}$, $\widehat{\boldsymbol{\alpha}}$, $\hat{\sigma}^2$, the baseline hazard function $\hat{\lambda}_0(t)$ are derived in the supplementary document. The MLE of $\boldsymbol{\gamma}$ has no closed form, hence we could use the numeric optimization algorithm[24] to optimize this parameter. We repeat the E-step and M-step until convergence achieved. The convergence criterion for MCEM in our numerical study is

$$\max\left\{\frac{|\Theta^{(t)} - \Theta^{(t-1)}|}{|\Theta^{(t)}| + \epsilon_2}\right\} < \epsilon_1$$

where we set $\epsilon_1 = 0.002$ and $\epsilon_2 = 0.001$. The standard error of $\hat{\Theta}$ is computed using the bootstrap method.[25]

# 4 Application to clinical transplant data

The clinical transplant data resource is from the United Network for Organ Sharing. As mentioned in the introduction, all patients ($N = 13{,}635$) have both an end-stage renal disease (ESRD) and a diabetic disease. In this data, all patients already have a kidney transplantation from a living or deceased donor, and all of them are on the waiting list for the pancreas transplant. A part of patients ($N = 2776$) may have a pancreas transplant at any time during the follow-up period. We apply the proposed joint model to this clinical transplant data in this section. The main result from the proposed joint model is shown in Section 4.1, and the effect of the side event of pancreas transplant on the all-cause graft loss is shown in Section 4.2.

## 4.1 Main results from the joint model

In order to demonstrate the feasibility of the proposed joint model (1), we apply it to some clinical transplant data in this section. As shown in Figure 1, the baseline of GFR and the slope of GFR are related to the time-to-event outcome as mentioned before. So we choose two latent factors $\beta_{i1}$ and $\beta_{i2}$, and the joint model can be specified as in the following

$$\begin{cases} Y_i(t) = \boldsymbol{\alpha}^T\mathbf{Z}_i + \boldsymbol{\beta}_i^T\boldsymbol{\xi}(t) + \epsilon_i, \quad i = 1, \ldots, n, \\ \lambda(t|\mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) = \lambda_0\{\int_0^t \phi(s, \mathbf{Z}_i, \mathbf{w}_i(s), \boldsymbol{\beta}_i, \boldsymbol{\gamma})\mathrm{d}s\}\phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) \end{cases} \tag{6}$$

where $\phi(t, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma}) = \exp[\boldsymbol{\gamma}_1^T\mathbf{Z}_i + \boldsymbol{\gamma}_2^T\boldsymbol{\beta}_i + \mathbf{w}_i(t|\boldsymbol{\gamma}_3)]$, and $Y_i(t)$ is the GFR value at various time points post kidney transplant. The GFR value is calculated according to the formula in the paper[26]: $GFR = 141 \times \min(\mathrm{Scr}/\mathrm{d}, 1)^e \times \max(\mathrm{Scr}/\mathrm{d}, 1)^{-1.209} \times 0.993^{\mathrm{Age}} \times (1.018 \text{ if female}) \times (1.159 \text{ if black})$, where Scr is the measured serum creatinine in mg/dL, and the serum creatinine is a chemical waste product from the muscle metabolism and blood. The parameter $d = 0.7$ if female or 0.9 if male, and the parameter $e = -0.329$ if female or –0.411 if male. Let $W_i$ be the time from the kidney transplant to the pancreas transplant or the time on the waiting-list for patients without a pancreas transplant. We specify the piecewise linear function $\mathbf{w}_i(t)$ in

$\phi(t_i, \mathbf{Z}_i, \mathbf{w}_i(t), \boldsymbol{\beta}_i, \boldsymbol{\gamma})$ in equation (6) as follows

$$\mathbf{w}_i(t|\boldsymbol{\gamma}_3) = \begin{cases} 0 & \text{if } t \le W_i \\ \gamma_{30} & \text{if } W_i < t \le W + \frac{14}{365} \\ \gamma_{31}\left(t - W_i - \frac{14}{365}\right) & \text{if } W_i + \frac{14}{365} < t \le W_i + \frac{45}{365} \\ \gamma_{32}\left(t - W_i - \frac{45}{365}\right) & \text{if } W_i + \frac{45}{365} < t \le W_i + \frac{90}{365} \\ \gamma_{33}\left(t - W_i - \frac{90}{365}\right) & \text{if } W_i + \frac{90}{365} < t \le W_i + \frac{180}{365} \\ \gamma_{34}\left(t - W_i - \frac{180}{365}\right) & \text{if } W_i + \frac{180}{365} < t \le W_i + \frac{365}{365} \\ \gamma_{35}\left(t - W_i - \frac{365}{365}\right) & \text{if } W_i + \frac{365}{365} < t \le W_i + \frac{730}{365} \\ \gamma_{36} & \text{if } t > W_i + \frac{730}{365} \end{cases} \tag{7}$$

Table 1 displays the coefficients and standard errors of all parameters in the longitudinal sub-model and the AFT survival sub-model. The baseline term ($\beta_1$) in the mixed-effects submodel is 48.94, which indicates that most patients have a good kidney function at the baseline. The slope term ($\beta_2 = -1.36$) of GFR is negative and statistically significant, which means that the kidney function progression decreases during the follow-up period time. The estimates for other coefficients in the mixed-effect submodel are also reasonable. For example, the value of GFR decreases by 0.17 as the age of patients increases by 1. In other words, the kidney function of older patients is worse than young patients. The average GFR value of patients with a decreased donor is 1.15 less than patients with a living donor.

**Table 1.** Estimates for parameters in Model (1).

| | The longitudinal submodel | | The survival submodel | |
|---|---|---|---|---|
| Parameters | Coef. (SE) | *P* value | Coef. (SE) | *P* value |
| Age (per year) | −0.17 (0.05) | <0.001 | 0.02 (0.01) | 0.044 |
| Female | 5.39 (0.73) | <0.001 | −0.23 (0.03) | 0.029 |
| Black | −3.69 (1.34) | <0.001 | 0.17 (0.08) | 0.020 |
| Other | −6.45 (1.08) | <0.001 | 0.23 (0.07) | <0.001 |
| TX era 1993 − 1997 | 7.56 (1.15) | <0.001 | −0.29 (0.04) | 0.080 |
| TX era 1998 − 2002 | 10.75 (1.16) | <0.001 | −0.76 (0.03) | <0.001 |
| TX era 2003 − 2007 | 16.52 (1.30) | <0.001 | −0.95 (0.03) | <0.001 |
| PKPRA 1 − 29 | −0.93 (0.17) | 0.024 | 0.06 (0.01) | 0.030 |
| PKPRA 30 − 100 | −2.35 (0.11) | 0.159 | 0.27 (0.13) | 0.049 |
| HLA mismatch 1 − 6 | −1.54 (0.61) | 0.045 | 0.24 (0.05) | 0.004 |
| Dialysis time 0.1 − 1 years | −0.27 (0.17) | 0.689 | 0.07 (0.01) | 0.005 |
| Dialysis time 1.1 − 2 years | −0.52 (0.42) | 0.166 | 0.08 (0.01) | 0.007 |
| Dialysis time 2.1 − 3 years | −0.73 (1.01) | 0.142 | 0.33 (0.03) | 0.028 |
| Dialysis time > 3 years | −0.96 (0.11) | 0.029 | 0.38 (0.05) | <0.001 |
| Decreased Donor | −1.15 (0.18) | 0.015 | 0.14 (0.05) | 0.024 |
| $\beta_1$ | 48.94 (2.34) | <0.001 | | |
| $\beta_2$ | −1.36 (0.12) | <0.001 | | |
| $\gamma_{21}$ | | | −0.07 (0.01) | <0.001 |
| $\gamma_{22}$ | | | −0.21 (0.04) | <0.001 |
| $\gamma_{30}$ | | | 1.22 (0.01) | <0.001 |
| $\gamma_{31}$ | | | −9.42 (0.04) | 0.035 |
| $\gamma_{32}$ | | | −2.51 (0.05) | 0.041 |
| $\gamma_{33}$ | | | −0.65 (0.45) | 0.542 |
| $\gamma_{34}$ | | | −0.35 (0.21) | 0.251 |
| $\gamma_{35}$ | | | −0.06 (0.01) | 0.045 |
| $\gamma_{36}$ | | | −0.28 (0.04) | <0.001 |
| Random-effect parameters | | Value (Std. Error) | Correlation | |
| SD ($\beta_1$) | | 14.91 (2.30) | –0.40 | |
| SD ($\beta_2$) | | 2.89 (0.12) | | |

Note: The standard errors of the estimates are given in brackets.

In the survival sub-model, the coefficients ($\gamma_{21}$ and $\gamma_{22}$) of the random intercept and slope ($\beta_1$ and $\beta_2$) of GFR are negative, and they are also statistically significant. These results indicate that the latent baseline level and the latent slope of GFR are related to the time-to-event outcome ACGL. In other words, the failure rate of ACGL increases as the value of GFR decreases during the followed-up time period, and patients in the higher baseline of GFR are less likely to have the time-to-event outcome ACGL. The coefficients of female patients are larger than male patients. It is reasonable that patient age is significantly related to all-cause graft loss, which indicates that patients are more likely to have an ACGL with the hazard ratio (1.02) as patient age increases per year. Compared with the white patients, the black patients are more likely to have the time-to-event outcome ACGL. The transplantation era and the dialysis duration before transplant are also related to the time-to-event outcome ACGL. For example, patients have a longer dialysis duration before kidney transplant, and the more likely patients have all-cause graft failure. Compared with patients who have a deceased donor, patients who have a living donor transplant are less likely to have the time-to-event outcome ACGL. The dynamic effect of the pancreas transplant on all-cause graft loss is presented in the next section.

## 4.2 Effect of pancreas transplant on allograft

In order to evaluate the average and time-varying relative risk of the pancreas transplant on all-cause graft loss, we can set the piecewise linear function $\mathbf{w}_i(t|\boldsymbol{\gamma}_3)$ in equation (3) in two separate forms. In order to evaluate the average relative risk of the pancreas transplant on all-cause graft loss, we set $\mathbf{w}_i(t|\boldsymbol{\gamma}_3)$ as

$$\mathbf{w}_i(t|\boldsymbol{\gamma}_3) = \begin{cases} 0 & \text{if } t \leq W_i, \\ \gamma_{31} & \text{if } t > W_i \end{cases}$$

Then the coefficient vector $\boldsymbol{\gamma}_3$ in the piecewise function has only one element $\gamma_{31}$. In fact, the coefficient $\gamma_{31}$ represents the average relative risk of the side event on the time-to-event outcome in this case when we set $K=0$ and $D_0 = 0$ in equation (3). We find that the pancreas transplant has a significantly statistical benefit effect on ACGL because the hazard ratio is $\exp(\gamma_{31}) = \exp(-0.13) = 0.88$ with the $p$-value 0.045. In other words, the pancreas transplant can reduce the risk of the time-to-event outcome ACGL.

However, the relative risk of this side event on the time-to-event outcome is non-proportional. Therefore, we need to display the relative risk curve at various time points post pancreas transplant. It is also useful to control the potential risk for the clinical practice if we can determine the relative risk at specified time points. In order to evaluate the time-varying relative risk of the pancreas transplant on all-cause graft loss, we set the piecewise linear function $\mathbf{w}_i(t|\boldsymbol{\gamma}_3)$ as the formula (7) after specifying the values of $D_k$ as $D_0 = 14$ days, $D_1 = 45$ days, $D_2 = 90$ days, $D_3 = 180$ days, $D_4 = 365$ days, and $D_5 = 730$ days. Then we obtain the coefficient vector $\boldsymbol{\gamma}_3$ at these specified time points from the date of pancreas transplant in comparison with patients without pancreas transplant, which is shown in Table 1.

For easy comparison, we transform the estimated coefficients into the hazard ratios. Figure 4 displays the hazard ratio curve of pancreas transplant at different time points. It shows that the hazard ratio is very high in the beginning because of the clinical surgery or organ acute rejection, then the hazard ratio decreases to 1.00 at 152 days from the date of pancreas transplant, and then becomes less than 1 thereafter. From the time point when the hazard ratio is equal to 1.00, the pancreas transplantation starts to have a survival benefit. It is a good clinical example to demonstrate the hazard ratio curve when the hazard ratios are not proportional.

# 5 Simulations

## 5.1 Simulation 1

The first simulation study is implemented to access the finite-sample performance of our proposed MCEM algorithm in Section 3. A multivariate mixed-effects model is chosen to simulate the longitudinal trajectories

$$Y_i(t) = \sum_{p=1}^{15} \alpha_p Z_{ip} + \beta_{i1} + \beta_{i2} t + \boldsymbol{\epsilon}_i$$

where $\beta_{i1} = \beta_1 + b_{i1},\ b_{i1} \sim \text{Normal}(0, \sigma_1^2),\ \beta_{i2} = \beta_2 + b_{i2}$, and $b_{i2} \sim \text{Normal}(0, \sigma_2^2),\ i = 1, \ldots, n$. In our application example, the longitudinal outcome is GFR. Here $\alpha_1, \alpha_2, \ldots, \alpha_{15}$ are the coefficients for age, gender, and other fixed covariates shown in Table 1. We set $\beta_1 = 48.94,\ \sigma_1 = 14.91,\ \beta_2 = -1.36$, and $\sigma_2 = 2.89$, which are the estimate

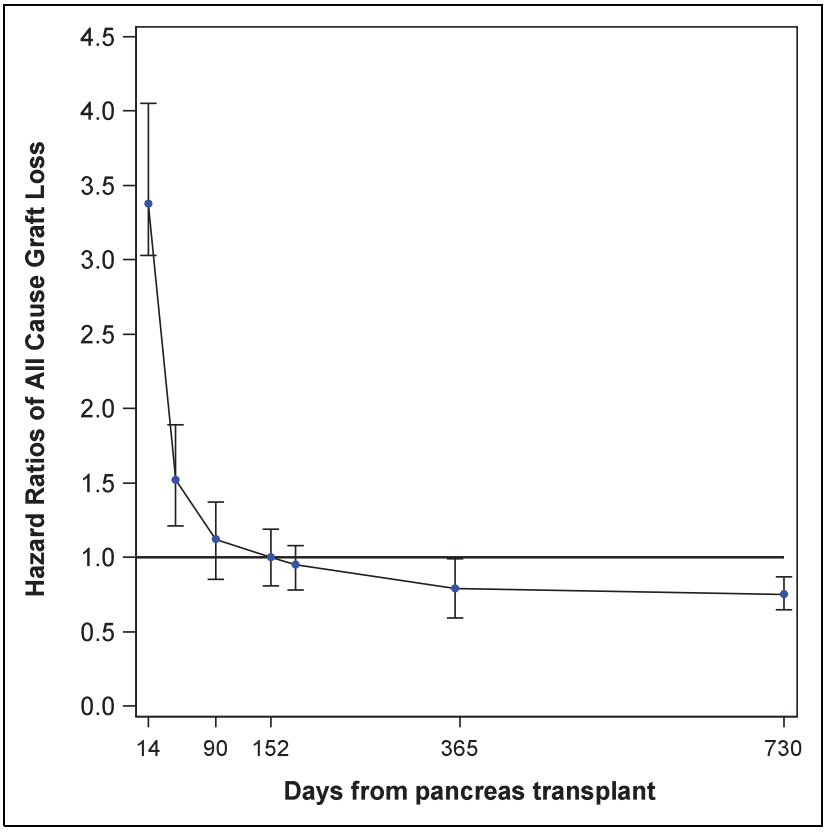


**Figure 4.** The curve of hazard ratios of all-cause graft loss for patients with a pancreas transplant with the 95% confidence intervals at 14, 45, 90, 152, 180, 365, 730 days from the date of pancreas transplant. The reference group are patients without a pancreas transplant. The hazard Ratio curve reaches 1.00 at 152 days from the date of pancreas transplant.

from the real data in Table 1. The measurement error $\epsilon_{ij} \sim \text{Normal}(0, \sigma_3^2)$, where we set $\sigma_3 = 0.85$ and $j = 1, \ldots, 12$. The scheduled measurement times of the repeated longitudinal outcome are set at the sequence year $(1, 2, \ldots, 12)$ for each subject, but there are no measurements available after death or censoring time. The time-to-event $T_i$ is specified as follows

$$Log(T_i) = \sum_{p=1}^{15} \gamma_{1p} Z_{ip} + \gamma_{21}\beta_{i1} + \gamma_{22}\beta_{i2} + w(t|\gamma_3) + \tau_i$$

where $\gamma_1, \gamma_2, \ldots, \gamma_{15}$ are the coefficients for age, gender, and other fixed covariates, $\gamma_{21}$ and $\gamma_{22}$ are the coefficient for the random effects $\beta_{1i}$ and $\beta_{2i}$, and the random error $\tau_i \sim$ Gumbel (0, 1). We set their true values as the estimates from the real data shown in Table 1. The piecewise linear function $\mathbf{w}_i(t|\boldsymbol{\gamma}_3)$ is specified as follows

$$\mathbf{w}_i(t|\boldsymbol{\gamma}_3) = \begin{cases} 0 & \text{if } t \le W_i \\ \gamma_{31} & \text{if } t > W_i \end{cases}$$

where $\gamma_{31} = -0.13$. The number of subjects are set as $n = 100$.

We estimate the joint model (1) with the Monte Carlo EM algorithm from the simulated data. The simulation procedure is repeated for 100 replicates. Table 2 shows the parameter estimates, together with biases and root mean square errors (RMSEs). It shows that the means of the parameter estimates by the MCEM algorithm are close to their true values. The average number of iterations till convergence is 12. We notice that the estimate for $\beta_1$ and $\beta_1$ has large RMSE, which is caused by the setting of our simulated data. In the simulation data, we set $\beta_1 = 48.94$, $\sigma_1 = 14.91$, $\beta_2 = -1.36$, and $\sigma_2 = 2.89$, which are the estimated from the real transplant data. We find that the MCEM algorithm can estimate parameters accurately in the proposed joint model, which is consistent with the literature.[13,14]

**Table 2.** Means, biases, root mean square errors (RMSEs) of the parameter estimates for the joint model (1) using our proposed MCEM algorithm in Simulation 1.

| | The longitudinal submodel | | | | The survival submodel | | | |
|---|---|---|---|---|---|---|---|---|
| Parameters | True | Mean | Bias | RMSE | True | Mean | Bias | RMSE |
| Age (per year) | −0.17 | −0.17 | 0.00 | 0.012 | 0.02 | 0.02 | −0.00 | 0.001 |
| Female | 5.39 | 5.41 | −0.02 | 0.236 | −0.23 | −0.23 | −0.00 | 0.010 |
| Black | −3.69 | −3.68 | −0.01 | 0.235 | 0.17 | 0.17 | 0.00 | 0.009 |
| Other | −6.45 | −6.45 | 0.00 | 0.220 | 0.23 | 0.23 | 0.00 | 0.010 |
| TX era 1993 − 1997 | 7.56 | 7.57 | −0.01 | 0.300 | −0.29 | −0.29 | 0.00 | 0.010 |
| TX era 1998 − 2002 | 10.75 | 10.80 | −0.05 | 0.245 | −0.76 | −0.76 | −0.00 | 0.010 |
| TX era 2003 − 2007 | 16.52 | 16.55 | −0.03 | 0.227 | −0.95 | −0.95 | 0.00 | 0.011 |
| PKPRA 1 − 29 | −0.93 | −0.93 | −0.00 | 0.114 | 0.06 | 0.06 | 0.00 | 0.003 |
| PKPRA 30 − 100 | −2.35 | −2.43 | −0.08 | 0.119 | 0.27 | 0.27 | 0.00 | 0.011 |
| HLA Mismatch 1 − 6 | −1.54 | −1.54 | 0.00 | 0.132 | 0.24 | 0.24 | −0.00 | 0.010 |
| Dialysis time 0.1 − 1 years | −0.27 | −0.27 | 0.00 | 0.121 | 0.07 | 0.07 | 0.00 | 0.005 |
| Dialysis time 1.1 − 2 years | −0.52 | −0.49 | −0.03 | 0.118 | 0.08 | 0.08 | −0.00 | 0.005 |
| Dialysis time 2.1 − 3 years | −0.67 | −0.66 | −0.01 | 0.147 | 0.33 | 0.33 | −0.00 | 0.004 |
| Dialysis time > 3 years | −0.96 | −0.94 | −0.02 | 0.163 | 0.38 | 0.38 | −0.00 | 0.005 |
| Decreased Donor | −1.15 | −1.14 | −0.01 | 0.109 | 0.14 | 0.14 | 0.00 | 0.004 |
| $\beta_1$ | 48.94 | 49.35 | −0.41 | 1.439 | | | | |
| $\beta_2$ | −1.36 | −1.47 | 0.11 | 0.505 | | | | |
| $\gamma_{21}$ | | | | | −0.07 | −0.07 | −0.00 | 0.002 |
| $\gamma_{22}$ | | | | | −0.21 | −0.21 | −0.00 | 0.005 |
| $\gamma_3$ | | | | | −0.13 | −0.13 | 0.00 | 0.005 |

## 5.2 Simulation 2

In order to study the effect of the various correlation construction between the longitudinal submodel and the survival model, we develop two simulation studies in this section.

The relationship between the longitudinal submodel and the survival model in our proposed joint model (1) is based on latent feathers. Some other studies treat the longitudinal outcome as a covariable in the survival model as shown in the model (8) such as[10,12,14]

$$\begin{cases} Y_i(t) = X_i(t) + \epsilon_i, \quad i = 1, \ldots, n, \\ \lambda(t|X(t)) = \lambda_0\{\int_0^t \gamma_1 X(s)\mathrm{d}s\} \exp(\gamma_1 X(t)) \end{cases} \tag{8}$$

The simulation data are generated in two scenarios. In the first scenario, we set $\gamma_1 = \gamma_2 = 1.00$. In the second scenario, we set $\gamma_1 = 1.00$ and $\gamma_2 = -1.00$. We choose these two different scenarios because we want to see the differences between the proposed model (1) and the model (8) when the effects of the intercept and slope of GFR curves are in the same or opposite direction. A mixed-effects model is chosen to mimic the longitudinal trajectories

$$Y_i(t) = \beta_{i1} + \beta_{i2} t + \epsilon_i$$

where $\beta_{i1} = \beta_1 + b_{i1},\ b_{i1} \sim \text{Normal}(0, \sigma_1^2),\ \beta_{i2} = \beta_2 + b_{i2},\ b_{i2} \sim \text{Normal}(0, \sigma_2^2),\ i = 1, \ldots, n$. Here we set the true values for these parameters as $\beta_1 = 2.50,\ \sigma_1 = 1,\ \beta_2 = -0.20$, and $\sigma_2 = 0.02$. The random measurement error $\epsilon_i \sim$ Normal(0, 1), and the number of subjects are set as $n = 100$. The preliminary scheduled measurement time of the longitudinal outcome is set at the sequence year $(1, 2, \ldots, 12)$ for each subject, but there are no measurements available after death or censoring time. The time-to-event $T_i$ is specified as follows

$$Log(T_i) = \gamma_1 \beta_{i1} + \gamma_2 \beta_{i2} + \tau_i$$

where the random error $\tau_i \sim$ Gumbel (0, 1). Note that our proposed joint model (1) and the alternative model (8) treat $Y_i(t)$ in two different ways. Our proposed joint model (1) chooses a mixed-effects submodel for $Y_i(t)$, and

**Table 3.** Means and standard deviations (STD) of the parameter estimates for our proposed joint model (1) and the model (8) in Simulation 2.

| | Scenario 1 | | Scenario 2 | |
|---|---|---|---|---|
| Parameters | Mean (STD) $\gamma_1$ | Mean (STD) $\gamma_2$ | Mean (STD) $\gamma_1$ | Mean (STD) $\gamma_2$ |
| True value | 1.00 | 1.00 | 1.00 | −1.00 |
| Fitted value in Model(1) | 0.98(0.05) | 1.00(0.05) | 1.02(0.05) | −0.98(0.01) |
| Fitted value in Model(8) | 0.64(0.03) | – | −0.24(0.06) | – |

shares two random parameters $\beta_{i1}$ and $\beta_{i2}$ with the AFT survival submodel. The alternative model (8) treats $X_i(t)$ as a covariate in the AFT survival component.

We estimate the joint model (1) with the Monte Carlo EM algorithm from the simulated data. The simulation procedure is repeated for 100 replicates. The average number of steps till convergence is 32. Table 3 displays the parameter estimates, together with their estimated standard errors. In Scenario 1, when the coefficients ($\gamma_1 = \gamma_2 = 1.00$) of the intercept ($\beta_{i1}$) and the slope ($\beta_{i2}$) are same, the estimated coefficient $\hat{\gamma}_1$ for the model in Tseng et al.[14] has the same sign as the true value $\gamma_1$, although there is a relatively large gap between them. In Scenario 2 when the coefficients ($\gamma_1 = 1.00$ and $\gamma_2 = -1.00$) are different, the estimated coefficient $\hat{\gamma}_1$ from model (8) is completely different from the true value. In summary, the results from Table 3 demonstrate that model (8) cannot describe both the relationship between the intercept ($\beta_{i1}$) and the slope ($\beta_{i2}$) of the longitudinal outcome with the time-to-event outcome by a single parameter $\gamma_1$. Especially in Scenario 2 when the intercept and the slope are in an opposite relationship with the time-to-event outcome ($\gamma_1 = 1.00$ and $\gamma_2 = -1.00$), it is impossible to describe the two relationships by a single parameter $\gamma_1$. This simulation example shows the advantages in our proposed joint model by using the features from the longitudinal submodel in the survival submodel.

The second advantage of our proposed joint model is that the estimated results offer a straightforward interpretation. For example, in Scenario 2, if patients have a higher baseline, i.e. a larger $\beta_{i1}$, then patients are more likely to have the time-to-event outcome. So physicians can tell patients which level they are in and the corresponding risk to have the time-to-event given a baseline value. Similarly, if patients have a larger value of the slope, i.e. a larger $\beta_{i2}$, then patients are less likely to have a time-to-event outcome. So physicians can tell patients the trend of the longitudinal outcome and the corresponding risk to have a time-to-event given the value of the slope.

## 5.3 Simulation 3

In this section, we investigate the effect of misspecification of the distribution of random effects on parameter estimates. A mixed-effects model is chosen to mimic the longitudinal trajectories

$$Y_i(t) = \beta_{i1} + \beta_{i2}t + \epsilon_i$$

where the random effects $\beta_{i1}$ and $\beta_{i2}$ are sampled in two scenarios. In the first scenario, $\beta_{i1}$ and $\beta_{i2}$ are sampled from the normal distribution $\beta_{i1} = \beta_1 + b_{i1},\ b_{i1} \sim \text{Normal}(0, \sigma_1^2),\ \beta_{i2} = \beta_2 + b_{i2},\ b_{i2} \sim \text{Normal}(0, \sigma_2^2),\ i = 1, \ldots, n$. Here we set the true values for these parameters as $\beta_1 = 2.50,\ \sigma_1 = 1,\ \beta_2 = -0.20$, and $\sigma_2 = 0.02$. In the second scenario, $\beta_{i1}$ and $\beta_{i2}$ are sampled from a bimodal mixture of normal distributions, where we set $\beta_{i1} \sim 0.55 \cdot N(3, 0.7^2) + 0.45 \cdot N(1, 0.5^2)$ and $\beta_{i2} \sim 0.55 \cdot N(-0.3, 0.03^2) + 0.45 \cdot N(-0.1, 0.01^2)$. The random measurement error $\epsilon_i \sim \text{Normal}(0, 1), i = 1, \ldots, n$. The number of subjects is set to be $n = 100$. The preliminary scheduled measurement times of the longitudinal outcome are set at the sequence year $(1, 2, \ldots, 12)$ for each subject, but there are no measurements available after the date of time-to-event outcome or censoring time. The time-to-event $T_i$ is specified as follows

$$Log(T_i) = \gamma_1\beta_{i1} + \gamma_2\beta_{i2} + \tau_i$$

where $\gamma_1 = 1.00,\ \gamma_2 = 1.00$, and the random error $\tau_i \sim$ Gumbel (0, 1).

We estimate the joint model (1) with the Monte Carlo EM algorithm from the simulated data. Therefore, the first scenario has the correct model assumption and the second scenario has the misspecified model assumption.

**Table 4.** The mean, bias, standard deviation (STD), and root mean squared error (RMSE) of the parameter estimates for the joint model (1) when the model assumption is correct or misspecified in Simulation 3.

| Model Assumption | Correct | | Misspecified | |
|---|---|---|---|---|
| Parameters | $\gamma_1$ | $\gamma_2$ | $\gamma_1$ | $\gamma_2$ |
| True value | 1.000 | 1.000 | 1.000 | 1.000 |
| Mean | 0.984 | 1.002 | 0.986 | 0.996 |
| Bias | –0.016 | 0.002 | –0.014 | –0.004 |
| STD | 0.053 | 0.049 | 0.051 | 0.053 |
| RMSE | 0.055 | 0.048 | 0.053 | 0.053 |
| 95%CI | 96% | 95% | 96% | 96% |

The simulation procedure is repeated for 100 replicates. The average number of steps till convergence is 30. Table 4 displays the summary of the simulation results in the two scenarios. In comparison with the simulation results when the distribution for the random effects is correctly specified, the simulation results are similar when the distribution for the random effects is incorrectly specified as the bimodal mixture of normal distributions.

## 6 Conclusions and discussion

This paper is motivated by a longitudinal and time-to-event transplant data set. Our proposed joint model has a longitudinal submodel and an AFT sub-model, and both submodels share a vector of latent variables with each other. Our proposed joint model has three major advantages. First, as shown in Table 3, the model[14] cannot correctly describe the relationships between the time-to-event outcome and the longitudinal process when the intercept and the slope of the longitudinal process are in an opposite relationship with the time-to-event outcome. Second, it is one of few joint models with an AFT regression rather than Cox regression. To calculate the dynamic hazard ratio curve when the proportional hazards assumption is not satisfied, this joint model includes a piecewise linear function in the AFT regression. Finally, this model can estimate the parameters of the longitudinal component by incorporating the time-to-event information through censoring, and similarly, the estimation of the time-to-event accommodates the longitudinal data information.

The proposed joint model is demonstrated with a real clinical transplantation application. The estimation results from our proposed joint model provide at least two useful guidelines for the clinical practice. First, it confirms that the latent baseline and slope of GFR trajectories are significantly related to ACGL. The slope of GFR is negatively correlated with ACGL, which means that patients are more likely to have ACGL when GFR decreases. In addition, patients with a lower baseline GFR are more likely to have ACGL. Second, the hazard ratio curve of the effect of pancreas transplant on ACGL helps to understand the risk process of the pancreas transplant for clinical physicians. For example, the hazard ratio is very high in the beginning because of the clinical surgery or acute rejection, decreases to 1 at 152 days post pancreas transplant, and then becomes less than 1. From the time point when the hazard ratio is equal to 1, the pancreas transplant starts to have some survival benefit in comparison with no pancreas transplant. Our proposed joint model can also be applied to other areas, although it is motivated by a clinical data of multiple organ transplantations.

### Acknowledgements

We are very grateful for the kind reviewing and very constructive comments of the Editor, the Associate Editor and two reviewers, which were extremely helpful for us to improve our work.

### Declaration of conflicting interests

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

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grants to L. Wang and J. Cao.

## Supplemental material

Supplementary material for this article is available online.

## References


1. Wolfe RA, Ashby VB, Milford EL, et al. Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of first cadaveric transplant. *New Engl J Med* 1999; **341**: 1725–1730.
2. Knoll G and Nichol G. Dialysis, kidney transplantation, or pancreas transplantation for patients with diabetes mellitus and renal failure: a decision analysis of treatment options. *J Am Soc Nephrol* 2003; **14**: 500–515.
3. Orsenigo E, Fiorina P, Cristallo M, et al. Long-term survival after kidney and kidney-pancreas transplantation in diabetic patients. *Transplant Proc* 2004; **36**: 1072–1075.
4. Kleinclauss F, Fauda M, Sutherland DER, et al. Pancreas after living donor kidney transplants in diabetic patients: Impact on long-term Kidney graft function. *Clin Transplant* 2009; **23**: 437–446.
5. Waki K, Terasaki PI and Kadowaki T. Long-term pancreas allograft survival in simultaneous pancreas-kidney transplantation by era. *Diab Care* 2010; **33**: 1789–1791.
6. Poommipanit N, Sampaio MS, Cho Y, et al. Pancreas after living donor kidney versus simultaneous pancreas-kidney transplant: an analysis of the Organ Procurement Transplant Network/United Network of Organ Sharing Database. *Transplantation* 2010; **89**: 1496–1503.
7. Pavlakis M, Khwaja K, Mandelbrot D, et al. Renal allograft failure predictors after PAK transplantation: results from the New England Collaborative Association of Pancreas Program. *J Transplant* 2010; **89**: 1347–1353.
8. Laird N and Ware J. Random-effects models for longitudinal data. *Biometrics* 1982; **38**: 963–974.
9. Pocock S, Geller N and Tsiatis A. The analysis of multiple endpoints in clinical trials. *Biometrics* 1987; **43**: 487–498.
10. Tsiatis AA, DeGruttola V and Wulfsohn MS. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *J Am Stat Assoc* 1995; **90**: 27–37.
11. Wulfsohn MS and Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics* 1997; **53**: 330–339.
12. Tsiatis AA and Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Stat Sinica* 2004; **14**: 809–834.
13. Wu L, Liu W and Hu XJ. Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics* 2010; **66**: 327–335.
14. Tseng Y, Hsieh F and Wang J. Joint modelling of accelerated failure time and longitudinal data. *Biometrics* 2005; **92**: 587–603.
15. Song P, Li M and Yuan Y. Joint regression analysis of correlated data using Gaussian copulas. *Biometrics* 2009; **65**: 60–68.
16. Rizopoulos D, Hatfield L, Carlin B, et al. Combining dynamic predictions from joint models for longitudinal and time-to-event data using Bayesian model averaging. *J Am Stat Assoc* 2014; **109**: 1385–1397.
17. Cox D. Regression models and life tables. *J Royal Stat Soc Ser B* 1972; **34**: 187–220.
18. Cox D and Cakes D. *Analysis of survival data*. London: Chapman & Hall/CRC Press, 1984.
19. James R. Estimation of the time-dependent accelerated failure time model in the presence of confounding factors. *Biometrika* 1992; **79**: 321–334.
20. Reid N. A conversation with Sir David Cox. *Stat Sci* 1994; **9**: 439–455.
21. Kay R and Kinnersley N. On the use of the accelerated failure time model as an alternative to the proportional hazards model in the treatment of time to event data: a case study in influenza. *Drug Inform J* 2002; **36**: 571–579.
22. Wei G and Tanner MA. Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *J Am Stat Assoc* 1990; **85**: 699–704.
23. Dempster AP, Laird NM and Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J Royal Stat Soc Ser B (Methodol)* 1977; **39**: 1–38.
24. Avriel M. *Nonlinear programming: analysis and methods*. New York: Prentice-Hall, 1976.
25. Efron B and Tibshirani RJ. *An introduction to the bootstrap*. London: CRC Press, 1994.
26. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med* 2009; **150**: 604–612.