Personalized Schedules for Kidney Transplant Patients

Firstname Lastname[1,*]

[1]Department of Biostatistics, Erasmus University Medical Center, the Netherlands
*email: a.tomer@erasmusmc.nl

1 Introduction

According to the inclusion criteria of the study, a total of 239 kidney transplant patients were included in the data set. The transplantation characteristics of these patients is presented in Table 1. The data set also includes periodical measurements of serum creatinine (SCr) and protein creatinine ratio (PCR), which are biomarkers used to check the state of the transplant. The median number of repeated SCr and PCR measurements per patient are 45 and 37, respectively. For SCr 95% of the observations are taken before 6 years, while for PCR they are taken before 5.4 years. The median time between two SCr measurements is 10 days, while the same for PCR is 14 days.

Table 1: Observed transplantation characteristics of the studied population (n = 239).

Quantitative characteristics		
Name	Abbreviation	Mean (SD)
Receiver age (at baseline)	rec_age	50.70 (13.09)
Donor age	d_age	49.73 (12.66)
Donor BMI	d_bmi	25.10 (4.43)
Receiver BMI	rec_bmi	25.43 (4.31)
Cold ischemia time (minutes)	cit	887.25 (522.95)
#HLA A, B and DR mismatches	hla	2.81 (1.57)
Panel reactive antibody (%)	pra	4.81 (14.20)
#Days on dialysis before transplant	dial_days	1334.91 (1283.93)
#Anti-hypertensive medicaments	ah_nr	1.58 (0.96)

Categorical characteristics		
Name	Abbreviation	Category (%)
Receiver gender	rec_gender	Female (42.68 %)
Donor gender	d_gender	Female (56.49 %)
Delayed graft function after transplant	dgf	No (67.78 %)
Previous transplantation	prev_tp	No (84.45 %)
Diabetes mellitus	dm	No (84.52 %)
Known cardiovascular events before transplant	hvdis	No (61.92 %)
Deceased donor	d_cadaveric	No (25.94 %)

In the cohort, 44 out of 239 patients observed graft failure (death-censored). The survival probability one year post transplantation is 97.87%. The graft survival probabilities and 95% CI estimated using Kaplan-Meier estimator are presented in Figure 1.

2 Joint Model

Our goal is to check if SCr and PCR both, are useful to predict graft failure. To this end, we model the two longitudinal outcomes and graft failure together using joint models (JMs) for time

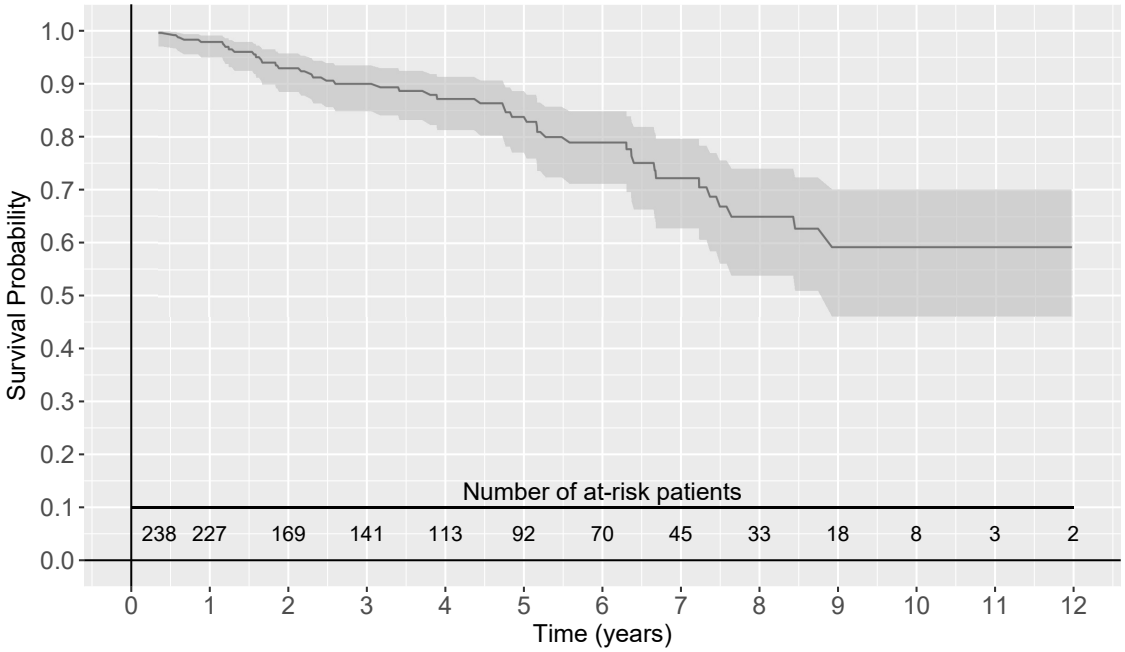

Figure 1: Graft survival probabilities and 95% CI estimated using Kaplan-Meier estimator.

to event and longitudinal data (Rizopoulos, 2012; Tsiatis and Davidian, 2004). In this model we use log transformed values of both SCr and PCR (Fournier et al., 2016). More specifically, we model the impact of log(SCr) value and log(SCr) velocity, log(PCR) value and log(PCR) velocity, and transplantation characteristics on the hazard of graft failure. In this regard, the JM consists of a multivariate longitudinal sub-model to model the evolution of SCr and PCR and a relative risk sub-model to model the impact of transplantation characteristics and biomarkers on the hazard of graft failure. The longitudinal evolution of the two outcomes over time is modeled flexibly using B-splines. The model formulation for PCR outcome is as follows (for SCr outcome it is same):

$$\begin{aligned}\log \mathrm{PCR}(t) = {} & \beta_0 + \beta_1 \mathrm{rec\_age} + \beta_2 \mathrm{d\_age} + \beta_3 \mathrm{d\_bmi} + \beta_4 \mathrm{rec\_bmi} + \beta_5 \mathrm{cit} \\ & + \beta_6 \mathrm{hla} + \beta_7 \mathrm{pra} + \beta_8 \mathrm{dial\_days} + \beta_9 \mathrm{ah\_nr} + \beta_{10} \mathrm{rec\_gender} + \beta_{11} \mathrm{d\_gender} \\ & + \beta_{12} \mathrm{dgf} + \beta_{13} \mathrm{prev\_tp} + \beta_{14} \mathrm{dm} + \beta_{15} \mathrm{hvdis} + \beta_{16} \mathrm{d\_cadaveric} \\ & + \sum_{k=1}^{4} \beta_{k+16} B_k(t, \mathcal{K}) + b_{i0} + \sum_{k=1}^{4} b_{ik} B_k(t, \mathcal{K}) + \varepsilon_i(t),\end{aligned} \tag{1}$$

where $B_k(t, \mathcal{K})$ denotes the k-th basis function of a B-spline with three internal knots at $\mathcal{K} = \{0.082, 0.219, 1\}$ (30 and 80 days recommended by the clinicians) years, and boundary knots at 0.039 and 6 years (minimum and 0.95 quantile of the time of measurements two outcomes). The quantitative transplantation characteristics are standardized to avoid convergence issues in parameter estimation. For the relative risk sub-model the hazard function we fit is given by:

$$h_i(t) = h_0(t) \exp\left\{\gamma_1 \mathrm{prev\_tp} + \gamma_2 \mathrm{hla} + \gamma_3 \mathrm{cit} + \gamma_4 \mathrm{dial\_days} + \alpha_1 m_{i1}(t) + \alpha_2 m'_{i1}(t)\right\}. \tag{2}$$

where α_1 and α_2 are measures of strength of the association between hazard of graft failure and $\log \mathrm{PCR}$ value $m_{i1}(t)$ and $\log \mathrm{PCR}$ velocity $m'_{i1}(t)$, respectively.

The parameters of the JM are estimated using the R package **JMbayes** (Rizopoulos, 2016), which uses the Bayesian methodology to estimate the model parameters (section B, supplementary material). Out of 239 patients, we use the data of only those 238 patients for whom both PCR and SCr data is available. The parameter estimates for the longitudinal sub-model for SCr and PCR are provided in Table 2 and Table 3, respectively. The effect of transplantation characteristics on both outcomes is small and ignorable, and hence not discussed in detail. Since the quantitative variables are standardized, the effect sizes correspond to one standard deviation increase in the corresponding variable. To avoid the tricky interpretation of variables corresponding to evolution

over time, instead the evolution of SCr and PCR over time is depicted in Figure 2, and Figure 3, respectively. @Hessel: We have to explain why the creatinine levels dip.

Table 2: Parameter estimates for the longitudinal model for SCr.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	5.226	0.080	5.064	5.378	<0.000
Receiver age	-0.063	0.022	-0.107	-0.019	0.010
Donor age	0.083	0.020	0.045	0.119	<0.000
Donor BMI	-0.011	0.021	-0.054	0.028	0.612
Receiver BMI	0.018	0.023	-0.025	0.060	0.420
#HLA mismatches between donor and recipient	0.020	0.022	-0.022	0.065	0.342
Panel reactive antibody percentage	0.048	0.027	-0.008	0.100	0.082
#Anti-hypertensive medicaments	0.040	0.020	0.001	0.080	0.048
Cold ischemia time	0.029	0.035	-0.039	0.102	0.390
#Days on dialysis before transplant	0.015	0.029	-0.042	0.071	0.580
Receiver gender: Male	0.197	0.042	0.111	0.276	<0.000
Previous transplant: Yes	0.016	0.064	-0.115	0.141	0.786
Donor gender: Male	0.053	0.042	-0.027	0.136	0.198
Delayed graft function: Yes	0.118	0.049	0.025	0.216	0.006
Diabetes Mellitus: Yes	-0.103	0.059	-0.217	0.012	0.076
Cardiovascular events before transplantation: Yes	-0.047	0.043	-0.129	0.044	0.272
Deceased donor: Yes	0.163	0.082	0.004	0.313	0.044
Spline: visit time [0.039, 0.082] years	-0.440	0.041	-0.517	-0.358	<0.000
Spline: visit time [0.082, 0.219] years	-0.182	0.053	-0.284	-0.081	<0.000
Spline: visit time [0.219, 1] years	-0.545	0.081	-0.712	-0.395	<0.000
Spline: visit time [1, 6] years	0.007	0.083	-0.155	0.176	0.946
σ	0.190	0.001	0.187	0.192	

Table 3: Parameter estimates for the longitudinal model for PCR.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Intercept	3.731	0.179	3.398	4.083	<0.000
Receiver age	0.030	0.052	-0.066	0.138	0.604
Donor age	0.209	0.047	0.118	0.301	<0.000
Donor BMI	-0.019	0.051	-0.121	0.084	0.716
Receiver BMI	-0.116	0.050	-0.219	-0.021	0.014
#HLA mismatches between donor and recipient	-0.013	0.049	-0.112	0.086	0.776
Panel reactive antibody percentage	0.047	0.061	-0.066	0.166	0.446
#Anti-hypertensive medicaments	0.056	0.047	-0.03	0.147	0.208
Cold ischemia time	0.062	0.082	-0.097	0.211	0.468
#Days on dialysis before transplant	0.006	0.066	-0.120	0.134	0.952
Receiver gender: Male	-0.026	0.094	-0.207	0.166	0.798
Previous transplant: Yes	0.035	0.149	-0.241	0.332	0.816
Donor gender: Male	0.114	0.096	-0.079	0.303	0.228
Delayed graft function: Yes	0.043	0.118	-0.174	0.275	0.740
Diabetes Mellitus: Yes	0.153	0.135	-0.124	0.396	0.256
Cardiovascular events before transplantation: Yes	-0.016	0.106	-0.221	0.199	0.890
Deceased donor: Yes	0.144	0.193	-0.246	0.509	0.462
Spline: visit time [0.039, 0.082] years	-0.821	0.090	-0.989	-0.638	<0.000
Spline: visit time [0.082, 0.219] years	-0.578	0.131	-0.838	-0.304	<0.000
Spline: visit time [0.219, 1] years	-0.898	0.160	-1.218	-0.587	<0.000
Spline: visit time [1, 6] years	0.460	0.234	0.015	0.927	0.036
σ	0.479	0.004	0.472	0.486	

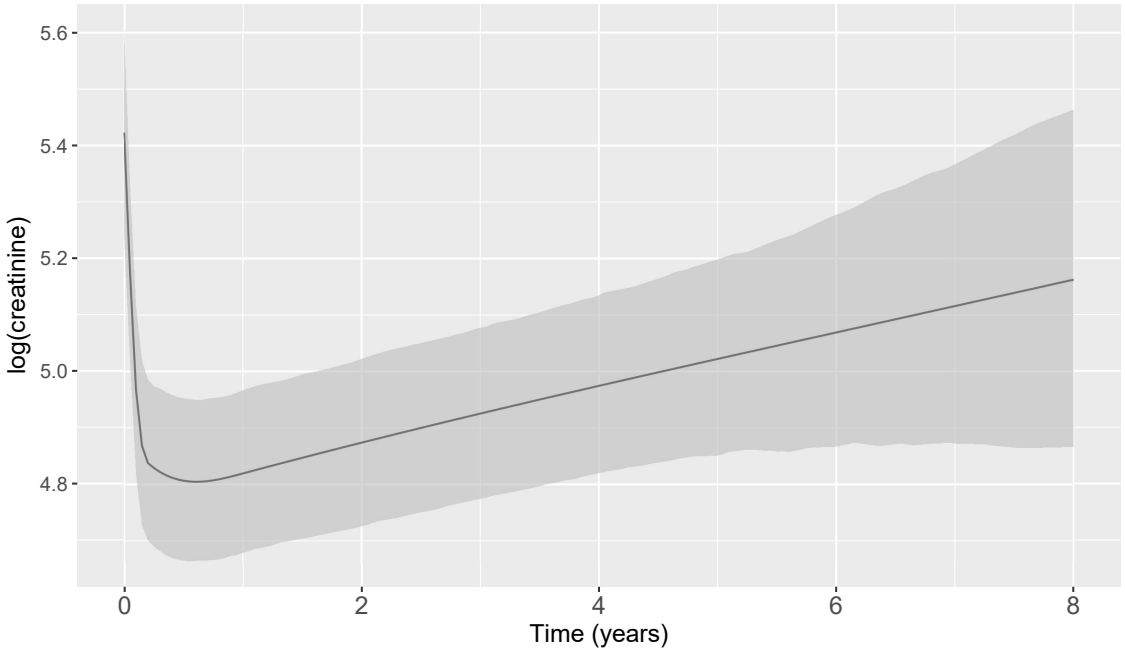

Figure 2: Fitted longitudinal evolution of SCr and 95% credible interval for a patient with the transplantation characteristics described in Table 1.

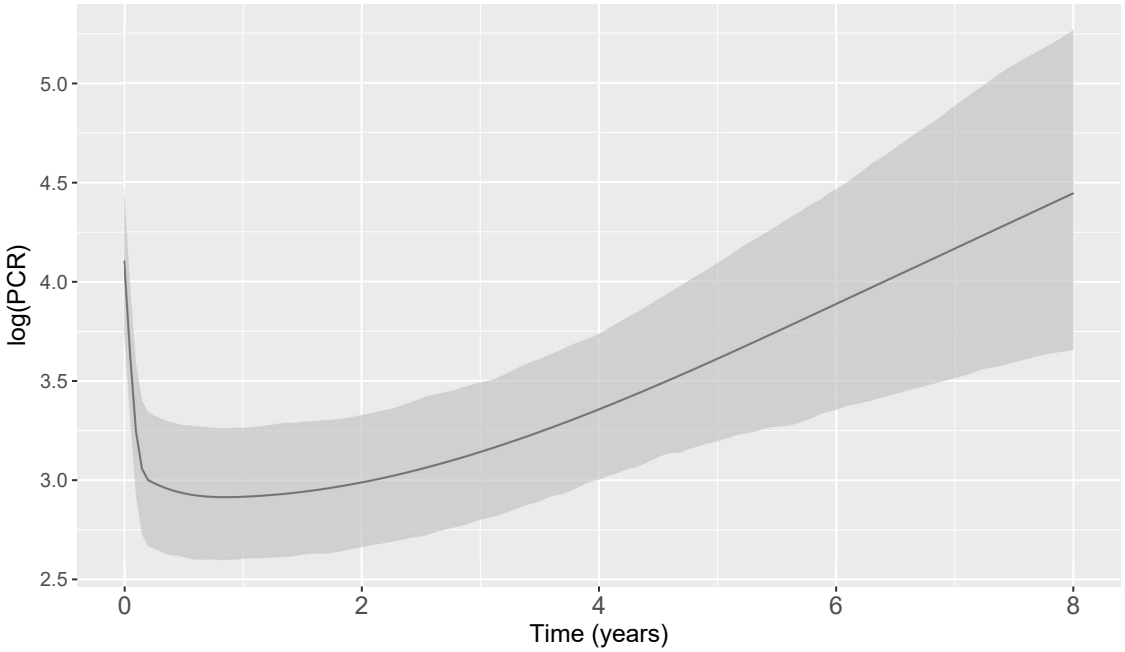

Figure 3: Fitted longitudinal evolution of PCR and 95% credible interval for a patient with the transplantation characteristics described in Table 1.

The parameter estimates for the relative risk sub-model are provided in Table 4. Since the quantitative variables are standardized, the effect sizes correspond to one standard deviation increase in the corresponding variable. We found that the log SCr levels are strongly associated with the hazard of GR. More specifically, for a given patient, if the SCr levels become twice and the remaining variables in the relative risk model remain the same, the hazard of graft failure increases three fold. log PCR levels and velocity are not strongly associated with hazard of GR. To further verify if they are required in the model in presence of both log SCr levels and velocity, we fitted two more JMs. In the first JM we modeled the association between log SCr levels and velocity

and hazard of graft failure. In the second JM we modeled the association between log PCR levels and velocity and hazard of graft failure. We then calculated time dependent AUC, that is, area under the curve (Rizopoulos, 2016; Rizopoulos, Molenberghs, and Lesaffre, 2017) values for all of the three JMs. The time dependent AUC were calculated periodically at an interval of 6 months. The first AUC was calculated at 6 months since transplantation and the last AUC was calculated at 3 years since transplantation. The resulting AUC values are plotted in Figure 4 and listed in Table 5. It can be seen that the model with both longitudinal outcomes performs the same as the model with only creatinine, to discriminate between patients who obtain graft failure versus others. Hence modeling PCR may not be necessary.

Table 4: Relative risk sub-model estimates for mean and 95% credible interval.

Variable	Mean	Std. Dev	2.5%	97.5%	P
Previous transplant: Yes	0.305	0.339	-0.099	0.986	0.352
#HLA mismatches between donor and recipient	0.048	0.093	-0.114	0.269	0.620
Cold ischemia time	-0.051	0.105	-0.277	0.133	0.644
#Days on dialysis before transplant	-0.013	0.102	-0.251	0.178	0.934
log PCR	0.145	0.125	-0.056	0.431	0.188
Slope(log PCR)	0.021	0.058	-0.076	0.145	0.828
log SCr	1.599	0.241	1.067	2.063	<0.000
Slope(log SCr)	0.203	0.123	-0.017	0.443	0.082

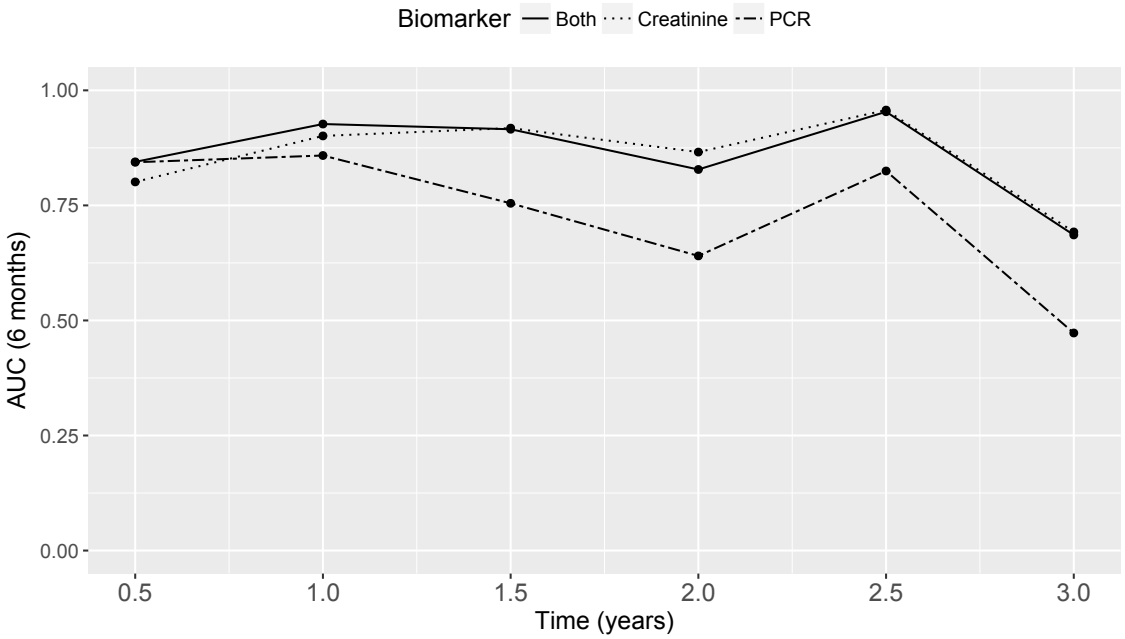

Figure 4: Area under curve characteristics for the JMs fitted to the kidney transplant data set.

Table 5: Area under curve characteristics for the JMs fitted to the kidney transplant data set.

Biomarkers	Year 0.5	Year 1	Year 1.5	Year 2	Year 2.5	Year 3
Both SCr and PCR	0.845	0.927	0.915	0.828	0.953	0.686
Only SCr	0.801	0.901	0.918	0.866	0.957	0.692
Only PCR	0.844	0.858	0.755	0.640	0.825	0.473

3 Personalized Schedules for Measurement of SCr

Currently, the schedule for measurement of SCr levels and fixed and common for all patients. SCr levels are measured 20 times in the first year after transplantation and every three months thereafter. Such fixed and frequent schedules are often burdensome for the patients. Patients who remain relatively stable after transplantation may not require frequent measurement of SCr in the first year. On the other hand, patients for whom the graft decays faster after the first year, a frequent schedule of SCr may be required to check the state of the graft. In this regard, instead of a common fixed schedule for all patients, we propose using a different schedule for every patient. More specifically, we propose using personalized schedules based on JMs Rizopoulos et al. (2016). This is because, JMs utilize random effects and thus they are inherently patient specific. In this direction, firstly a full specification the joint distribution of SCr levels and time of graft failure is obtained. It is then used to define a patient-specific posterior predictive distribution of time of graft failure, given the observed SCr measurements. The optimal time of the next SCr measurement is the one at which the expected information gained from an extra SCr measurement is maximum. In order to create reasonable predictions, SCr measurements for the first 3 months are taken as per the fixed schedule. This time period corresponds to the time around which we observed an increase in the SCr profile.

Since the SCr measurements are already taken for the kidney transplant patients, in order to demonstrate the efficacy of the personalized schedules we conduct a small simulation. To this end, we first assume a population of kidney transplant patients, whose SCr and hazard of graft failure follow a JM of the form described in Section 2, with parameters equal to the posterior mean of parameters estimated from the joint model fitted to the kidney transplant dataset. From this population we sample 625 patients, which are further split into a training (575 patients) and test (50 patients) part. For the training patients we generate a graft failure time T_i^* as well as a random and non-informative censoring time C_i. For the test patients the graft failure time T_j^* and an intervention time T_j^I is generated. The intervention time is the time at which the 6 month dynamic risk of graft failure of the patient becomes larger than a certain threshold κ. The choice of κ dictates the amount of time at hand between intervention and graft failure. In this simulation we evaluate two κ values, namely 0.05 and 0.025.

Our goal is to compare personalized schedule with the currently used fixed schedule of SCr measurements. To this end, we first fit a joint model of the specification described in Section 2 to the training data set and obtain a MCMC sample from the posterior distribution of the parameters of the JM. Using the fitted JM, we then iteratively schedule SCr measurements for the test patients, until the dynamic risk of graft failure (Rizopoulos, 2011) of the patients becomes larger than the threshold κ. Let N_j^I denote the number of SCr measurements conducted for the j-th test patient. The time difference between the observed intervention time due to the schedule (T_j^S) and the true intervention time, that is, the intervention offset is denoted by $O_j^I = T_j^S - T_j^I$. Lastly, the failure offset $O_j^* = T_j^S - T^* j$ is the time at hand between the observed intervention time and the time of graft failure. Using the test patients, we calculate these measures for both personalized and fixed schedules. It is to be noted that in the ideal scenario, N_j^I will be one, and offset O_j^I will be zero.

A boxplot of the observed values of the number of SCr measurements N_j^I, intervention offset O_j^I and failure offset O_j^* are presented in Figure 5, Figure 6 and Figure 7, respectively. The mean number of SCr measurements for personalized schedule is 14.46 whereas it is 27.60 for fixed schedule. In addition the standard deviation for number of SCr measurements is 9.17 for fixed schedule and 4.03 for personalized schedule. That is, personalized schedule not only schedule less N_j^I on average but the variation in N_j^I from patient to patient is also less. The mean absolute intervention offset for personalized schedules is 0.445 whereas for fixed schedules it is 0.448. The personalized schedule also has less standard deviation for absolute O_j^I, with it being 0.285 for personalized schedule and being 0.338 for fixed schedule. There is indeed a risk that either of the schedule can exceed the true graft failure time. In this regard, 12% of the times the graft failure is not detected for the test patients when fixed schedule is used. This rate is 14% when personalized schedule is used. Furthermore, the mean absolute failure offset is 2.71 for personalized schedule and 2.76 for fixed schedule. The standard deviation for absolute O_j^* is 1.97 for personalized schedule and 2.21 for fixed schedule.

In order to reduce the risk of overshooting the true graft failure time we propose that a smaller κ of 0.025 is used. The boxplot for the failure offset for this scenario is displayed in Figure 10. In this

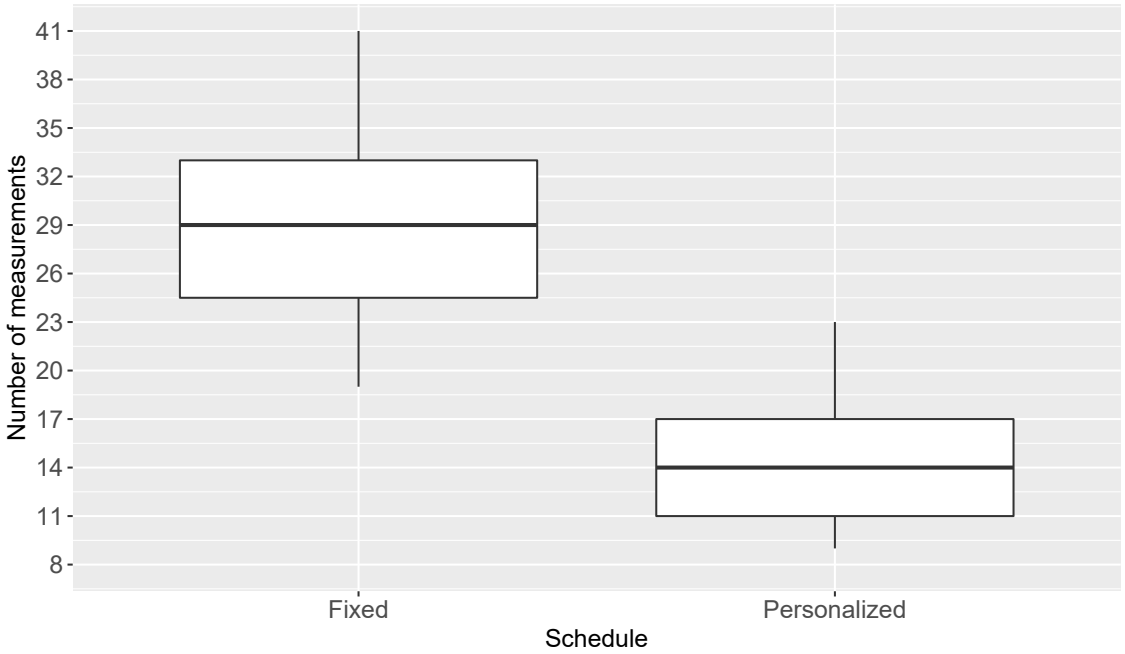

Figure 5: Boxplot of the number of SCr measurements N_j^I for the test patients, for $\kappa = 0.05$.

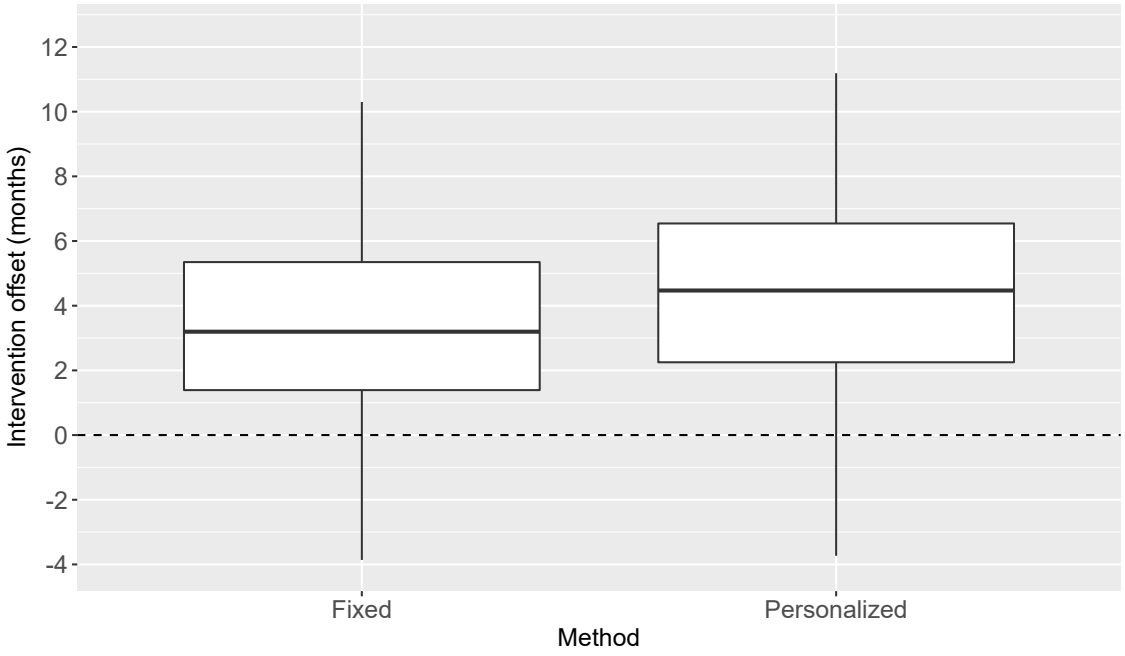

Figure 6: Boxplot of the intervention offset O_j^I for the test patients, for $\kappa = 0.05$. The zero offset mark is displayed with the dashed line.

scenario only for 6% of the patients the graft failure time is exceeded. Boxplot for number of SCr measurements N_j^I and intervention offset O_j^I are displayed in Figure 8 and Figure 9, respectively. However in this scenario, although the personalized schedule conducts less SCr measurements, it also exceeds the true intervention time more often than the fixed schedule.

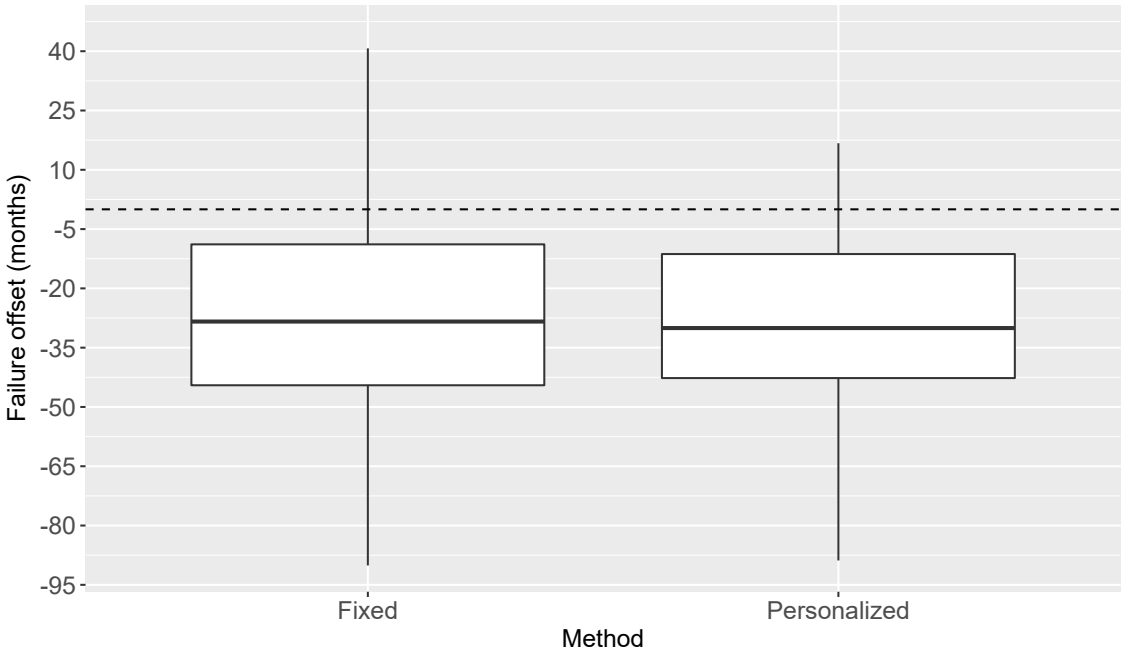

Figure 7: Boxplot of the failure offset O_j^* for the test patients, for $\kappa = 0.05$. The zero offset mark is displayed with the dashed line.

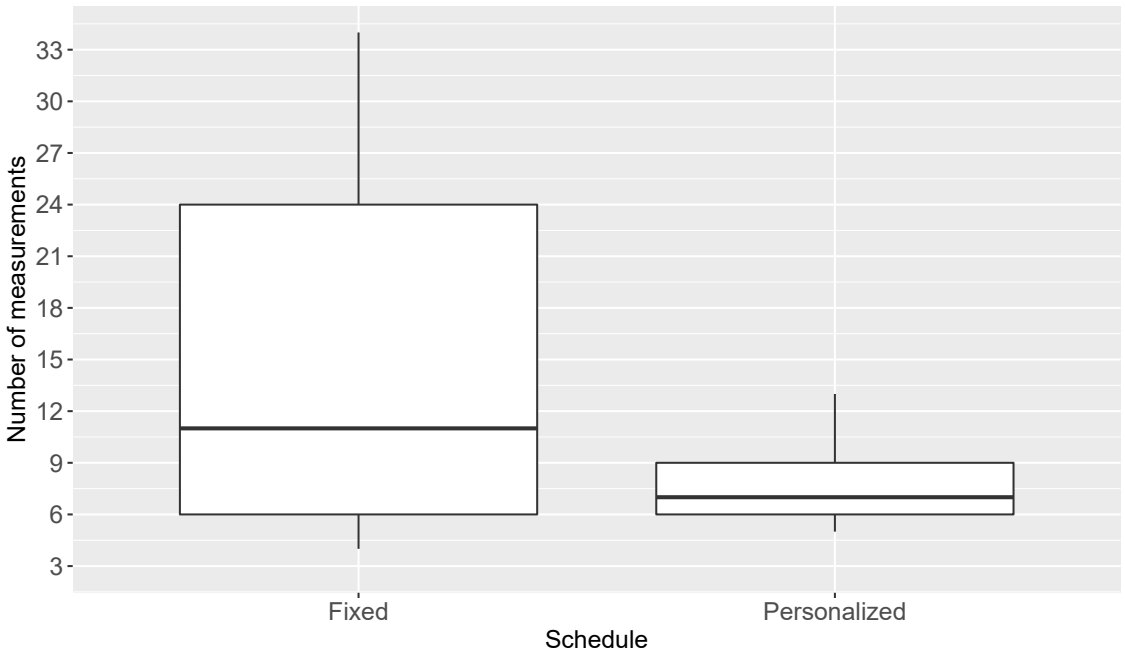

Figure 8: Boxplot of the number of SCr measurements N_j^I for the test patients, for $\kappa = 0.025$.

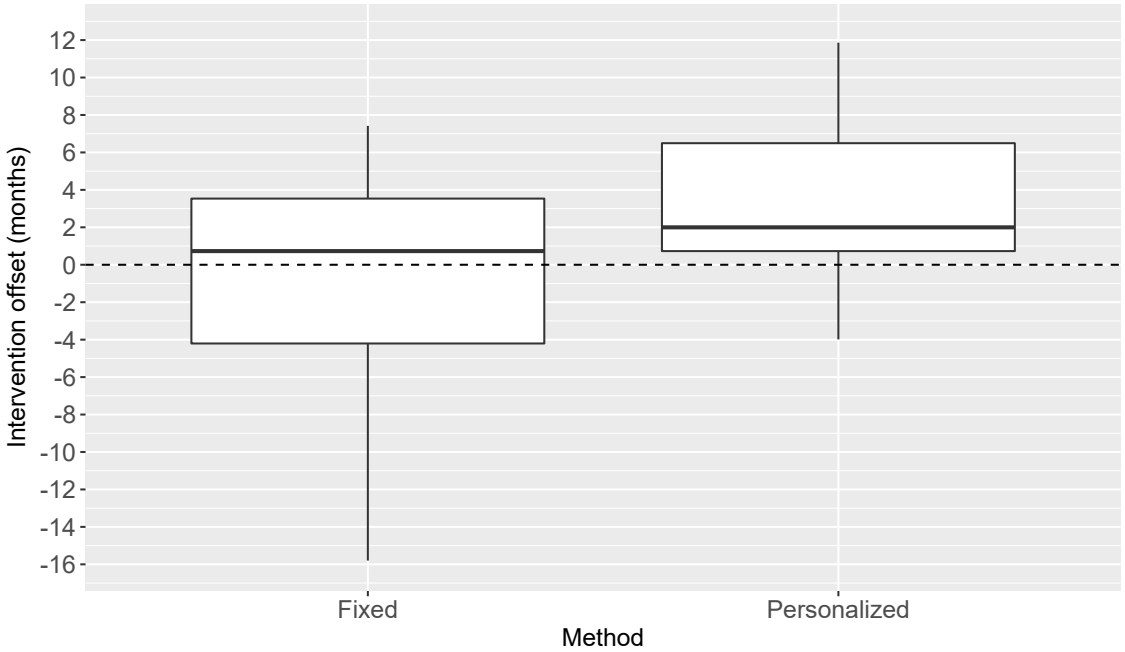

Figure 9: Boxplot of the intervention offset O_j^I for the test patients, for $\kappa = 0.025$. The zero offset mark is displayed with the dashed line.

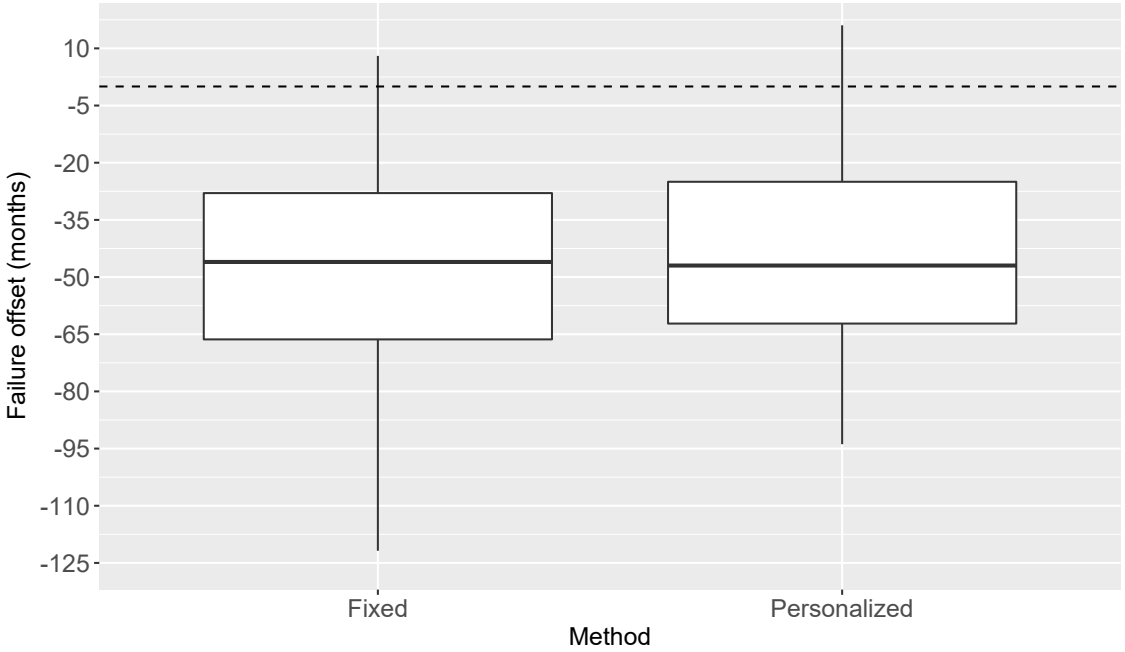

Figure 10: Boxplot of the failure offset O_j^* for the test patients, for $\kappa = 0.025$. The zero offset mark is displayed with the dashed line.

4 Supplementary

We start with the definition of the joint modeling framework that will be used to fit a model to the available dataset, and then to plan biomarker measurements for future patients. Let T_i^* denote the true graft failure time, and C_i denote the censoring time for the i-th patient enrolled in the kidney transplant aftercare. Let $T_i = \min(T_i^*, C_i)$ denote the observed graft failure time and $\delta_i = I(T_i^* < C_i)$ the event indicator for the i-th patient. $I(\cdot)$ is an indicator function that takes the value 1 when $T_i^* < C_i$ and 0 otherwise. Let $\boldsymbol{y}_{i1}$ and $\boldsymbol{y}_{i2}$ denote the log transformed $n_{i1} \times 1$ and $n_{i2} \times 1$ vectors of protein creatinine ratio (PCR) and serum creatinine (SCr) levels, respectively, for the i-th patient. For a sample of n patients the observed data is denoted by $\mathcal{D}_n = \{T_i, \delta_i, \boldsymbol{y}_{i1}, \boldsymbol{y}_{i2}; i = 1, \ldots, n\}$.

The two outcomes, PCR and SCr are continuous in nature and thus to model them the joint model utilizes a multivariate linear mixed effects model (LMM). The PCR outcome is modeled as (model for SCr is same):

$$\begin{aligned} y_{i1}(t) &= m_{i1}(t) + \varepsilon_{i1}(t) \\ &= \boldsymbol{x}_i^T(t)\boldsymbol{\beta}_1 + \boldsymbol{z}_i^T(t)\boldsymbol{b}_{i1} + \varepsilon_{i1}(t), \end{aligned}$$

where $\boldsymbol{x}_i(t)$ denotes the row vector of the common design matrix for fixed effects of both PCR and SCr, respectively, and $\boldsymbol{z}_i(t)$ denotes the same for random effects. The corresponding fixed effects are denoted by $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and the complete vector of random effects by $\boldsymbol{b}_i = (\boldsymbol{b}_{i1}, \boldsymbol{b}_{i2})^T$. The complete vector of random effects are assumed to be normally distributed with mean zero and variance-covariance matrix $\boldsymbol{D}$. The true and unobserved PCR and SCr levels at time t are denoted by $m_{i1}(t)$ and $m_{i2}(t)$, respectively. Unlike $y_{i1}(t), y_{i2}(t)$, the former are not contaminated with the measurement errors $\varepsilon_{i1}(t)$ and $\varepsilon_{i2}(t)$, respectively. The errors are assumed to be normally distributed with mean zero and variance σ_1^2 and σ_2^2, respectively, and are independent of the random effects $\boldsymbol{b}_{i1}, \boldsymbol{b}_{i2}$.

To model the effect of the two longitudinal outcomes on hazard of graft failure, joint models utilize a relative risk sub-model. The hazard of graft failure for patient i at any time point t, denoted by $h_i(t)$, depends on a function of subject specific linear predictors $m_{i1}(t)$, m_{i2} and/or the random effects:

$$\begin{aligned} h_i(t \mid \mathcal{M}_{i1}(t), \mathcal{M}_{i2}(t), \boldsymbol{w}_i) &= \lim_{\Delta t \to 0} \frac{\Pr\{t \le T_i^* < t + \Delta t \mid T_i^* \ge t, \mathcal{M}_{i1}(t), \mathcal{M}_{i2}(t), \boldsymbol{w}_i\}}{\Delta t} \\ &= h_0(t) \exp\left[\boldsymbol{\gamma}^T \boldsymbol{w}_i + f_1\{\mathcal{M}_{i1}(t), \boldsymbol{b}_{i1}, \boldsymbol{\alpha}_1\} + f_2\{\mathcal{M}_{i1}(t), \boldsymbol{b}_{i2}, \boldsymbol{\alpha}_2\}\right], \quad t > 0, \end{aligned}$$

where $\mathcal{M}_{i1}(t) = \{m_{i1}(v), 0 \le v \le t\}$, and $\mathcal{M}_{i2}(t) = \{m_{i2}(v), 0 \le v \le t\}$ denote the history of the underlying PCR and SCr levels, respectively, up to time t. The vector of baseline covariates is denoted by $\boldsymbol{w}_i$, and $\boldsymbol{\gamma}$ are the corresponding parameters. The function $f_1(\cdot), f_2(\cdot)$ parametrized by vectors $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2$ specify the functional form of PCR and SCr levels (Brown, 2009; Rizopoulos, 2012; Rizopoulos et al., 2014; Taylor et al., 2013) that are used in the linear predictor of the relative risk model. Some functional forms of PCR, relevant to the problem at hand are the following (functional forms for SCr are similar):

$$\begin{cases} f_1\{M_{i1}(t), \boldsymbol{b}_{i1}, \boldsymbol{\alpha}_1\} = \alpha_1 m_{i1}(t), \\ f_1\{M_{i1}(t), \boldsymbol{b}_{i1}, \boldsymbol{\alpha}_1\} = \alpha_{11} m_{i1}(t) + \alpha_{12} m'_{i1}(t), \quad \text{with } m'_{i1}(t) = \frac{\mathrm{d}m_{i1}(t)}{\mathrm{d}t}. \end{cases}$$

These formulations of $f_1(\cdot)$ postulate that the hazard of graft failure at time t may be associated with the underlying level $m_{i1}(t)$ of PCR (and/or SCr) at t, or with both the level and velocity $m'_{i1}(t)$ of PCR (and/or SCr) at t. Lastly, $h_0(t)$ is the baseline hazard at time t, and is modeled flexibly using P-splines. More specifically:

$$\log h_0(t) = \gamma_{h_0,0} + \sum_{q=1}^{Q} \gamma_{h_0,q} B_q(t, \boldsymbol{v}),$$

where $B_q(t, \boldsymbol{v})$ denotes the q-th basis function of a B-spline with knots $\boldsymbol{v} = v_1, \ldots, v_Q$ and vector of spline coefficients γ_{h_0}. To avoid choosing the number and position of knots in the spline, a relatively high number of knots (e.g., 15 to 20) are chosen and the corresponding B-spline regression coefficients γ_{h_0} are penalized using a differences penalty (Eilers and Marx, 1996).

4.1 Parameter Estimation

We estimate parameters of the joint model using Markov chain Monte Carlo (MCMC) methods under the Bayesian framework. Let $\boldsymbol{\theta}$ denote the vector of the parameters of the joint model. The joint model postulates that given the random effects, time to graft failure and longitudinal responses taken over time are all mutually independent. Under this assumption the posterior distribution of the parameters is given by:

$$\begin{aligned} p(\boldsymbol{\theta}, \boldsymbol{b} \mid \mathcal{D}_n) &\propto \prod_{i=1}^{n} p(T_i, \delta_i, \boldsymbol{y}_{i1}, \boldsymbol{y}_{i2} \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}) \\ &\propto \prod_{i=1}^{n} p(T_i, \delta_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) p(\boldsymbol{y}_{i1} \mid \boldsymbol{b}_{i1}, \boldsymbol{\theta}) p(\boldsymbol{y}_{i2} \mid \boldsymbol{b}_{i2}, \boldsymbol{\theta}) p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) p(\boldsymbol{\theta}), \\ p(\boldsymbol{b}_i \mid \boldsymbol{\theta}) &= \frac{1}{\sqrt{(2\pi)^q \det(\boldsymbol{D})}} \exp(\boldsymbol{b}_i^T \boldsymbol{D}^{-1} \boldsymbol{b}_i), \end{aligned}$$

where the likelihood contribution of PCR conditional on random effects is (contribution of SCr can be derived similarly):

$$\begin{aligned} p(\boldsymbol{y}_{i1} \mid \boldsymbol{b}_{i1}, \boldsymbol{\theta}) &= \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^{n_i}} \exp\left(- \frac{\|\boldsymbol{y}_{i1} - \boldsymbol{X}_{i1}\boldsymbol{\beta}_1 - \boldsymbol{Z}_{i1}\boldsymbol{b}_{i1}\|^2}{\sigma_1^2} \right), \\ \boldsymbol{X}_{i1} &= \{\boldsymbol{x}_{i1}(t_{i11})^T, \ldots, \boldsymbol{x}_{i1}(t_{i1n_i})^T\}^T, \\ \boldsymbol{Z}_{i1} &= \{\boldsymbol{z}_{i1}(t_{i11})^T, \ldots, \boldsymbol{z}_{i1}(t_{i1n_i})^T\}^T. \end{aligned}$$

The likelihood contribution of the time to graft failure outcome is given by:

$$p(T_i, \delta_i \mid \boldsymbol{b}_i, \boldsymbol{\theta}) = h_i(T_i \mid \mathcal{M}_{i1}(s), \mathcal{M}_{i2}(s), \boldsymbol{w}_i)^{\delta_i} \exp\left\{ - \int_0^{T_i} h_i(s \mid \mathcal{M}_{i1}(s), \mathcal{M}_{i2}(s), \boldsymbol{w}_i) \mathrm{d}s \right\}. \quad (3)$$

The integral in (3) does not have a closed-form solution, and therefore we use a 15-point Gauss-Kronrod quadrature rule to approximate it.

We use independent normal priors with zero mean and variance 100 for the fixed effects $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$, and inverse Gamma prior with shape and rate both equal to 0.01 for the parameters σ_1^2, σ_2^2. For the variance-covariance matrix $\boldsymbol{D}$ of the random effects we take inverse Wishart prior with an identity scale matrix and degrees of freedom equal to q (number of random effects). For the relative risk model's parameters $\boldsymbol{\gamma}$ and the association parameters $\boldsymbol{\alpha}$, we use a global-local ridge-type shrinkage prior. For example, for the s-th element of $\boldsymbol{\alpha}$ we assume (similarly for $\boldsymbol{\gamma}$):

$$\alpha_s \sim \mathcal{N}(0, \tau\psi_s), \quad \tau^{-1} \sim \text{Gamma}(0.1, 0.1), \quad \psi_s^{-1} \sim \text{Gamma}(1, 0.01).$$

The global smoothing parameter τ has sufficiently mass near zero to ensure shrinkage, while the local smoothing parameter ψ_s allows individual coefficients to attain large values. For the penalized version of the B-spline approximation to the baseline hazard, we use the following prior for parameters γ_{h_0} (Lang and Brezger, 2004):

$$p(\gamma_{h_0} \mid \tau_h) \propto \tau_h^{\rho(\boldsymbol{K})/2} \exp\left(- \frac{\tau_h}{2} \gamma_{h_0}^T \boldsymbol{K} \gamma_{h_0} \right),$$

where τ_h is the smoothing parameter that takes a Gamma(1, 0.005) hyper-prior in order to ensure a proper posterior for γ_{h_0}, $\boldsymbol{K} = \Delta_r^T \Delta_r + 10^{-6}\boldsymbol{I}$, where Δ_r denotes the r-th difference penalty matrix, and $\rho(\boldsymbol{K})$ denotes the rank of $\boldsymbol{K}$.

4.2 Estimated Covariance Matrix of Random Effects

The estimated covariance matrix of random effects D is presented below. Covariance parameters corresponding to random effects of PCR outcome are denoted by $b_{1k}, k = \{0, \dots 4\}$, whereas $b_{2k}, k = \{0, \dots 4\}$ denote the random effects for SCr outcome.

$$
\mathbf{D} = \begin{array}{c} \begin{array}{cccccccccc} b_{10} & b_{11} & b_{12} & b_{13} & b_{14} & b_{20} & b_{21} & b_{22} & b_{23} & b_{24} \end{array} \\ \left[\begin{array}{cccccccccc} 1.050 & & & & & & & & & \\ -0.833 & 1.651 & & & & & & & & \\ -0.483 & 0.137 & 3.158 & & & & & & & \\ -1.114 & 1.284 & 1.264 & 3.749 & & & & & & \\ -0.407 & 0.844 & -0.609 & 2.930 & 5.609 & & & & & \\ 0.251 & -0.221 & -0.110 & -0.289 & -0.085 & 0.328 & & & & \\ -0.228 & 0.258 & 0.042 & 0.253 & 0.093 & -0.277 & 0.361 & & & \\ -0.089 & 0.087 & 0.567 & 0.447 & -0.016 & -0.185 & 0.137 & 0.592 & & \\ -0.322 & 0.347 & 0.468 & 0.936 & 0.575 & -0.504 & 0.494 & 0.531 & 1.229 & \\ -0.060 & 0.178 & 0.064 & 0.774 & 1.357 & -0.162 & 0.165 & 0.069 & 0.537 & 0.837 \end{array}\right] \end{array} \begin{array}{c} b_{10} \\ b_{11} \\ b_{12} \\ b_{13} \\ b_{14} \\ b_{20} \\ b_{21} \\ b_{22} \\ b_{23} \\ b_{24} \end{array}
$$

References

Brown, Elizabeth R. (2009). "Assessing the association between trends in a biomarker and risk of event with an application in pediatric HIV/AIDS". In: *The Annals of Applied Statistics* 3.3, pp. 1163–1182.

Eilers, Paul HC and Brian D Marx (1996). "Flexible smoothing with B-splines and penalties". In: *Statistical Science* 11.2, pp. 89–121.

Fournier, Marie-cécile et al. (2016). "A joint model for longitudinal and time-to-event data to better assess the specific role of donor and recipient factors on long-term kidney transplantation outcomes". In: *European Journal of Epidemiology* 31.5, p. 469.

Lang, Stefan and Andreas Brezger (2004). "Bayesian P-splines". In: *Journal of Computational and Graphical Statistics* 13.1, pp. 183–212.

Rizopoulos, Dimitris (2011). "Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data". In: *Biometrics* 67.3, pp. 819–829.

— (2012). *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. CRC Press.

— (2016). "The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC". In: *Journal of Statistical Software* 72.7, pp. 1–46.

Rizopoulos, Dimitris, Geert Molenberghs, and Emmanuel M.E.H. Lesaffre (2017). "Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking". In: *Biometrical Journal*. doi:10.1002/bimj.201600238.

Rizopoulos, Dimitris et al. (2014). "Combining Dynamic Predictions From Joint Models for Longitudinal and Time-to-Event Data Using Bayesian Model Averaging". In: *Journal of the American Statistical Association* 109.508, pp. 1385–1397.

Rizopoulos, Dimitris et al. (2016). "Personalized screening intervals for biomarkers using joint models for longitudinal and survival data". In: *Biostatistics* 17.1, pp. 149–164.

Taylor, Jeremy M.G. et al. (2013). "Real-time individual predictions of prostate cancer recurrence using joint models". In: *Biometrics* 69.1, pp. 206–213.

Tsiatis, Anastasios A and Marie Davidian (2004). "Joint modeling of longitudinal and time-to-event data: an overview". In: *Statistica Sinica* 14.3, pp. 809–834.