Universiteit Leiden
The Netherlands

Linked shrinkage prior: a Bayesian approach to detect interaction effects in survival analysis

Yufang Wang

Thesis advisor:
Prof. M.A. van de Wiel, Amsterdam UMC
Dr. H. Peters-Sengers, Amsterdam UMC

Defended on [date], [2025]

MASTER THESIS

STATISTICS AND DATA SCIENCE

UNIVERSITEIT LEIDEN

Summary

Previous studies on graft loss in kidney transplantation mainly used hypothesis-driven approaches, which focused on single interaction effects and ignoring other two-way interactions. This thesis adopts a data-driven approach to simultaneously investigate all two-way interactions. Building on the linked shrinkage prior proposed by van de Wiel et al., 2024, which connects interaction and main effects, we apply this prior to Bayesian proportional hazards models by using both full and partial likelihood estimators. For the baseline hazard function $h_0(t)$ in full likelihood estimator, we account for the following situation: (1) true event times T^* follow an Exponential distribution; (2) true event times T^* follow an Weibull distribution; and (3) true event times T^* do not follow a specific distribution, we use B-Splines method to approximate the corresponding log baseline hazard function. The method is validated through simulations and applied to kidney transplantation data. Results show that linked shrinkage prior improves detection of interaction effects compared to non-informative prior. However, its advantage diminishes with full likelihood estimators combined with B-Splines approximation of log baseline hazard function in the real-world kidney transplantation data. We conclude that linked shrinkage prior is effective for detecting interaction terms in survival analysis, but careful handling of baseline hazard estimation is crucial, with partial likelihood estimator being a safer choice.

Introduction

In kidney transplantation research, graft loss refers to the time duration before a transplanted kidney fails. This outcome is typically modeled as a time-to-event variable, and has traditionally been explained by average donor and recipient characteristics—such as HLA mismatch, donor type (donation after circulatory death, DCD vs. donation after brain death, DBD) and donor sex (female vs. male). Recent findings by Steenvoorden, Evers, et al. (2024) report that interaction effects between donor type and donor sex may also play a significant role in explaining graft loss. In their study, the authors began with a biological hypothesis and focused exclusively on the interaction between donor type and donor sex. All other interactions in their study were deliberately excluded. While this hypothesis-driven approach can increase statistical power for detecting the targeted interaction, it is inherently limited by its reliance on empirical theories, which may not always be comprehensive, accurate or objective. As a result, interactions not anticipated by these theories are overlooked. As a complementary research approach, including all possible two-way interactions in time-to-event analyses may provide more comprehensive and/or objective insights into graft loss from a data-driven perspective.

However, when including all possible two-way interactions in time-to-event analyses, three common problems may arise: (1) models, e.g., cox regression, may fail to converge (see, e.g., Beca et al., 2021; Johnson et al., 1982); (2) issues related to multiple testing (see, e.g., van de Wiel

et al., 2024); and (3) large variance or standard error of parameter estimates may occur. These challenges arise because (1) the number of interaction terms grows quadratically as the number of covariates increases; and (2) interaction terms are often highly correlated with their associated main effects.

In this thesis, we aim to address this gap by developing an interpretable and statistically robust approach that can handle a large number of parameters while still allowing for valid parameter inference. Specifically, we focus on the accurate estimation and selection of interaction effects to better explain graft loss in kidney transplantation.

Background

Before addressing the accurate estimation and selection of interaction effects to explain graft loss in kidney transplantation, it is helpful to first briefly introduce a commonly used framework to analyze time-to-event outcomes, i.e., *the proportional hazards (PH) models*. Note that for the sake of simplicity, in this thesis, we only focus on right-censored data.

This section first introduces the structure of time-to-event data, then explains key quantities in proportional hazards (PH) models. Next, it reviews two commonly used estimators and concludes with assumption checking using Schoenfeld residuals.

Data Structure

In time-to-event data analysis, the data structure is typically the following. For each individual i ($i = 1, \ldots, n$), we represent the observed data as Equation 1.

$$\mathcal{D}_i = \{\mathbf{x}_i, T_i, \delta_i\}, \tag{1}$$

where:

- $\mathbf{x}_i \in \mathbb{R}^{1\times p}$ is consisted of covariate vectors,
- T_i is the observed time to event or censoring,
- $\delta_i = \mathbb{I}(T_i \leq C_i)$ is the event indicator.

Specifically, $\delta_i = 1$ if the event is observed (i.e., $T_i^* = T_i$), and $\delta_i = 0$ if the observation is right-censored (i.e., $T_i^* > T_i$), where T_i^* denotes the true event time. C_i denotes the censoring time (the event of interest has already occurred for the individual before that person is observed in the study at time C_i.

Hazard function and other relevant quantities

In PH models, a central quantity is the *hazard function*, which describes the instantaneous risk of an event occurring at time t for individual i, given that the event has not yet occurred. Formally, the hazard function is defined as:

$$h_i(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T_i^* < t + \Delta t \mid T_i^* > t)}{\Delta t} = h_0(t) \exp(\boldsymbol{\eta}_i),$$

where $h_0(t)$ is the baseline hazard function and $\boldsymbol{\eta}_i$ represents the linear combination of covariates for individual i, i.e., $\mathbf{x}_i\boldsymbol{\beta}$. This formulation assumes (1) *non-informative censoring*, i.e., for each individual i, $T_i \perp C_i \mid \mathbf{x}_i$, meaning that for each individual i, the censoring time C_i is independent of the event time t_i given the covariates $\mathbf{x}_i$; and (2) *proportional hazard*, i.e., the effect of a covariate is constant over time.

In addition to the hazard function, several other quantities as follows are also fundamental in PH models:

- **Cumulative hazard function:**

$$H_i(t) = \int_0^t h_i(u)\, du = H_0(t) \exp(\boldsymbol{\eta}_i),$$

  which represents the accumulated risk up to time t for individual i.

- **Survival function:**

$$S_i(t) = \exp(-H_i(t)) = S_0(t)^{\exp(\boldsymbol{\eta}_i)},$$

  giving the probability that individual i survives beyond time t.

- **Density function:**

$$f_i(t) = S_i(t) \cdot h_i(t),$$

  which describes the probability distribution of the event time.

Estimators in PH models

Maximum Likelihood Estimator (MLE)

With the time-to-event data structure and the relevant quantities defined in the proportional hazards (PH) model, we now move on to one of the two primary estimators for the regression coefficients $\boldsymbol{\beta}$ in the linear predictor $\boldsymbol{\eta}$—namely, the *maximum likelihood estimator (MLE)*. For n independent subjects, the MLE is given in Equation 2.

$$\log L(\mathcal{D}_i \mid \boldsymbol{\beta}, \text{other parameters}) = \sum_{i=1}^{n} \Big\{ \mathbb{I}(\delta_i = 0) \log[S_i(t)] + \mathbb{I}(\delta_i = 1) \left(\log[h_i(t)] + \log[S_i(t)]\right) \Big\}, \tag{2}$$

The indicator function $\mathbb{I}(\cdot)$ ensures that the MLE accounts appropriately for both censored and uncensored observations.

Note that as shown in Equation 2, defining the MLE involves specifying one of the following functions: $S_0(t)$, $H_0(t)$, or $h_0(t)$. In this thesis, we focus on specifying $h_0(t)$ by considering both parametric and flexible parametric approaches for the sake of simplicity and practicality. That is, we consider the following:

1. **(Parametric) Exponential**: assume true event times T^* follow Exp(λ).
2. **(Parametric) Weibull**: assume true event times t follow Weibull(λ, v).
3. **(Flexible Parametric) B-Splines-based**: approximate the log baseline hazard function using B-Splines, i.e., $\log h_0(t) \approx \sum_{l=1}^{L} \gamma_l B_l(t; k, \zeta)$, where k denotes the knot vector, ζ is the degree of the B-spline basis, and γ_l are the coefficients associated with the B-spline basis functions.

Based on these approaches, we can obtain other quantities as summarized in Table 1. Note that, in contrast to the closed-form expressions of the cumulative hazard function and survival function under the assumption that t follows an Exponential or Weibull distribution, approximating log baseline hazard function via the B-Splines method does not yield closed-form solutions for $H(t)$ or $S(t)$. Thus, for B-Splines approach, in this thesis, we approximate $H(t)$ from baseline hazard function $h(t)$ using the Gauss-Kronrod quadrature method with 15 nodes (see Equation 3). We choose this method because it provides a smooth and accurate approximation of the integral based on the baseline hazard function.

$$\int_{u=0}^{T_i} h_i(u)du \approx \frac{T_i}{2} \sum_{q=1}^{Q} w_q h_i \left(\frac{T_i(1 + w_q)}{2} \right) \tag{3}$$

where w_q are the quadrature weights and $Q = 15$ is the number of nodes.

Table 1: PH models under different baseline hazard assumptions

Assumption for t	**Exponential** $h_0(t) = \lambda$	**Weibull** $h_0(t) = \lambda v t^{v-1}$	**B-Splines Approximation** $h_0(t) = \exp\left(\sum_{l=1}^{L} \gamma_l B_l(t; k, \zeta)\right)$
Parameter(s)	Scale $\lambda > 0$	Scale $\lambda > 0$ Shape $v > 0$	Coefficients γ_l for B-spline basis
$h(t)$	$\lambda \exp(\eta_i)$	$\lambda v t^{v-1} \exp(\eta_i)$	$\exp\left(\sum_{l=1}^{L} \gamma_l B_l(t; k, \zeta) + \eta_i\right)$
$H(t)$	$\lambda \exp(\eta_i) t$	$\lambda \exp(\eta_i) t^v$	$\int_0^t h(u)\, du$
$S(t)$	$\exp(-\lambda \exp(\eta_i) t)$	$\exp(-\lambda \exp(\eta_i) t^v)$	$\exp(-H(t))$

Note: $B_l(t; k, \zeta)$ in Table 1 denote the l^{th} basis function ($l = 1, \ldots, L$) of a B-spline of degree ζ, evaluated at the knot vector $k = \{k_1, \ldots, k_J\}$. Each basis term has an associated coefficient γ_l, which is estimated from the data. In this thesis, we employ the `bSpline` function from the `splines2` R package (Wang & Yan, 2021) to construct cubic B-Splines (i.e., $\zeta = 3$) for modeling the log-baseline hazard as a smooth function of time. The B-spline basis excludes an intercept. The two boundary knots are positioned at 0 and the maximum of observing times. Thus, the estimated baseline hazard function can span the entire follow-up period, similar to the coverage of the classic Kaplan-Meier estimator. Additionally, two internal knots are placed at evenly spaced quantiles of the observed event times, which concentrates knots in regions with higher event density.

Maximum Partial Likelihood Estimator (MPLE)

Besides the MLE, another important estimator of $\boldsymbol{\beta}$ in the linear combination $\boldsymbol{\eta}$ is the *maximum partial likelihood estimator (MPLE)*, which has been introduced by (Cox, 1975) as Equation 4. Afterwards, this work has been updated by (Efron, 1977) to account for tied event times, i.e., multiple events occur at the same time (see Equation 5). Compared to the MLE in Equation 2, MPLE in Equation 4 and 5 can estimated $\boldsymbol{\beta}$ without specifying the baseline hazard $h_0(t)$ and thus become more computationally efficient and applicable than MLE.

$$\log PL(\mathcal{D}_i \mid \boldsymbol{\beta}) = \sum_{i=1}^{n} \delta_i \left[\mathbf{x}_i \boldsymbol{\beta} - \log \left(\sum_{j \in R_i} \exp(\mathbf{x}_j \boldsymbol{\beta}) \right) \right] \tag{4}$$

where R_i is the risk set when i-th event occurs.

$$\log PL_{\text{Efron}}(\boldsymbol{\beta}) = \sum_{j=1}^{J} \left[\sum_{i \in D_j} \mathbf{x}_i \boldsymbol{\beta} - \sum_{r=0}^{d_j - 1} \log \left(\sum_{k \in R_j} \exp(\mathbf{x}_k \boldsymbol{\beta}) - \frac{r}{d_j} \sum_{i \in D_j} \exp(\mathbf{x}_i \boldsymbol{\beta}) \right) \right] \tag{5}$$

where J is the number of distinct event times, d_j is the number of events at the j-th distinct event time, D_j is the set of tied events at the j-th distinct event time, and R_j is the risk set at the j-th distinct event time.

Assumption checking

As discussed earlier, the Proportional Hazards (PH) models and their associated functions/quantities rely on two key assumptions: non-informative censoring and proportional hazard. Non-informative censoring typically depends on how the data is recorded. In contrast, proportional hazard needs to be properly checked based on the output of the model. One common approach to assess the adequacy of the proportional hazards assumption for both categorical and continuous covariates is to plot the Schoenfeld residuals against the event times.

The Schoenfeld residual is a partial residual, defined as the vector at i-th event time,

$$\hat{r}_i = (\hat{r}_{i1}, \hat{r}_{i2}, \ldots, \hat{r}_{ip})^T$$

where

$$\hat{r}_{ij} = x_{ij} - \hat{E}(x_{ij} \mid R_i),$$

with

$$\hat{E}(x_{ij}|R_i) = \frac{\sum_{k \in R_i} x_{kj} \exp(\mathbf{x}_k \hat{\boldsymbol{\beta}})}{\sum_{k \in R_i} \exp(\mathbf{x}_k \hat{\boldsymbol{\beta}})}, \tag{6}$$

where R_i denotes the set of individuals at risk at i-th event time (i.e., under observation when the i-th event occurs; see Schoenfeld, 1982), and $\hat{\boldsymbol{\beta}}$ is the estimated coefficient vector.

In the presence of tied event times, Equation 6 will be updated according to Equation 5, ending up with Equation 7.

$$\hat{E}_{ij}^{\text{Efron}} = \frac{1}{d_i} \sum_{r=1}^{d_i} \frac{\sum_{k \in R_i} x_{kj} e^{\mathbf{x}_k \boldsymbol{\beta}} - \frac{r-1}{d_i} \sum_{k \in D_i} x_{kj} e^{\mathbf{x}_k \hat{\boldsymbol{\beta}}}}{\sum_{k \in R_i} e^{\mathbf{x}_k \hat{\boldsymbol{\beta}}} - \frac{r-1}{d_i} \sum_{k \in D_i} e^{\mathbf{x}_k \hat{\boldsymbol{\beta}}}} \tag{7}$$

where R_i is the risk set at i-th distinct event time; D_i is the set of tied events at i-th distinct event time; $d_i = |D_i|$ is the number of tied events at i-th distinct event time; and $\hat{\boldsymbol{\beta}}$ is the estimated coefficient vector.

If the proportional hazards assumption stands, the Schoenfeld residuals have a sample path of a random walk, i.e., their values fluctuate randomly over time without exhibiting systematic trends. Therefore, they are useful in assessing time trend or lack of proportionality.

Related work

Building on the background of proportional hazards (PH) models and their estimators, we now review existing methods for addressing the mentioned problem in the section of Introduction

when detecting interaction effects in survival or regression settings. Two prominent approaches in regression setting are the *two-stage method* and the *linked shrinkage prior*. The two-stage method begins with a model containing only main effects. Interaction terms are added only if their corresponding main effects are statistically significant (Hao & Zhang, 2017). While this approach helps mitigate the quadratic growth of interaction terms as the number of main terms increases, it may be less applicable for small datasets due to concerns about double-dipping—using the same data to identify significant main effects and then to test interactions.

In contrast, the linked shrinkage prior offers a more integrated solution. To be more concrete, as shown in Equation 8, van de Wiel et al. (2024) have linked main and interaction effects through the product $\tau_j \tau_k$ in the prior for β_{jk}. This setting allows symmetric treatment of strong and weak main effects. Besides, they also include a shared shrinkage parameter τ_{int} for all interaction terms, which follows $\mathcal{U}(0.01, 1)$. This shared shrinkage parameter τ_{int} on the one hand aligns with the common belief that interaction effects are, on average, weaker than main effect. On the other hand, a lower bound on τ_{int} prevents complete shrinkage to zero, which is undesirable in sparse settings. In the presence of categorical covariates, van de Wiel et al. (2024) modify the summation over j, k to exclude interactions between different levels of the same factor. In other words, they consider only interaction effects between distinct factors, explicitly excluding interactions among mutually exclusive levels within a single categorical variable. Together, the linked shrinkage prior demonstrates clear advantages in detecting interaction effects in regression models. Note that other methods, such as the horseshoe prior—originally developed for high-dimensional data analysis rather than detecting interaction effects—can also be applied in this context.

$$
\begin{aligned}
Y_i &= \alpha + \sum_{j=1}^{p} \beta_j x_{ij} + \sum_{j<k} \beta_{jk} x_{ij} x_{ik} + \varepsilon_i \\
\alpha &\sim \mathcal{N}(0, 10^2) \\
\beta_j &\sim \mathcal{N}(0, \sigma^2 \tau_j) \\
\beta_{jk} &\sim \mathcal{N}(0, \sigma^2 \tau_j \tau_k \tau_{\text{int}}) \\
\varepsilon_i &\sim \mathcal{N}(0, \sigma^2) \\
\tau_j &\sim \mathcal{C}^+(0, 1) \\
\tau_{\text{int}} &\sim \mathcal{U}(0.01, 1) \\
\sigma^2 &\sim \text{IG}(1, 0.001)
\end{aligned}
\tag{8}
$$

where $\mathcal{D} = \{(Y_i, \mathbf{x}_i)\}_{i=1}^{n}$ is n independent observations; Y_i is the response variable, and $\mathbf{x}_i$ is the covariate vector; x_{ij} is the jth covariate for individual i, with $j = 1, \ldots, p$.

The Current Study

In this study, we apply the linked shrinkage prior to Bayesian proportional hazard (PH) models to enable accurate estimation and selection of interaction effects in explaining graft loss in kidney transplantation. To this end, we employ both the maximum likelihood estimator (MLE) in Equation 2 and the maximum partial likelihood estimator (MPLE) in Equation 5. When employing MLE, for priors of the coefficients $\boldsymbol{\beta}$ in $\boldsymbol{\eta}$, we adopt the prior structure from Equation 8 by replacing the response variable Y_i with the linear combination of covariates $\boldsymbol{\eta}$. For ancillary parameters that employed to model $h_0(t)$, in this thesis, we specify their prior distributions as shown in Table 2.

Table 2: Prior distributions for ancillary parameters under different baseline hazard specifications

Parameter	**Exponential**	**Weibull**	**B-Splines Approximation**
Scale parameter λ	Lognormal$(0, 100)$	Lognormal$(0, 100)$	—
Shape parameter v	—	Lognormal$(0, 100)$	—
B-spline coefficients γ_l	—	—	$\mathcal{N}(0, 100)$

Regarding MPLE, as mentioned before, MPLE in Equations 4 and 5 effectively removes the nuisance baseline hazard function. Although this simplification improves computational efficiency, it poses challenges for the classical Bayesian method. That is, the posterior distribution is derived from the full likelihood estimator and the prior distribution via $P(\theta \mid \text{Data}) \propto L(\text{Data} \mid \theta)P(\theta)$. Nonetheless, Kalbfleisch (1978) demonstrated that using MPLE in Partial Bayesian PH models approximates the marginal posterior of $\boldsymbol{\beta}$ which assumes that $H_0(t)$ follows a gamma process prior. This result was later extended by Kim and Kim (2009) to accommodate tied event times. Together, there is strong theoretical justification for applying the linked shrinkage prior to Bayesian PH models by incorporating MPLE. Thus, we employ MPLE in this study and define the priors for $\boldsymbol{\beta}$ in the same way as in the MLE-based application.

Validating reformulated method on simulated study

Simulated data generation

To validate the performance of applying linked shrinkage prior to proportional hazards (PH) models, we conduct a simulation study using data generated via the `simsurv` package (Brilleman et al., 2021). This setup allows us to flexibly specify both the baseline hazard function $h_0(t)$ and the linear combinations of covariates $\boldsymbol{\eta}$. For the baseline hazard, we define the density

function $f_0(t)$ using a Weibull distribution with parameters $\lambda = 2$ and $v = 10$. To reflect realistic conditions, we set the event occurrence rate to 75% and adjust the censoring mechanism accordingly. For the linear combination of covariates $\boldsymbol{\eta}$, we generate $p = 10$ moderately collinear continuous covariates for $n = 200$ samples, resulting in $q = \frac{p(p-1)}{2} = 45$ two-way interaction terms.

We account for three possible scenarios of interaction terms regarding the main effects: interaction effects (1) between two non-zero main effects; (2) between one non-zero and one zero main effect; and (3) between two zero main effects. To this end, five main effects are assigned non-zero values (ranging from small to large), while the remaining five are set to zero. A separate testing dataset of $n' = 100$ samples is generated using the same configuration for the sake of prediction. The detailed information about the values of $\boldsymbol{\beta}$ in the simulation setting is as follows:

- **Sample size:** $n = 200$ (estimation/training), $n' = 100$ (testing)
- **Covariates:** $p = 10$ continuous variables
- **Interactions:** $q = \frac{p(p-1)}{2} = 45$ two-way interactions
- **Main effects:** $\beta_1 = 0.5$, $\beta_2 = -0.4$, $\beta_3 = 0.3$, $\beta_4 = 0.2$, $\beta_5 = 0.1$; β_6 to $\beta_{10} = 0$
- **Non-zero interactions:**
  - Between two non-zero main effects: $\beta_{12} = 0.3$, $\beta_{13} = 0.2$, $\beta_{15} = -0.1$, $\beta_{23} = -0.1$, $\beta_{24} = 0.3$, $\beta_{34} = 0.2$
  - Between one non-zero and one zero main effect: $\beta_{16} = 0.3$, $\beta_{19} = 0.2$, $\beta_{36} = -0.2$, $\beta_{47} = 0.1$
  - Between two zero main effects: $\beta_{67} = 0.3$, $\beta_{68} = 0.1$, $\beta_{9,10} = -0.2$
- **Remaining interactions:** The other $45 - 12 = 33$ interaction terms are set to zero.
- **Covariate generation:** Each covariate vector $\mathbf{x}_i = (x_{i1}, \ldots, x_{i10})$ is drawn from a multivariate normal distribution: $\mathbf{x}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, where $\Sigma_{jj} = 1$ and $\Sigma_{jk} = 0.3$ for $j \neq k$.

Performance metrics

With the simulated data generated as described previously, we evaluate our application of linked shrinkage prior to Bayesian proportional hazards (PH) models with the following metrics:

1. **Parameter estimation:**
   - *Bias:*

$$\text{Bias} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}^{(b)} - \beta\right)$$

- *Variance:*

$$\text{Variance} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}^{(b)} - \bar{\hat{\beta}}\right)^2, \quad \bar{\hat{\beta}} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}^{(b)}$$

- *Root Mean Squared Error (rMSE):*

$$\text{rMSE} = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left(\hat{\beta}^{(b)} - \beta\right)^2}$$

Where B denotes the number of such simulation procedures in the study; $\hat{\beta}^b$ represents the median value of the corresponding posterior distribution in the b-th simulation; and β is the ground-truth value specified in the simulation. Note that, due to the bias–variance trade-off, a well-performing model should aim to minimize rMSE rather than bias and variance simultaneously. In this study, we report bias and variance to decompose rMSE.

2. **Variable selection:**

- *Type I Error/False Discovery Rate (FDR):*

$$\text{FDR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

- *Power (1 - Type II Error):*

$$\text{Type II Error} = \frac{\text{FN}}{\text{FN} + \text{TP}}$$

Where FP, TN, FN, and TP are defined in Table 3. The parameter inference is based on its corresponding 95% credible interval.

Table 3: Confusion Matrix

	Actually Positive	**Actually Negative**
Predicted Positive	True Positives (TP)	False Positives (FP)
Predicted Negative	False Negatives (FN)	True Negatives (TN)

3. **Prediction:**

- *Concordance Index (C-index):* Measures the model's ability to correctly rank survival times. It is defined as the proportion of all usable patient pairs for which the predicted and observed outcomes are concordant. A higher C-index indicates better discriminative performance.

Note that the prediction metric, i.e., C-index, presented in this study are not the primary focus. They are included for two main reasons: (1) to serve as a complementary elements that support the overall analysis; (2) to evaluate of model performance in the context of real-world data.

Results of method performance: different ways of specifying $h_0(t)$

Using simulated data and the evaluation metrics described earlier, we first examine the performance of our applied linked shrinkage prior to proportional hazards (PH) models under different specifications of $h_0(t)$. That is, as mentioned before, we first consider full Bayesian parametric models based on MPLE: (1) assuming $T^* \sim \text{Exp}(\lambda)$ (from now on `Exponential`); (2) assuming $T^* \sim \text{Weibull}(\lambda, v)$ (from now on `Weibull`); (3) approximating $\log h_0(t)$ using B-Splines (from now on `B-Splines`). Then, we consider partial Bayesian model based on MPLE (from now on `PartialLikelihood`). Since the simulated true event times T^* follow a Weibull distribution, we expect `Weibull` to achieve the best performance across all evaluation criteria (Lázaro et al., 2021). Conversely, `Exponential` is anticipated to perform the worst due to its misspecified distributional assumption.

As a result (see Figure 1), `Weibull` indeed demonstrates the best performance in estimating $\boldsymbol{\beta}$, selecting relevant terms, and achieving high prediction accuracy. In contrast, `Exponential` exhibites the worst performance across these metrics. `B-Splines` and `PartialLikelihood` fall between these two extremes. We also visualize one of the estimated baseline hazard functions obtained from `Exponential`, `Weibull` and `B-Splines` (see Figure 2) and observed a similar result. Note that the good performance of `Weibull`, `B-Splines`, and `PartialLikelihood` comes at the cost of slightly higher variance for $\hat{\boldsymbol{\beta}}$ compared to `Exponential`. This is reasonable given their much smaller bias and lower rMSE relative to `Exponential`. Together, our applied linked shrinkage prior to proportional hazards (PH) models is correctly specified. However, given that, in practice, the distribution of true event times T^* remains commonly unknown, in the following sections we only focus on `B-Splines` and `PartialLikelihood`.

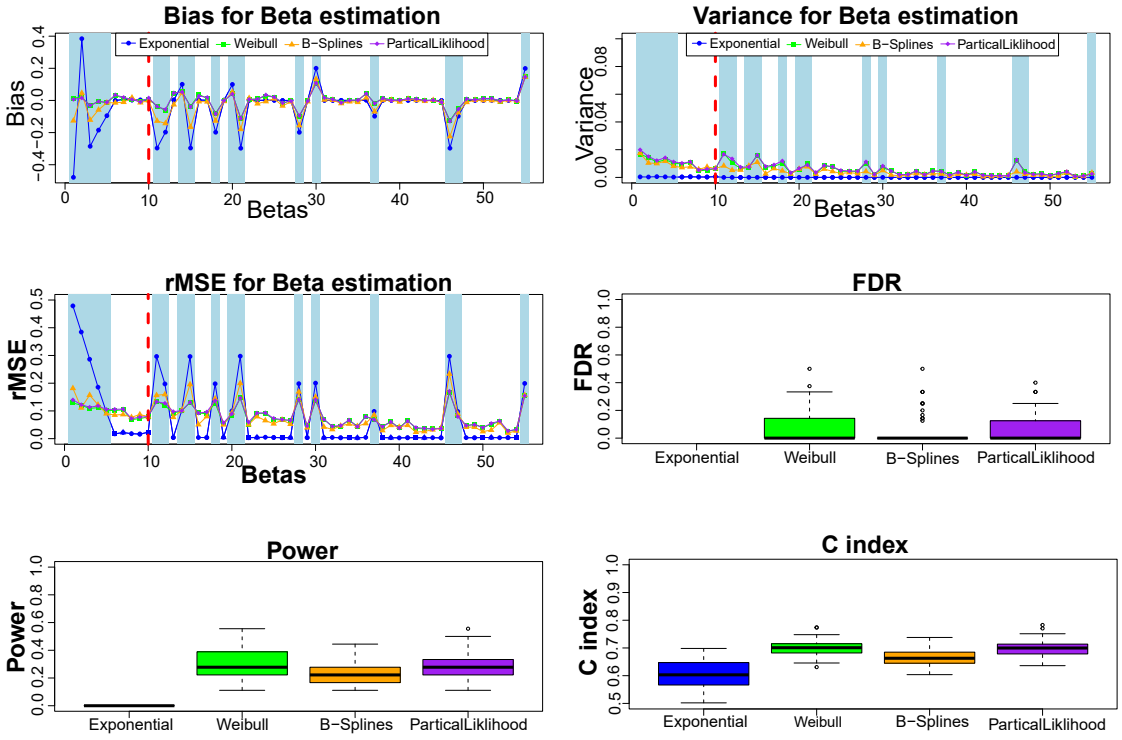


Figure 1: Performance of the applied linked shrinkage prior to Bayesian PH models across different specifications of $h_0(t)$. To the left of the red dashed vertical line are the main effects, while to the right are the interaction effects. `Weibull` specification demonstrates the best performance in estimating $\boldsymbol{\beta}$, selecting relevant terms, and achieving high prediction accuracy. `Exponential` specification exhibits the weakest performance across these metrics, whereas `B-Splines` and `PartialLikelihood` fall between these two extremes.

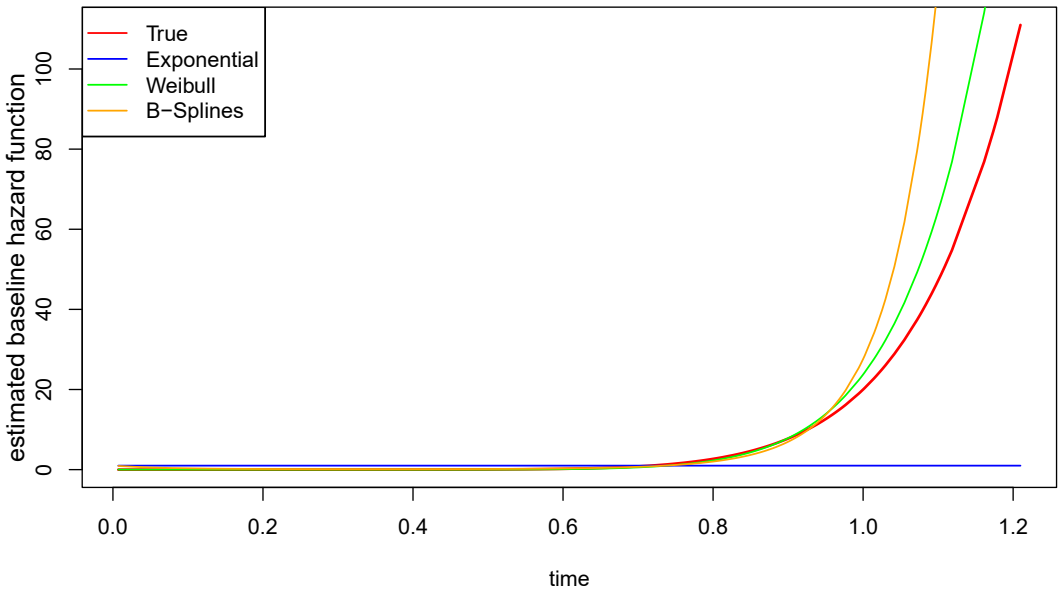


Figure 2: An example of $\hat{h}_0(t)$ from different methods. `Weibull` demonstrates the best performance in estimating $h_0(t)$. `Exponential` exhibites the weakest performance in estimating $h_0(t)$. `B-Splines` falls between these two extremes.

Results of method performance: linked shrinkage vs. non-informative prior distributions

With `B-Splines` and `PartialLikelihood`, we now move on to validating the advantage of the linked shrinkage prior in detecting interaction effects using the simulated data. To this end, we employ the corresponding models that incorporates non-informative prior as a benchmark. For the non-informative prior, we simply set all prior distributions for $\boldsymbol{\beta}$ to $\mathcal{N}(0, 100)$, including those for both main effects and interaction effects. From now on, the corresponding models are `B-Splines_no` and `PartialLikelihood_no`. Note that the estimated coefficients $\hat{\boldsymbol{\beta}}$ obtained under this non-informative prior setting is equivalent to those derived from the classical cox regression from a frequentist perspective.

As shown in Figure 3, the `B-Splines` and `PartialLikelihood` consistently outperform `B-Splines_no` and `PartialLikelihood_no`. Again, no consistent performance difference is observed between `B-Splines` and `PartialLikelihood`. We also visualize an example of an estimated baseline hazard function $\hat{h}_0$ obtained from `B-Splines` and `B-Splines_no`. Accordingly, `B-Splines` performs better than `B-Splines_no` in estimating baseline hazard function. Together,

`B-Splines` and `PartialLikelihood` outperform their counterparts that employ non-informative priors in detecting interaction effects.

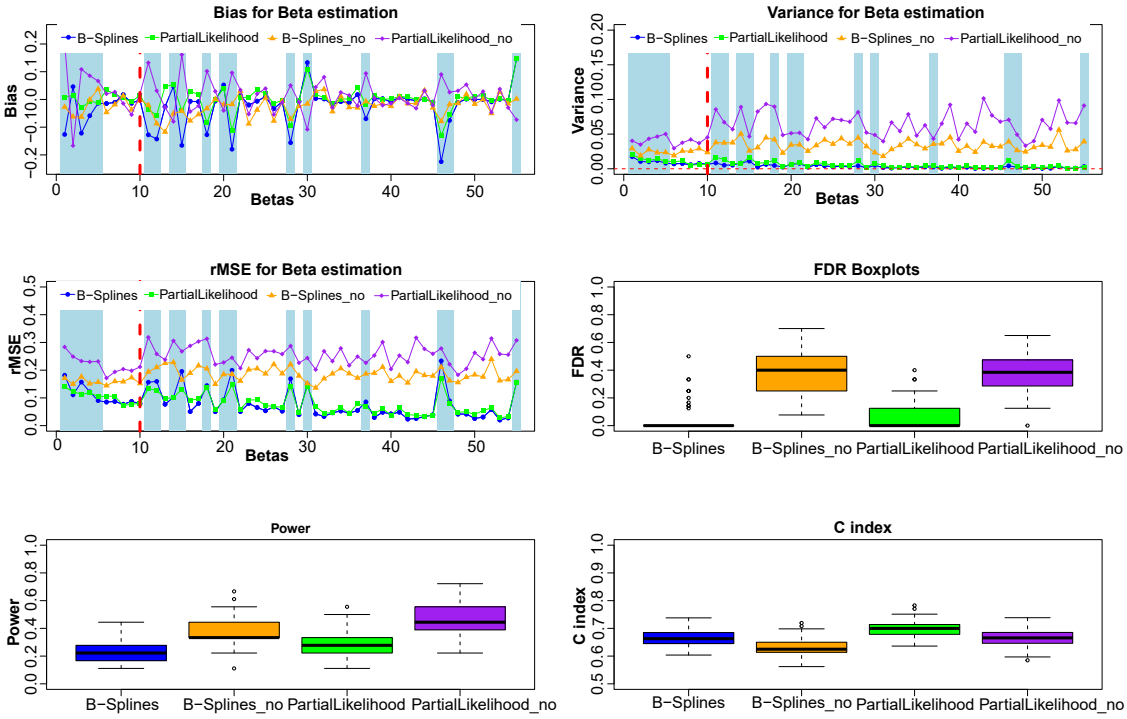


Figure 3: Performance of the applied linked shrinkage prior to Bayesian PH models compared to non-informative prior. To the left of the red dashed vertical line are the main effects, while to the right are the interaction effects. `B-Splines` and `PartialLikelihood` consistently outperform outperform their counterparts that employ non-informative priors across all evaluation metrics. No consistent difference is observed between `B-Splines` and `PartialLikelihood`.

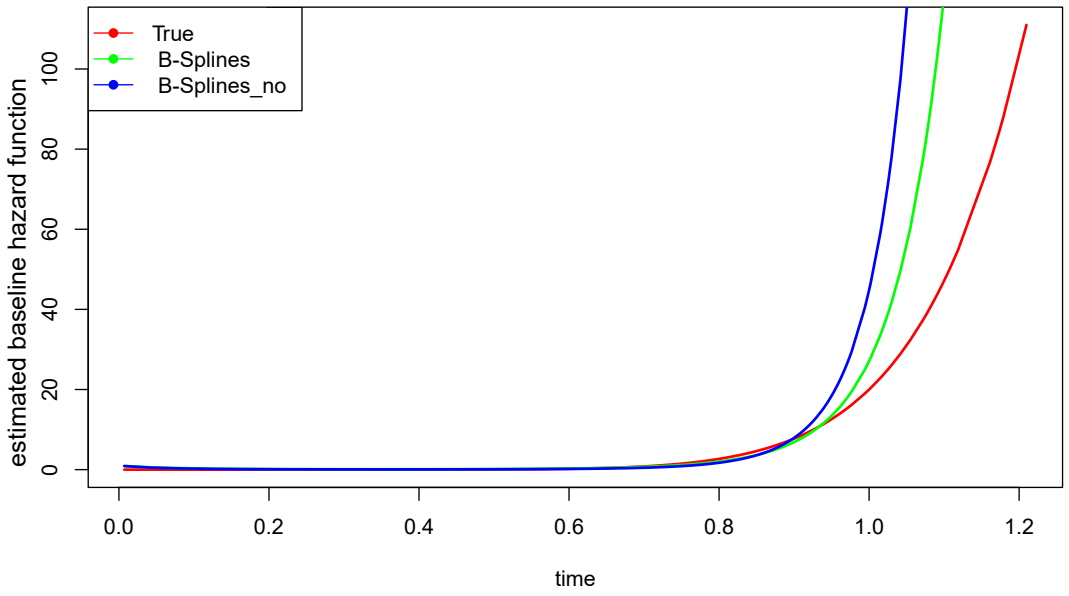


Figure 4: An example of $\hat{h}_0(t)$ from `B-Splines` and `B-Splines_no`. `B-Splines` performs better than `B-Splines_no` in estimating baseline hazard function.

Detecting interaction effects in kidney transplantation

With the validated `B-Splines` and `PartialLikelihood`, it is now time to proceed with the following: (1) apply `B-Splines` and `PartialLikelihood` on real-world data; and (2) focus on our original aim of this study, i.e., accurate estimation and selection of interaction effects to better explain graft loss in kidney transplantation. To this end, we also include `B-Splines_no` and `PartialLikelihood_no` as a benchmark. Regarding data, we obtain kidney transplantation records from the national registry of the Dutch National Transplant Foundation (NOTR), which is the exact dataset employed in Steenvoorden, Evers, et al. (2024) and Steenvoorden, Vogt, et al. (2024).

Data

With the kidney transplantation records obtained from NOTR, we first follow the same procedure as described in Steenvoorden, Vogt, et al. (2024) for data preprocessing. In the end, the preprocessed data (from now on `complexData`) includes: (1) 3008 observations, with 310 recorded events of graft loss; and (2) 14 covariates, comprising 8 continuous and 6 categorical covariates. The

continuous covariates are: `recipientage`, `dialysesvintage`, `donorage`, `lastcreatininemgdl`, `cold_ischaemic_period`, `HLA_mismatch`, `donorBMI`, and `recipient_BMI`. The categorical covariates are the following: `recipientsex` (`Female`/`Male`), `donorsex` (`Female`/`Male`), `smoking` (`Yes`/`No`), `typecadavericDBDDCD` (`DCD`/`DBD`), `initialPrimaryDiseaseET` (`Diabetes`/`Others`), and `donorcauseofdeath` (`Cardiac`/`Others`).

We also consult domain experts to exclude certain covariates based on expert knowledge, leading to a simple dataset (from now on `simpleData`). The excluded covariates are the following: `smoking` (`Yes`/`No`), `donorcauseofdeath` (`Cardiac`/`Others`), `dialysesvintage`, `recipientsex`, and `recipient_BMI`. The `simpleData` is included to achieve good convergence for PH models that incorporate non-informative priors.

Additionally, we include extra noise variables (i.e., `noise1` and `noise2`) for both `complexData` and `simpleData` due to their potential help in detecting false positives. Together, the final number of covariates for `complexData` is $p = 14 + 2 = 16$, while the final number of covariates for `simpleData` is $p = 9 + 2 = 11$. We consider all two-way interactions among covariates, resulting in $q = \binom{16}{2} = 120$ interaction parameters for `complexData` and $q = \binom{11}{2} = 55$ for `simpleData`. In total, there are 136 terms for `complexData` and 66 terms for `simpleData`.

To avoid zero or extremely early event times, observing time T_i in Equation 1 is shifted slightly to the right by adding 0.5. In this analysis, graft loss is considered the event of interest, while death is treated as a censored observation.

Descriptive data analysis for real-world data

Before analyzing the real-world dataset, we first visualize Kaplan–Meier curve of the kidney transplantation data (see Figure 5) and calculate the median follow-up time to evaluate the maturity and reliability of the follow-up for graft loss via the reverse Kaplan–Meier method (see Figure 6). Accordingly, there is a sharp decline in survival probability at the beginning, and the median follow-up time is 1437 days (approximately 3.9 years). This indicates that at least 50% of the patients were followed for this duration.

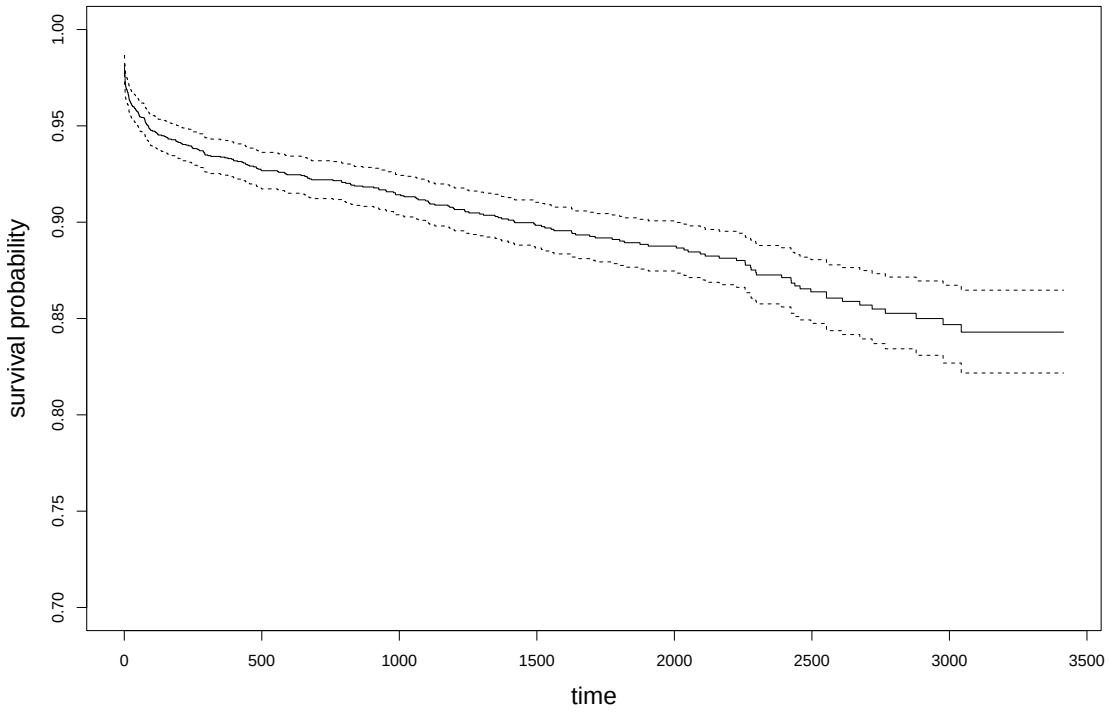


Figure 5: Kaplan–Meier curve of the real-world kidney transplantation dataset. There is a sharp decline in survival probability at the beginning.

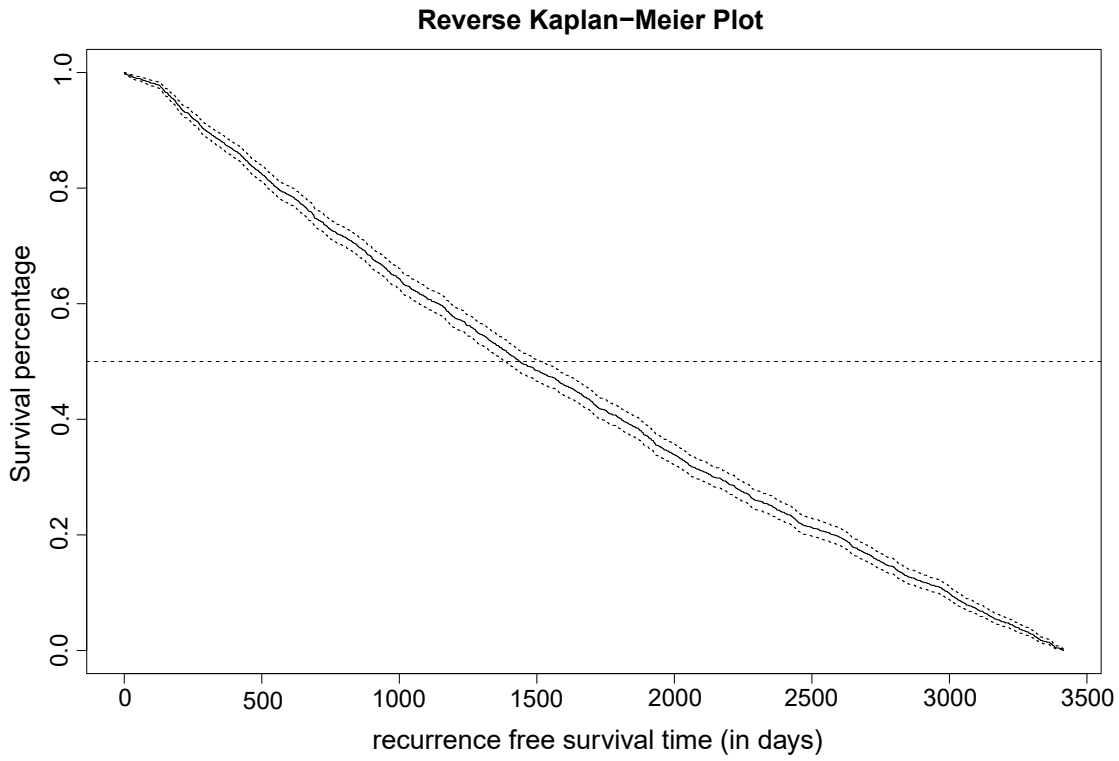


Figure 6: Reverse Kaplan–Meier curve of real-world kidney transplantation dataset. The median follow-up time is 1437 days (approximately 3.9 years).

We also assess collinearity among the covariates by analyzing pairwise correlations and evaluating the occurrence of each category in categorical covariates, separately for graft loss and censored events. As a result, we did not detect any strong collinearity, as shown in Figure 7. The occurrence of each category for graft loss and censored events is also within a reasonable range (see Table 4).

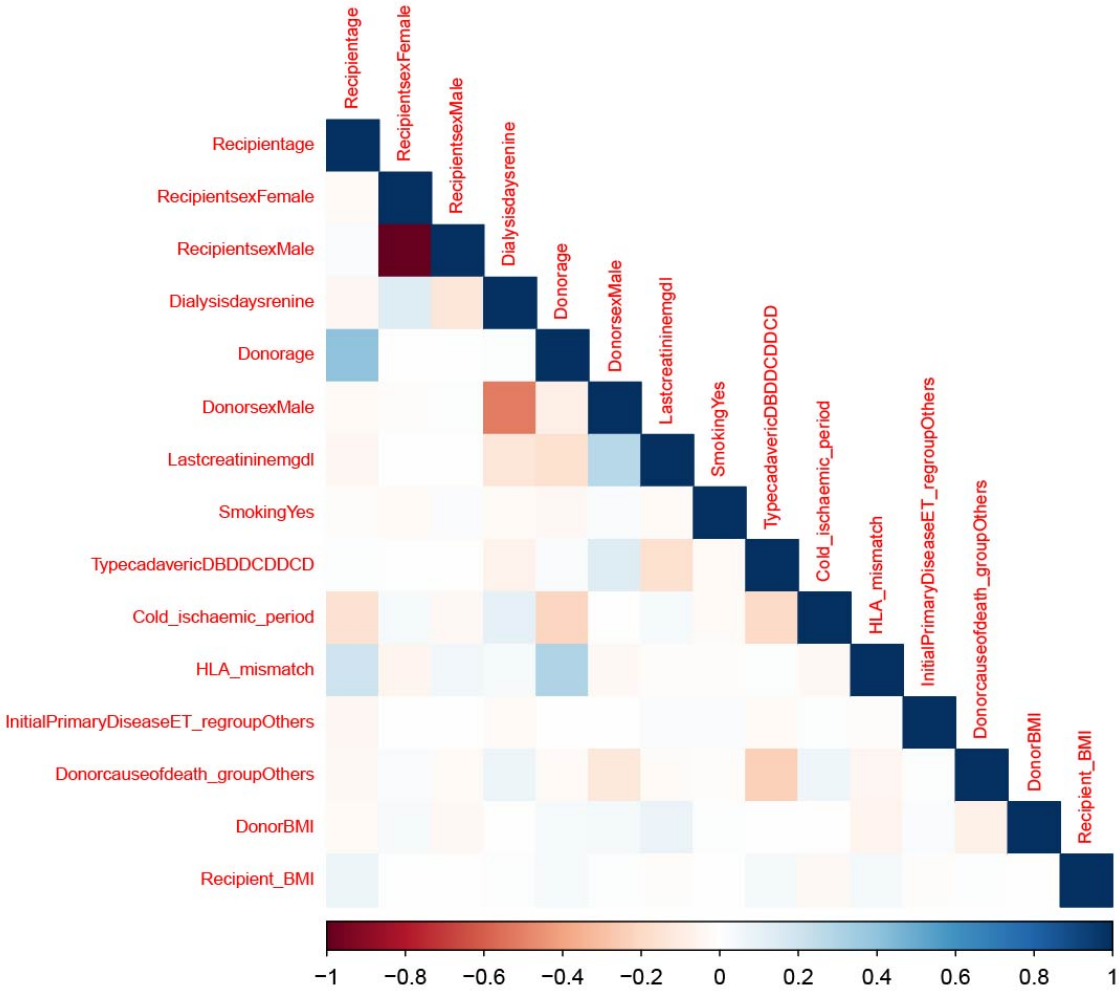


Figure 7: Pair-wise correlations among covariates in the dataset. No strong collinearity has been detected.

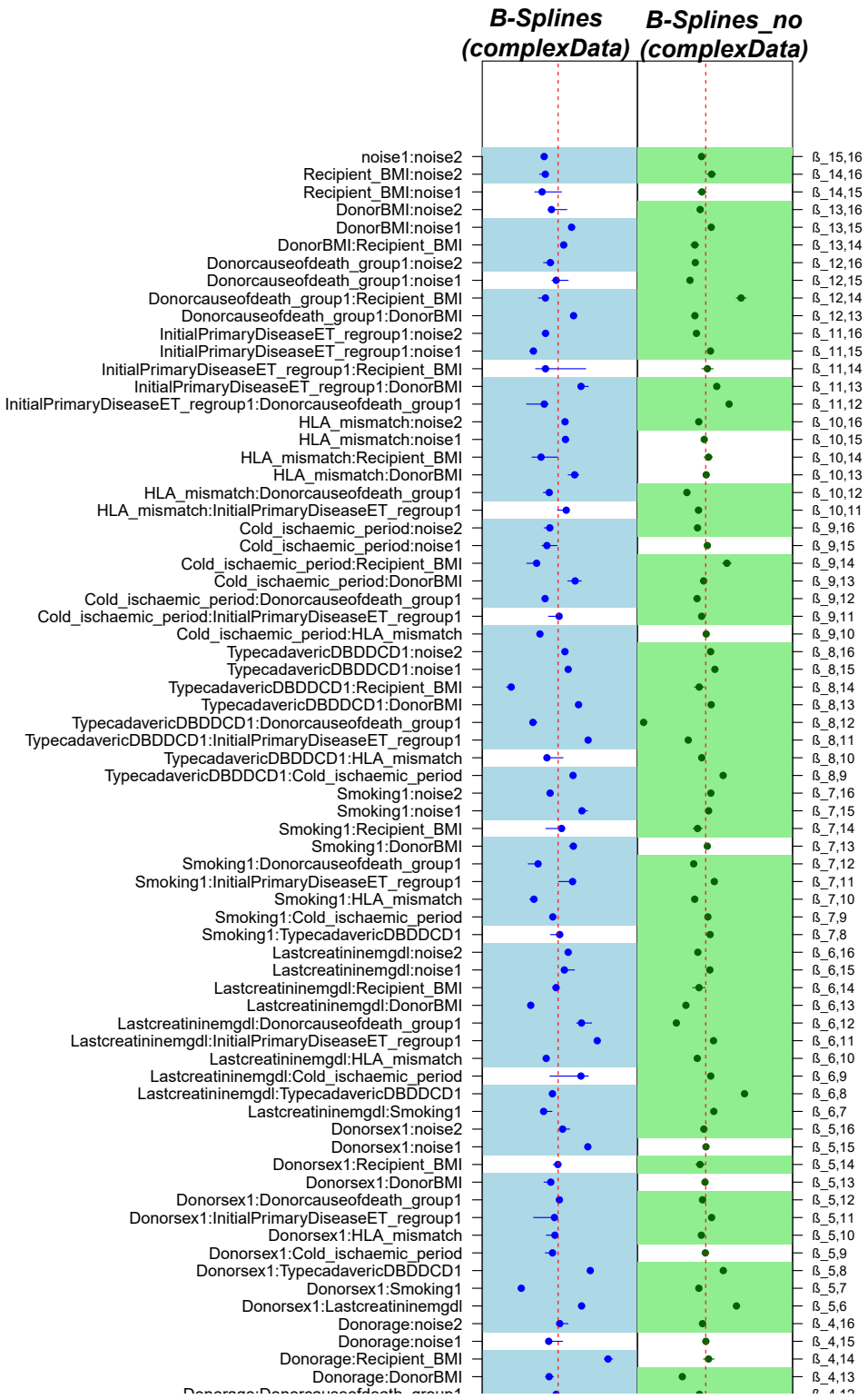
Table 4: Summary of the occurrence of each category in categorical covariates, separately for graft loss and censored events.

	Recipient sex		Donor sex		Smoking		Typecadaveric DBDDCD		Initial Primary Disease		Donor cause of death	
	Female	Male	Female	Male	Yes	No	DCD	DBD	Diabetes	Others	Cardiac	Others
Graft loss	124	186	161	149	184	149	175	135	22	288	54	256
Censored	979	1719	1193	1505	1575	1123	1619	1079	223	2475	487	2211

Results of kidney transplantation data

As we mentioned before, in this section, our methodological aim is to apply `B-Splines` and `PartialLikelihood` to real-world kidney transplantation data. To this end, first, as benchmark models, we include `B-Splines_no` and `PartialLikelihood_no`. Second, we generate some noise variables and include them in the kidney transplantation dataset. Last, we employ both `B-Splines` and `PartialLikelihood` to analyze `complexData` and `simpleData`.

As a result, for `complexData`, `B-Splines` and `B-Splines_no` detect a large number of covariates (see Figure 8), including interactions between `noise1`, `noise2`, and other covariates. A similar pattern is observed when applying `B-Splines` and `B-Splines_no` to `simpleData` (see Figure 9). Taken together, these findings suggest that `B-Splines` is not reliable when applied to kidney transplantation data. Therefore, in the remainder of this section, we focus on `PartialLikelihood`.


B-Splines (complexData)
B-Splines_no (complexData)
noise1:noise2
Recipient_BMI:noise2
Recipient_BMI:noise1
DonorBMI:noise2
DonorBMI:noise1
DonorBMI:Recipient_BMI
Donorcauseofdeath_group1:noise2
Donorcauseofdeath_group1:noise1
Donorcauseofdeath_group1:Recipient_BMI
Donorcauseofdeath_group1:DonorBMI
InitialPrimaryDiseaseET_regroup1:noise2
InitialPrimaryDiseaseET_regroup1:noise1
InitialPrimaryDiseaseET_regroup1:Recipient_BMI
InitialPrimaryDiseaseET_regroup1:DonorBMI
InitialPrimaryDiseaseET_regroup1:Donorcauseofdeath_group1
HLA_mismatch:noise2
HLA_mismatch:noise1
HLA_mismatch:Recipient_BMI
HLA_mismatch:DonorBMI
HLA_mismatch:Donorcauseofdeath_group1
HLA_mismatch:InitialPrimaryDiseaseET_regroup1
Cold_ischaemic_period:noise2
Cold_ischaemic_period:noise1
Cold_ischaemic_period:Recipient_BMI
Cold_ischaemic_period:DonorBMI
Cold_ischaemic_period:Donorcauseofdeath_group1
Cold_ischaemic_period:InitialPrimaryDiseaseET_regroup1
Cold_ischaemic_period:HLA_mismatch
TypecadavericDBDDCD1:noise2
TypecadavericDBDDCD1:noise1
TypecadavericDBDDCD1:Recipient_BMI
TypecadavericDBDDCD1:DonorBMI
TypecadavericDBDDCD1:Donorcauseofdeath_group1
TypecadavericDBDDCD1:InitialPrimaryDiseaseET_regroup1
TypecadavericDBDDCD1:HLA_mismatch
TypecadavericDBDDCD1:Cold_ischaemic_period
Smoking1:noise2
Smoking1:noise1
Smoking1:Recipient_BMI
Smoking1:DonorBMI
Smoking1:Donorcauseofdeath_group1
Smoking1:InitialPrimaryDiseaseET_regroup1
Smoking1:HLA_mismatch
Smoking1:Cold_ischaemic_period
Smoking1:TypecadavericDBDDCD1
Lastcreatininemgdl:noise2
Lastcreatininemgdl:noise1
Lastcreatininemgdl:Recipient_BMI
Lastcreatininemgdl:DonorBMI
Lastcreatininemgdl:Donorcauseofdeath_group1
Lastcreatininemgdl:InitialPrimaryDiseaseET_regroup1
Lastcreatininemgdl:HLA_mismatch
Lastcreatininemgdl:Cold_ischaemic_period
Lastcreatininemgdl:TypecadavericDBDDCD1
Lastcreatininemgdl:Smoking1
Donorsex1:noise2
Donorsex1:noise1
Donorsex1:Recipient_BMI
Donorsex1:DonorBMI
Donorsex1:Donorcauseofdeath_group1
Donorsex1:InitialPrimaryDiseaseET_regroup1
Donorsex1:HLA_mismatch
Donorsex1:Cold_ischaemic_period
Donorsex1:TypecadavericDBDDCD1
Donorsex1:Smoking1
Donorsex1:Lastcreatininemgdl
Donorage:noise2
Donorage:noise1
Donorage:Recipient_BMI
Donorage:DonorBMI
ß_15,16
ß_14,16
ß_14,15
ß_13,16
ß_13,15
ß_13,14
ß_12,16
ß_12,15
ß_12,14
ß_12,13
ß_11,16
ß_11,15
ß_11,14
ß_11,13
ß_11,12
ß_10,16
ß_10,15
ß_10,14
ß_10,13
ß_10,12
ß_10,11
ß_9,16
ß_9,15
ß_9,14
ß_9,13
ß_9,12
ß_9,11
ß_9,10
ß_8,16
ß_8,15
ß_8,14
ß_8,13
ß_8,12
ß_8,11
ß_8,10
ß_8,9
ß_7,16
ß_7,15
ß_7,14
ß_7,13
ß_7,12
ß_7,11
ß_7,10
ß_7,9
ß_7,8
ß_6,16
ß_6,15
ß_6,14
ß_6,13
ß_6,12
ß_6,11
ß_6,10
ß_6,9
ß_6,8
ß_6,7
ß_5,16
ß_5,15
ß_5,14
ß_5,13
ß_5,12
ß_5,11
ß_5,10
ß_5,9
ß_5,8
ß_5,7
ß_5,6
ß_4,16
ß_4,15
ß_4,14
ß_4,13

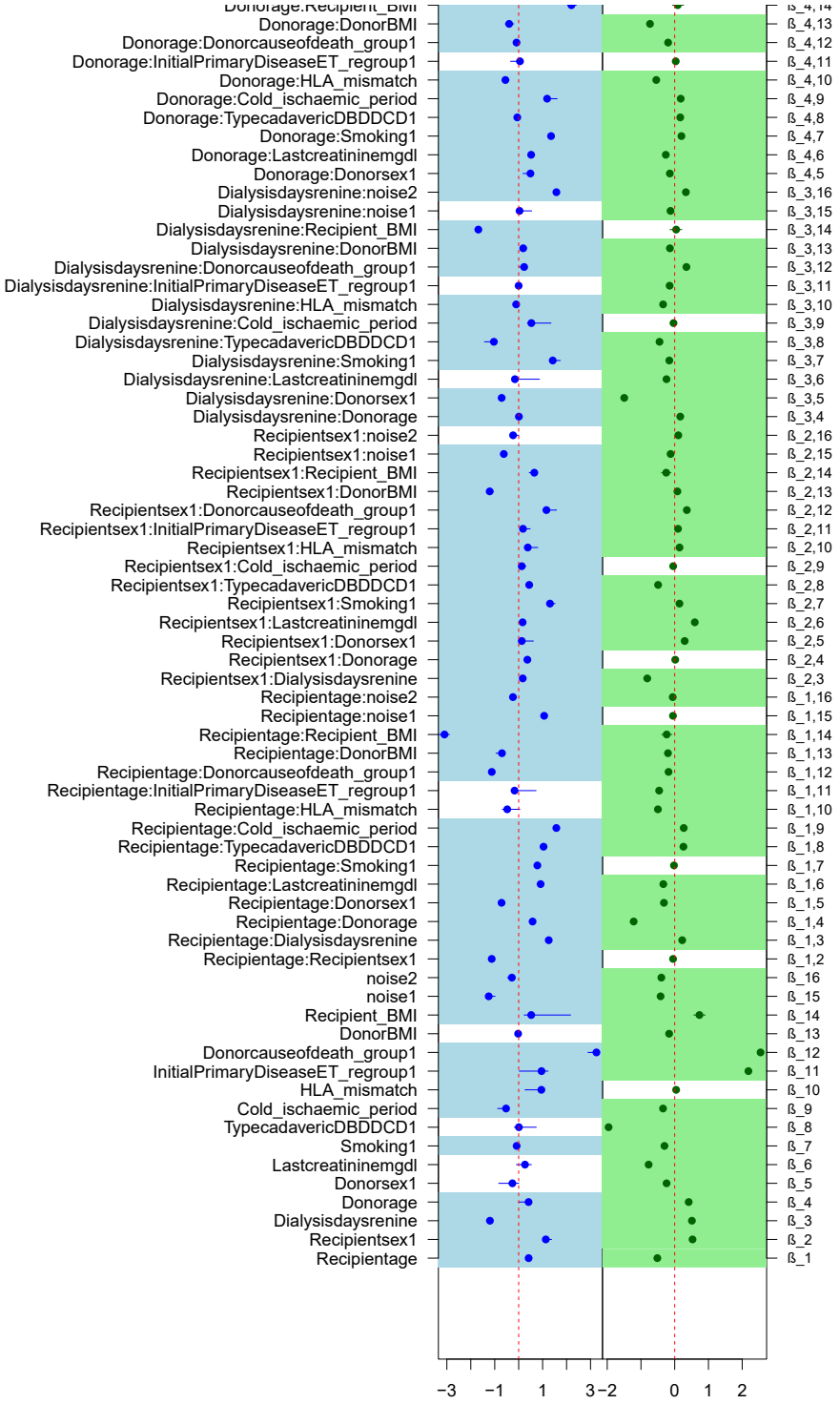


Figure 8: Result of `complexData` from `B-Splines` and `B-Splines_no`. Shadowed terms are the selected effects by the corresponding methods. `B-Splines` and `B-Splines_no` give similar results and detect a lot of covariates, including the interactions between `noise1`, `noise2` and other covariates. Together, both `B-Splines` and `B-Splines_no` are not trustworthy when analyzing `complexData`.

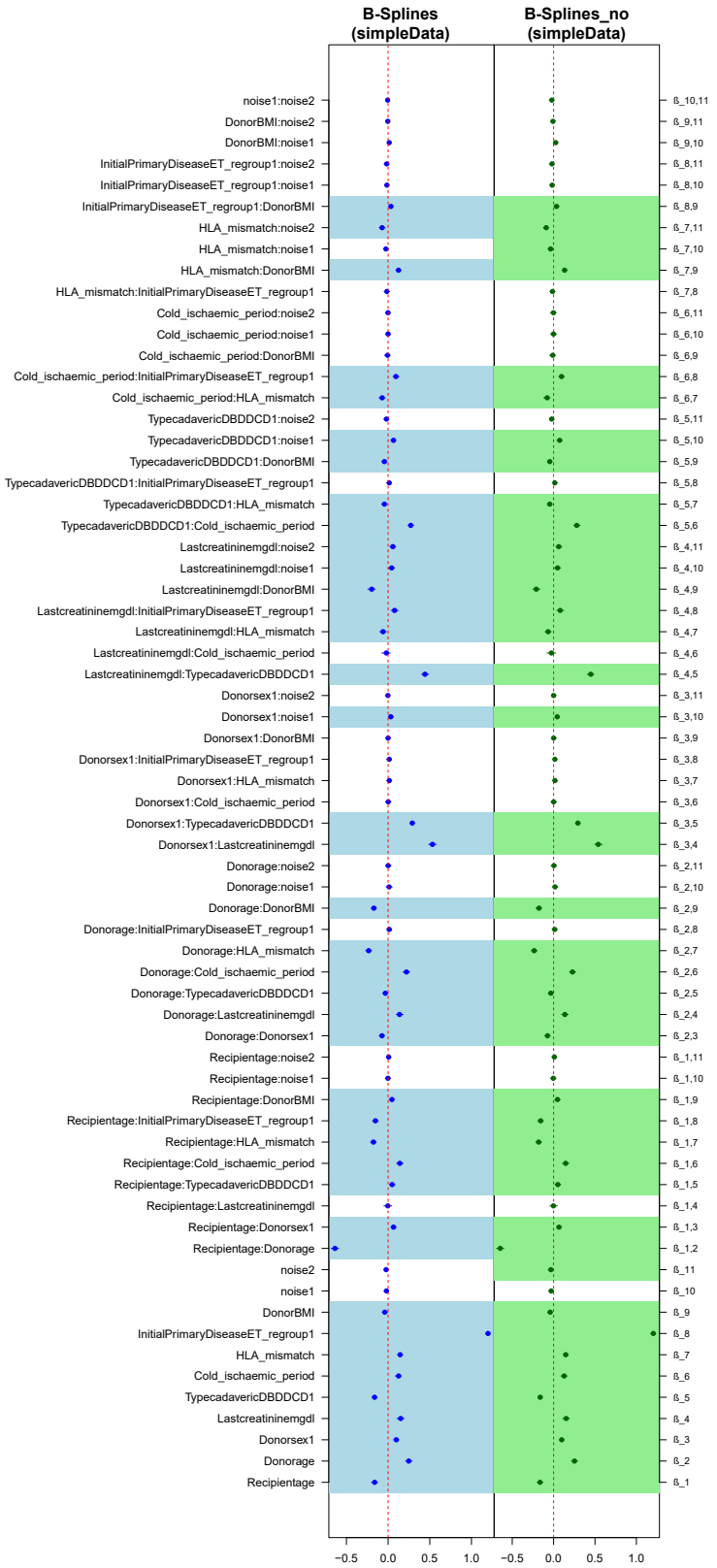


Figure 9: Result of `simpleData` from `B-Splines` and `B-Splines_no`. Shadowed terms are the selected effects by the corresponding methods. `B-Splines` and `B-Splines_no` detect a lot of covariates, including the main effects of `noise1` and `noise2` and their interactions with other covariates. Thus, both `B-Splines` and `B-Splines_no` are not trustworthy when analyzing `simpleData`.

We also visualize one of the estimated baseline hazard functions in Figure 10. Accordingly, graft loss mainly occurs within the first year after kidney transplantation. Note that since the estimated baseline hazard functions obtained from `B-Splines`, `B-Splines_no` are identical, we only present one plot for these methods. Such identical $\hat{h}_0(t)$ across `B-Splines` and `B-Splines_no` is inconsistent with the results observed in the simulation study.

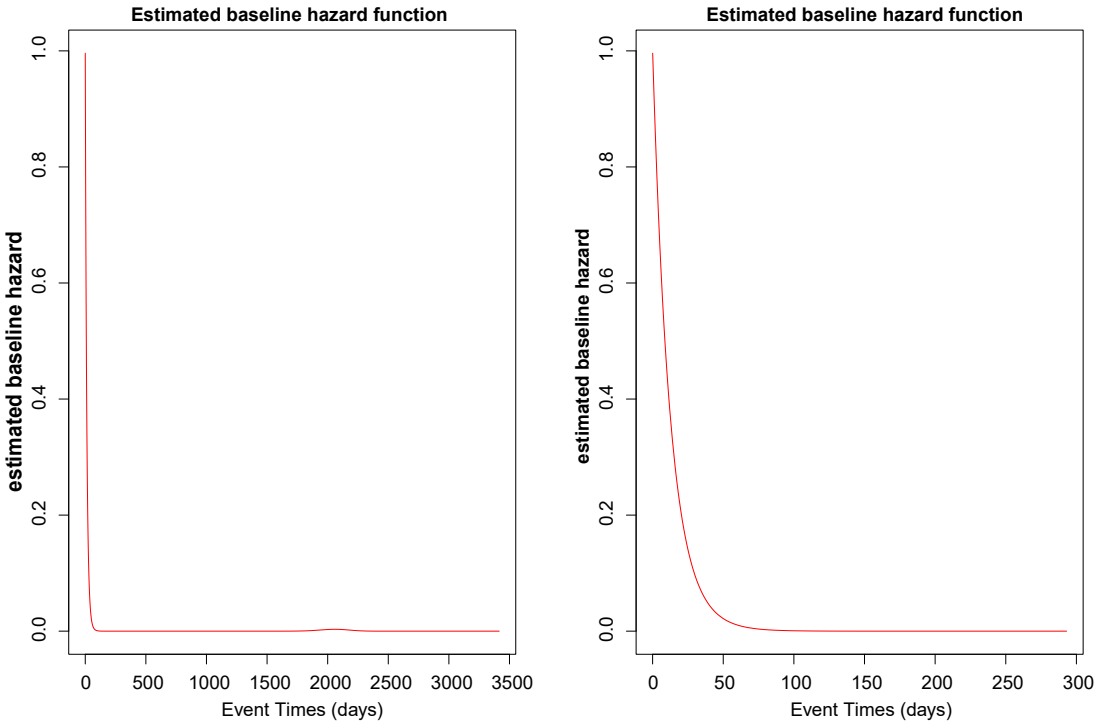


Figure 10: Estimated baseline hazard function. Note that since the estimated baseline hazard functions obtained from `B-Splines`, `B-Splines_no` are identical, we only present one plot for these methods. Graft loss mainly occurs in the first year after kidney transplantation

Regarding `PartialLikelihood`, its non-informative prior counterpart `PartialLikelihood_no` also detects interaction effects with noise variables (see Figure 11 and 12). This is the case when applying `PartialLikelihood_no` to both `complexData` and `simpleData`. In contrast, `PartialLikelihood` does not detect any effects related to the noise variables. More concretely, when applying `PartialLikelihood` to `complexData`, the main effects `dialysesvintage` and `Donorage`, as well as the interaction effects `dialysesvintage:Donorsex1` and `Donorsex1:TypecadavericDBDDCD1`, are detected. For `simpleData`, `PartialLikelihood` identifies the main effects `Donorage` and `HLA_mismatch`. Note that the interaction effect `Donorsex:TypecadavericDBDDCD` has been previously reported by Steenvoorden, Vogt, et al. (2024).

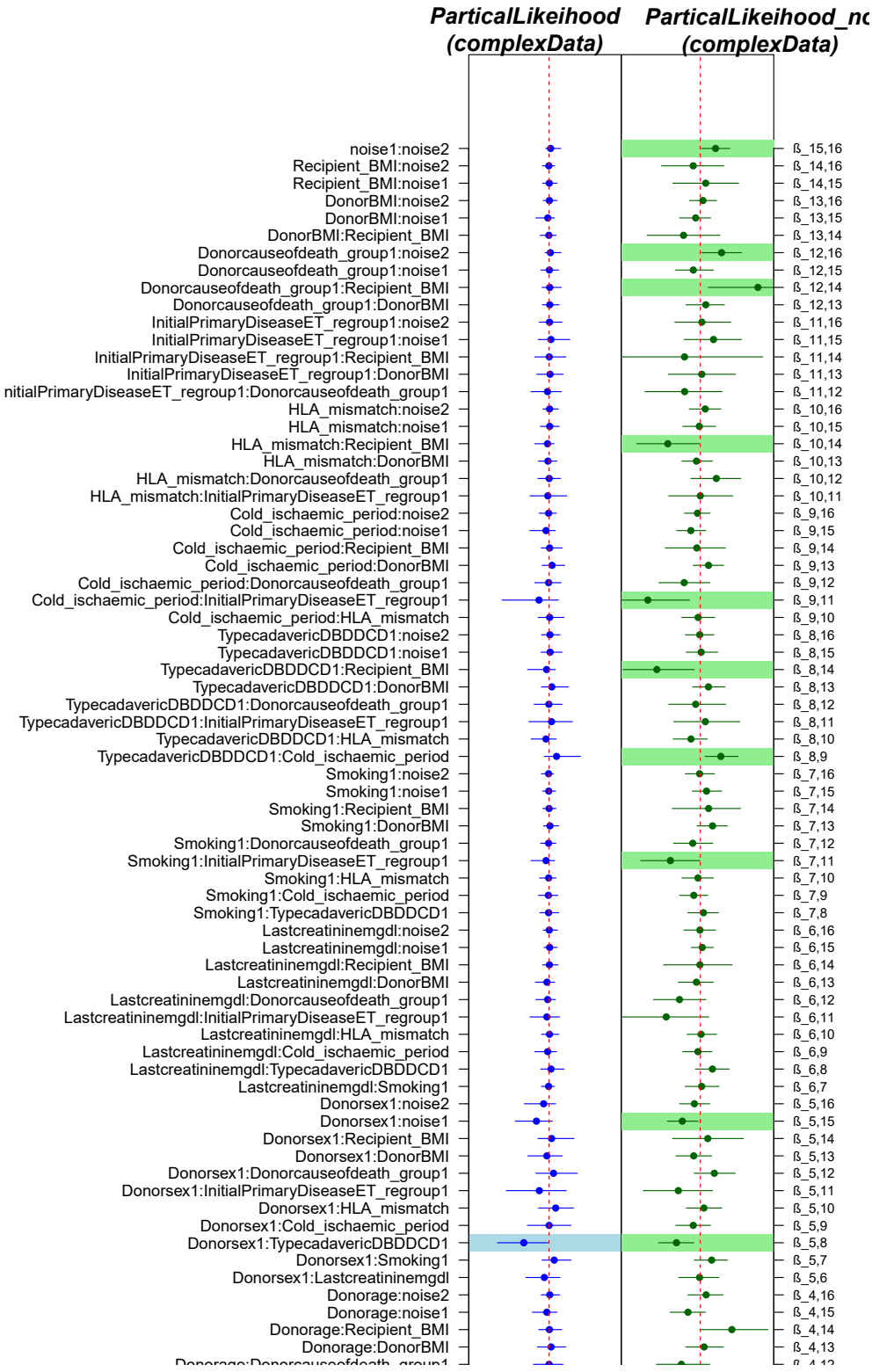

ParticalLikeihood (complexData)
ParticalLikeihood_nc (complexData)
noise1:noise2
Recipient_BMI:noise2
Recipient_BMI:noise1
DonorBMI:noise2
DonorBMI:noise1
DonorBMI:Recipient_BMI
Donorcauseofdeath_group1:noise2
Donorcauseofdeath_group1:noise1
Donorcauseofdeath_group1:Recipient_BMI
Donorcauseofdeath_group1:DonorBMI
InitialPrimaryDiseaseET_regroup1:noise2
InitialPrimaryDiseaseET_regroup1:noise1
InitialPrimaryDiseaseET_regroup1:Recipient_BMI
InitialPrimaryDiseaseET_regroup1:DonorBMI
nitialPrimaryDiseaseET_regroup1:Donorcauseofdeath_group1
HLA_mismatch:noise2
HLA_mismatch:noise1
HLA_mismatch:Recipient_BMI
HLA_mismatch:DonorBMI
HLA_mismatch:Donorcauseofdeath_group1
HLA_mismatch:InitialPrimaryDiseaseET_regroup1
Cold_ischaemic_period:noise2
Cold_ischaemic_period:noise1
Cold_ischaemic_period:Recipient_BMI
Cold_ischaemic_period:DonorBMI
Cold_ischaemic_period:Donorcauseofdeath_group1
Cold_ischaemic_period:InitialPrimaryDiseaseET_regroup1
Cold_ischaemic_period:HLA_mismatch
TypecadavericDBDDCD1:noise2
TypecadavericDBDDCD1:noise1
TypecadavericDBDDCD1:Recipient_BMI
TypecadavericDBDDCD1:DonorBMI
TypecadavericDBDDCD1:Donorcauseofdeath_group1
TypecadavericDBDDCD1:InitialPrimaryDiseaseET_regroup1
TypecadavericDBDDCD1:HLA_mismatch
TypecadavericDBDDCD1:Cold_ischaemic_period
Smoking1:noise2
Smoking1:noise1
Smoking1:Recipient_BMI
Smoking1:DonorBMI
Smoking1:Donorcauseofdeath_group1
Smoking1:InitialPrimaryDiseaseET_regroup1
Smoking1:HLA_mismatch
Smoking1:Cold_ischaemic_period
Smoking1:TypecadavericDBDDCD1
Lastcreatininemgdl:noise2
Lastcreatininemgdl:noise1
Lastcreatininemgdl:Recipient_BMI
Lastcreatininemgdl:DonorBMI
Lastcreatininemgdl:Donorcauseofdeath_group1
Lastcreatininemgdl:InitialPrimaryDiseaseET_regroup1
Lastcreatininemgdl:HLA_mismatch
Lastcreatininemgdl:Cold_ischaemic_period
Lastcreatininemgdl:TypecadavericDBDDCD1
Lastcreatininemgdl:Smoking1
Donorsex1:noise2
Donorsex1:noise1
Donorsex1:Recipient_BMI
Donorsex1:DonorBMI
Donorsex1:Donorcauseofdeath_group1
Donorsex1:InitialPrimaryDiseaseET_regroup1
Donorsex1:HLA_mismatch
Donorsex1:Cold_ischaemic_period
Donorsex1:TypecadavericDBDDCD1
Donorsex1:Smoking1
Donorsex1:Lastcreatininemgdl
Donorage:noise2
Donorage:noise1
Donorage:Recipient_BMI
Donorage:DonorBMI
ß_15,16
ß_14,16
ß_14,15
ß_13,16
ß_13,15
ß_13,14
ß_12,16
ß_12,15
ß_12,14
ß_12,13
ß_11,16
ß_11,15
ß_11,14
ß_11,13
ß_11,12
ß_10,16
ß_10,15
ß_10,14
ß_10,13
ß_10,12
ß_10,11
ß_9,16
ß_9,15
ß_9,14
ß_9,13
ß_9,12
ß_9,11
ß_9,10
ß_8,16
ß_8,15
ß_8,14
ß_8,13
ß_8,12
ß_8,11
ß_8,10
ß_8,9
ß_7,16
ß_7,15
ß_7,14
ß_7,13
ß_7,12
ß_7,11
ß_7,10
ß_7,9
ß_7,8
ß_6,16
ß_6,15
ß_6,14
ß_6,13
ß_6,12
ß_6,11
ß_6,10
ß_6,9
ß_6,8
ß_6,7
ß_5,16
ß_5,15
ß_5,14
ß_5,13
ß_5,12
ß_5,11
ß_5,10
ß_5,9
ß_5,8
ß_5,7
ß_5,6
ß_4,16
ß_4,15
ß_4,14
ß_4,13

Figure 11: Results of `complexData` from and `PartialLikelihood` and `PartialLikelihood_no`. Shadowed terms are the selected effects by the corresponding methods. `PartialLikelihood_no` detects a lot of covariates, including their interactions with noise variables. In contrast, `PartialLikelihood` only detects `dialysesvintage`, `Donorage`, `dialysesvintage:Donorsex`, and `Donorsex:TypecadavericDBDDCD`. Together, when analyzing `complexData`, `PartialLikelihood_no` is not trustworthy, and we should stick to the `PartialLikelihood`.

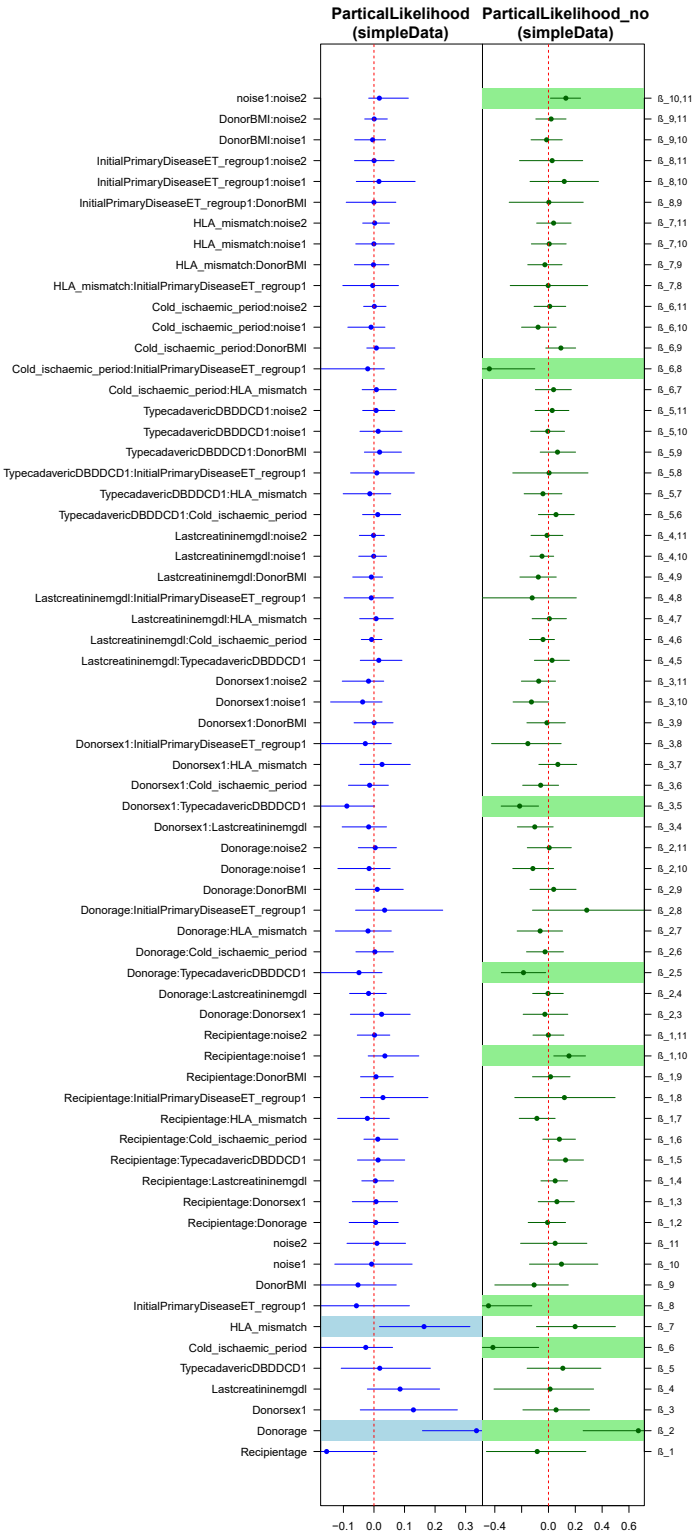


Figure 12: Results of `simpleData` from `PartialLikelihood` and `PartialLikelihood_no`. Shadowed terms are the selected effects by the corresponding methods. `PartialLikelihood_no` detects a lot of covariates, including the interactions with noise variable. In contrast, `PartialLikelihood` only detects the following covariates: `Donorage` and `HLA_mismatch`. Together, when analyzing `simpleData`, `PartialLikelihood_no` is not trustworthy. We should stick to the `PartialLikelihood`.

We also incorporate another three key aspects into our evaluation of `PartialLikelihood`: (1) model predictions to assess performance in the context of kidney transplantation data; (2) the correlation of the median coefficient values for the posterior distributions obtained from `PartialLikelihood` and `PartialLikelihood_no`; and (3) the correlation of the standard deviations for the posterior distributions obtained from `PartialLikelihood` and `PartialLikelihood_no`. For prediction performance, Figure 13 shows that `PartialLikelihood` achieves better predictive accuracy compared to `PartialLikelihood_no`. Moreover, predictions based on `complexData` outperform those derived from `simpleData`. Regarding the correlation (see Figure 14), `PartialLikelihood` does not yield systematically different median coefficient values but produces larger standard deviations for $\boldsymbol{\beta}$ than `PartialLikelihood_no`. This similarity in median coefficient values and discrepancy in standard deviations for the posterior distributions are less pronounced when applying the methods to `simpleData` than to `complexData`.

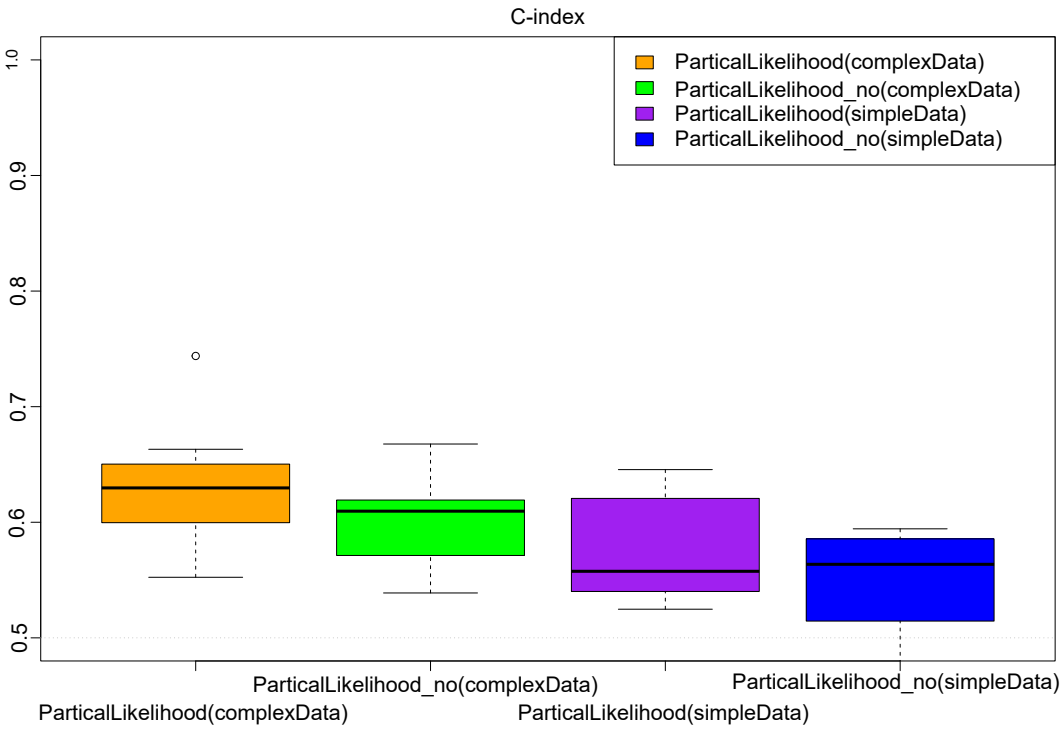


Figure 13: Prediction performance of `PartialLikelihood` and `PartialLikelihood_no` on `complexData` and `simpleData`. `PartialLikelihood` demonstrates overall superior predictive accuracy compared to `PartialLikelihood_no`. Predictions derived from `complexData` outperform those based on `simpleData`

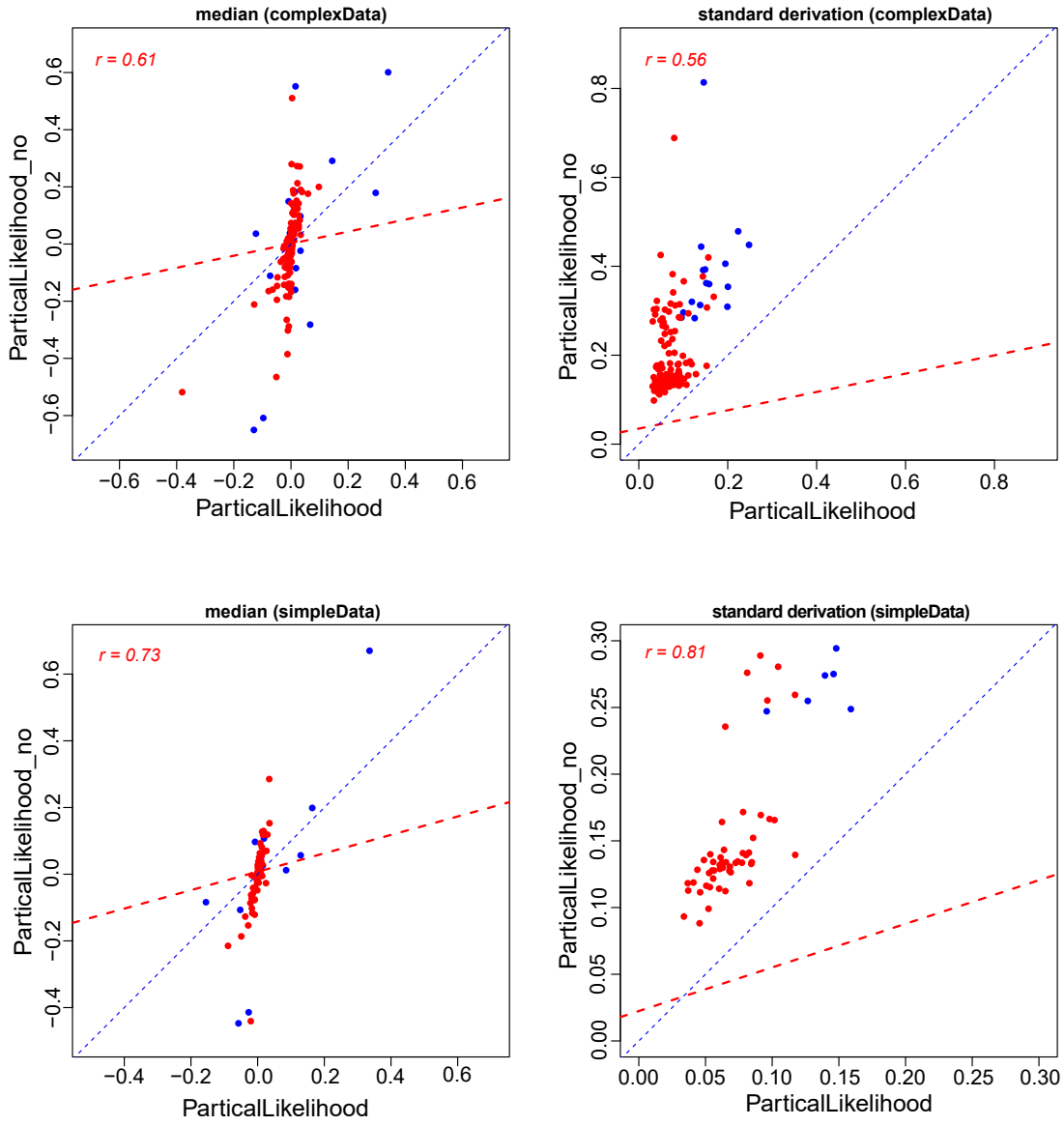


Figure 14: The correlation between the median coefficient values and their standard deviations for the posterior distributions obtained from `PartialLikelihood` and `PartialLikelihood_no`. The red dashed line represents the regression fit, while the blue dashed line is the diagonal line, which indicates the reference line where the two values are equal. The blue dots are the main effects and the red dots are the interaction ones. `PartialLikelihood` does not yield systematically different median coefficient values but produces larger standard deviations for $\boldsymbol{\beta}$ than `PartialLikelihood_no`. This similarity in median coefficient values and discrepancy in standard deviations for the posterior distributions are less pronounced when applying the methods to `simpleData` than to `complexData`.

Relationship between the $\hat{\beta}$ obtained from `complexData` and `simpleData`

As mentioned earlier, certain effects are detected by `PartialLikelihood` when applied to `complexData`, but remain undetected in the analysis of `simpleData` (see Table 5 for the summarized results). These effects include `dialysisvintage`, `HLA_mismatch`, `dialysisvintage:Donorsex`, and `Donorsex:DonorType`. Note that the variables `dialysisvintage` and `dialysisvintage:Donorsex` are present only in `complexData` and do not exist in `simpleData`. Therefore, we focus on `HLA_mismatch` and `Donorsex:DonorType`.

As shown in Table 5, the credible intervals for `HLA_mismatch` and `Donorsex:DonorType` obtained from both `complexData` and `simpleData` are very close to zero, suggesting that the corresponding posterior distributions are quite similar. To further verify this observation, we examine the correlation of the median and standard deviations of the posterior distributions for overlapping $\boldsymbol{\beta}$ obtained from `complexData` and `simpleData` (see Figure 15). Accordingly, the median and standard deviations of the posterior distributions for overlapping $\boldsymbol{\beta}$ obtained from `complexData` and `simpleData` are also largely similar.

Table 5: Detected Effects by `PartialLikelihood` on `complexData` vs. `simpleData`

	PartialLikelihood(complexData)			PartialLikelihood(simpleData)		
	median	credible interval	Selected	median	credible interval	Selected
dialysisvintage	0.293	[0.071, 0.472]	Yes	-	-	-
Donorage	0.340	[0.153, 0.598]	Yes	0.336	[0.159, 0.558]	Yes
HLA_mismatch	0.144	[-0.002, 0.314]	No	0.164	[0.018, 0.314]	Yes
dialysisvintage:Donorsex	-0.129	[-0.262, -0.005]	Yes	-	-	-
Donorsex:DonorType	-0.129	[-0.262, -0.005]	Yes	-0.089	[-0.232, 0.002]	No

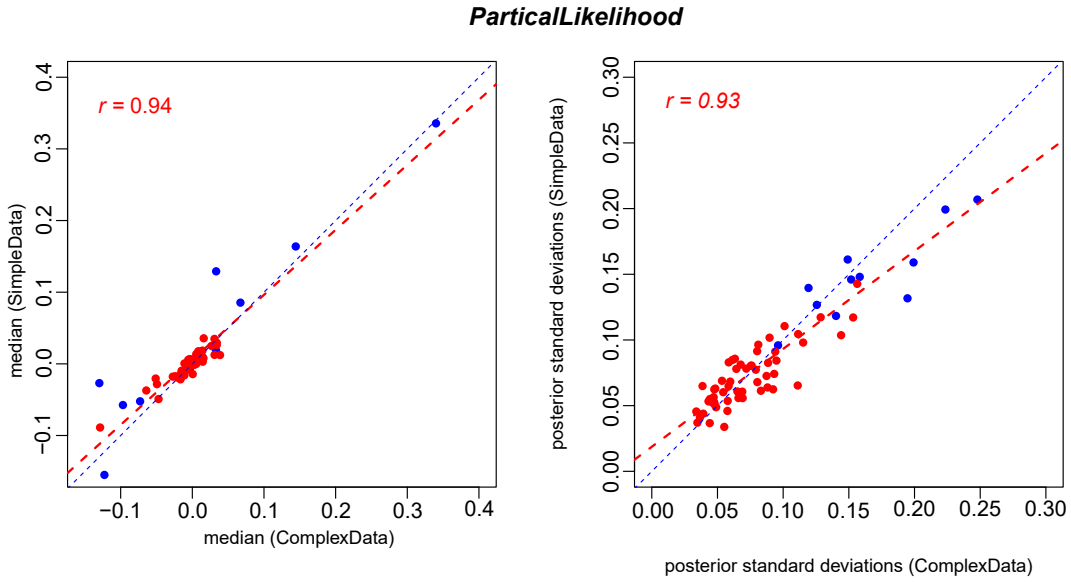


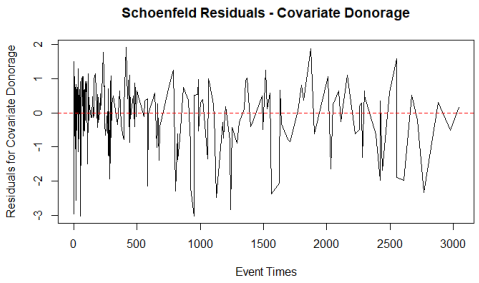
Figure 15: Correlations of the overlapping $\hat{\boldsymbol{\beta}}$ estimates from `complexData` and `simpleData` using `PartialLikelihood`. a. Correlation of the posterior distribution medians. b. Correlation of the posterior distribution standard deviations. The red dashed line represents the regression fit, while the blue dashed line is the diagonal line, which indicates the reference line where the two values are equal. The blue dots are the main effects and the red dots are the interaction ones. Overall, the two sets of estimates exhibit a strong correlation. That is, posterior distributions for overlapping $\boldsymbol{\beta}$ obtained from `complexData` and `simpleData` are similar.

Assumption checking

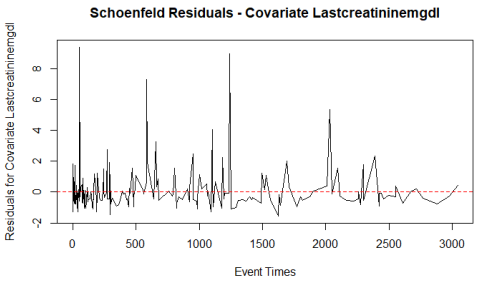
As mentioned before, two key assumptions of PH models is non-informative censoring and proportional hazards. With respect to the non-informative censoring, the censored time C_i has been recorded conditionally independently of the observed time T_i given the covariates $\mathbf{x}_i$. With respect to the proportional hazards, we employ the Schoenfeld residual to check the proportional hazard assumption. As a result, we found that some variable might violate PH assumptions, i.e., `Donorage`, `Lastcreatininemgdl`, `Cold_ischaemic_period`, and `DonorBMI` (see Figure 16). This violation of the PH assumptions can also be reflected on its corresponding interactions.

Discussion

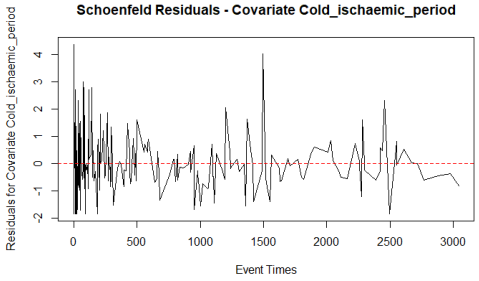
In summary, to have an accurate estimation and selection of interaction effects to better explain graft loss in kidney transplantation, we apply the previously proposed linked shrinkage prior distribution to Bayesian proportional hazards (PH) models by integrating both full and partial



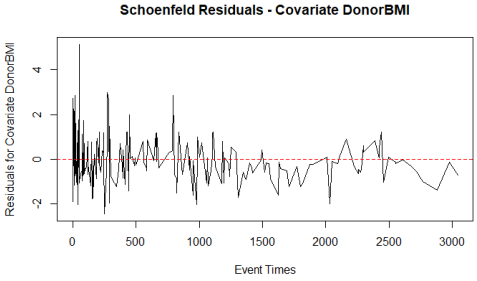
(a) `Donorage`



(b) `Lastcreatininemgdl`



(c) `Cold_ischaemic_period`



(d) `DonorBMI`

Figure 16: Results of PH assumption checking. The Schoenfeld residuals for `Donorage`, `Lastcreatininemgdl`, `Cold_ischaemic_period`, and `DonorBMI` exhibit time-dependent patterns, indicating a violation of the PH assumption for these variables.

likelihood estimators. For the full likelihood estimator, we consider three different situations for the baseline hazard function: (1) assuming that the true event times T^* follows Exp(λ); (2) assuming that the true event times T^* follows Weibull(λ, v); and (3) approximating $\log h_0(t)$ using B-Splines methods. In contrast, the partial likelihood estimator circumvents the need to specify the baseline hazard. These methods and the corresponding characteristics have been summarized in Table 6.

Table 6: Overview of model specifications and associated characteristics

Model	**Likelihood** (Full vs. Partial)	**Baseline Hazard** $h_0(t)$	**Prior for** $\boldsymbol{\beta}$	**Ancillary Parameters and Priors**
`Exponential`	Full	λ	Linked shrinkage	$\lambda \sim \text{Lognormal}(0, 100)$
`Weibull`	Full	$\lambda v t^{v-1}$	Linked shrinkage	$\lambda \sim \text{Lognormal}(0, 100)$ $v \sim \text{Lognormal}(0, 100)$
`B-Spline`	Full	$\exp\left(\sum_{l=1}^{L} \gamma_l B_l(t; k, \zeta)\right)$	Linked shrinkage	$\gamma_l \sim \mathcal{N}(0, 100)$
`PartialLikelihood`	Partial	—	Linked shrinkage	—
`B-Spline_no`	Full	$\exp\left(\sum_{l=1}^{L} \gamma_l B_l(t; k, \zeta)\right)$	$\mathcal{N}(0, 100)$	$\gamma_l \sim \mathcal{N}(0, 100)$
`PartialLikelihood_no`	Partial	—	$\mathcal{N}(0, 100)$	—

In the simulation setting, we demonstrate that when the distribution of event times t is unknown, both `B-Splines` and `PartialLikelihood` outperform models that include all two-way interactions and incorporate the non-informative prior in terms of parameter estimation, variable selection, hypothesis testing power, and prediction accuracy. However, no consistent performance difference is observed between `B-Splines` and `PartialLikelihood`. In the analysis of real-world kidney transplantation data, we find that the advantage of the linked shrinkage prior exists in `PartialLikelihood` but not in `B-Splines`. Moreover, `PartialLikelihood` yields similar posterior distributions of overlapping $\boldsymbol{\beta}$ before and after reducing the number of covariates in the kidney transplantation data.

The outperformance of `B-Splines` and `PartialLikelihood` in the simulation study underscores the effectiveness of the linked shrinkage prior in enhancing parameter estimation, particularly for models incorporating all two-way interactions. This observation is consistent with findings reported by van de Wiel et al. (2024) in regression contexts. For the real-world kidney transplantation analysis, we discuss their results in the remainder of this section from two perspectives: (1) the outperformance of `PartialLikelihood` but the underperformance of `B-Splines`; and (2) posterior distributions of overlapping $\boldsymbol{\beta}$ produced by `PartialLikelihood` in the kidney transplantation data analysis.

The outperformance of `PartialLikelihood` but the underperformance of `B-Splines` in kidney transplantation data

The outperformance of `PartialLikelihood` and the underperformance of `B-Splines` highlight that the linked shrinkage prior offers clear advantages in improving parameter estimation for models incorporating all two-way interactions. However, this advantage becomes limited when

using B-Splines to approximate $\log h_0(t)$ for the kidney transplantation dataset analyzed in this study. To understand these limitations of `B-Splines`, we closely examine the theoretical differences between `B-Splines` and `PartialLikelihood`. We identify three key differences in their approaches to modeling $h_0(t)$ that might explain the limitation of `B-Splines`.

First, `B-Splines` employs a number of knots placed in suboptimal locations when modeling $h_0(t)$, which is not the case for `PartialLikelihood`. In this study, the number of knots is three: two evenly spaced internal knots based on quantiles and two boundary knots, resulting in a B-spline basis with three columns. This three-column basis is unlikely to cause overfitting of the baseline hazard function $h_0(t)$. Therefore, we argue that the placement and number of knots are not responsible for the disadvantaged performance of `B-Splines` compared to `PartialLikelihood` in kidney transplantation data analysis. Thus, we discard this possible explanation.

Second, `B-Splines` estimates $h_0(t)$ directly, whereas `PartialLikelihood` circumvents modeling the baseline hazard function. The modeling of $h_0(t)$ in `B-Splines` may implicitly absorb unaccounted time-varying effects (Sleeper & Harrington, 1990), such as `Cold_ischemic_period` in the kidney transplantation data. In contrast, the coefficients estimated by `PartialLikelihood` represent the averaged effect for time-varying covariates and the actual effect for time-constant covariates. Taken together, the lack of adjustment for time-varying effects may explain why `B-Splines` underperforms in kidney transplantation data analysis.

Last, `B-Splines` assumes that the true baseline hazard function $h_0(t)$ is smooth, which is not required by `PartialLikelihood`. Given that both the Kaplan–Meier curve and the estimated $\hat{h}_0(t)$ from `B-Splines` exhibit a sharp decline at the beginning of the follow-up period, we argue that the true $h_0(t)$ may not be smooth and is therefore not fully captured by `B-Splines`. Thus, it is plausible that the lack of smoothness in the true baseline hazard contributes to the poor performance of `B-Splines`. However, at this stage, we cannot draw a definitive conclusion. The underlying mechanism for the failure of `B-Splines` requires further investigation in future work.

The posterior distributions of overlapping $\boldsymbol{\beta}$ produced by `PartialLikelihood`

Regarding the posterior distributions of overlapping $\boldsymbol{\beta}$ produced by `PartialLikelihood`, we discuss them from two perspectives. First, the similarity in the posterior distributions of overlapping $\boldsymbol{\beta}$ before and after reducing the number of covariates in the kidney transplantation data suggests that reducing covariates does not improve estimation stability for `PartialLikelihood`, even though it might seem intuitive that model simplification could enhance stability. Instead, simplification in the context of `PartialLikelihood` introduces the risk of omitting important variables and compromising the interpretability and completeness of the model. These omitted variables and their interactions may lead to biased conclusions about the relationships between predictors and outcomes.

Second, the comparison between the posterior distributions of overlapping $\boldsymbol{\beta}$ produced by

`PartialLikelihood` and `PartialLikelihood_no` indicates that the linked shrinkage prior primarily affects the dispersion rather than the location of the posterior. This observation is consistent with the theoretical aim of the linked shrinkage prior, which is to reduce variances by borrowing information across parameters without imposing a strong directional shift on the mean or median.

Conclusion

Linked shrinkage priors are useful for improving the detection of interaction terms in time-to-event data analysis, but they require appropriate methods to model the baseline hazard function $h_0(t)$. Using the partial likelihood estimator may be a safer choice compared to using full likelihood estimator with the B-Splines method for approximating the baseline hazard function. However, the underlying reason for this advantage of the partial likelihood estimator needs to be investigated in future studies. Furthermore, the linked shrinkage prior that incorporates the partial likelihood estimator has demonstrated reasonable robustness before and after reducing the number of covariates in the kidney transplantation data.

Limitations

As shown earlier, the proportional hazards (PH) assumption does not hold for every covariate in the kidney transplantation dataset. This implies, on one hand, that the current results are limited and cannot be directly interpreted to explain graft loss. On the other hand, it highlights the need to extend the Bayesian survival model used in this study to account for potential time-varying effects in future research. One way to address such effects is by introducing $\boldsymbol{\beta}(t)$ for covariates that violate the PH assumption. Alternatively, focusing only on graft loss occurring in the early stage—such as within one year after transplantation—could also be a reasonable approach.

References

Beca, J. M., Chan, K. K., Naimark, D. M., & Pechlivanoglou, P. (2021). Impact of limited sample size and follow-up on single event survival extrapolation for health technology assessment: A simulation study. *BMC Medical Research Methodology*, *21*(1), 282.

Brilleman, S. L., Wolfe, R., Moreno-Betancur, M., & Crowther, M. J. (2021). Simulating survival data using the simsurv r package. *Journal of Statistical Software*, *97*, 1–27.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, *62*(2), 269–276.

Efron, B. (1977). The efficiency of cox's likelihood function for censored data. *Journal of the American statistical Association*, *72*(359), 557–565.

Hao, N., & Zhang, H. H. (2017). A note on high-dimensional linear regression with interactions. *The American Statistician*, *71*(4), 291–297.

Johnson, M. E., Tolley, H. D., Bryson, M. C., & Goldman, A. S. (1982). Covariate analysis of survival data: A small-sample study of cox's model. *Biometrics*, 685–698.

Kalbfleisch, J. D. (1978). Non-parametric bayesian analysis of survival time data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *40*(2), 214–221.

Kim, Y., & Kim, D. (2009). Bayesian partial likelihood approach for tied observations. *Journal of statistical planning and inference*, *139*(2), 469–477.

Lázaro, E., Armero, C., & Alvares, D. (2021). Bayesian regularization for flexible baseline hazard functions in cox survival models. *Biometrical Journal*, *63*(1), 7–26.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69*(1), 239–241.

Sleeper, L. A., & Harrington, D. P. (1990). Regression splines in the cox model with application to covariate effects in liver disease. *Journal of the American Statistical Association*, *85*(412), 941–949.

Steenvoorden, T. S., Evers, L., Vogt, L., Rood, J. A., Kers, J., Baas, M. C., Christiaans, M. H., Lindeman, J. H., Sanders, J.-S. F., de Vries, A. P., et al. (2024). The differential impact of early graft dysfunction in kidney donation after brain death and after circulatory death: Insights from the dutch national transplant registry. *American Journal of Transplantation*.

Steenvoorden, T. S., Vogt, L., Hilhorst, M., Bemelman, F. J., Kers, J., & Peters-Sengers, H. (2024). Impact of donor sex on graft outcome in deceased donor kidney transplantation: Th-po777. *Journal of the American Society of Nephrology*, *35*(10S), 10–1681.

van de Wiel, M. A., Amestoy, M., & Hoogland, J. (2024). Linked shrinkage to improve estimation of interaction effects in regression models. *Epidemiologic Methods*, *13*(1), 20230039.

Wang, W., & Yan, J. (2021). Shape-restricted regression splines with r package splines2. *Journal of Data Science*, *19*(3).