Joint Survival Analysis of Time to Drug Change and a Terminal Event with Application to Drug Failure Analysis using Transplant Registry Data

Elizabeth Renouf[1,*], C.B. Dean[1], David R. Bellhouse[1] and Vivian C. McAlister[2]

[1]*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, N6A 5B7, Canada*

[2]*Department of Surgery, University of Western Ontario, London, Ontario, N6A 5B7, Canada*

Abstract: Statistical approaches for drug effectiveness studies after liver transplant have used a survival model with changes in treatment as a time-dependent covariate. However, the approach requires that changes in the time-dependent covariate be unrelated to survival outcome. Usually this is not the case, as one drug may be discontinued and an alternative chosen due to the declining health status of the patient. Other approaches examine only subjects who remain on the same drug over a time window, which discards valuable data and may lead to biased effects since this excludes data related to early deaths and to individuals who perform poorly on the drug and had to switch treatments. Because of these issues there are conflicting results seen in the evaluation of immunosuppressive drug effectiveness after liver transplant. We propose a joint survival outcome model with a time-to-drug-change event and a terminal event in graft failure that is useful in drug effectiveness studies where subjects are discontinued from an immunosuppressant (in favour of alternative treatment) due to health reasons. We also include a longitudinal biomarker component. The model takes account of the dependencies across out- comes through shared random effects. Using a Markov chain Monte Carlo approach, we fit the joint model to data from liver transplant recipients from the Scientific Registry for Transplant Recipients.

Keywords: Joint models, longitudinal, survival, transplant, joint outcome.

1. INTRODUCTION

In some fields of medicine, the dynamics of treatment and treatment failures require special consideration in statistical analysis. Transplantation is one of these fields. Transplant recipients must maintain a lifelong immunosuppressive regime, and a failure of one regime requires a change to an alternative drug. A failure in immunosuppressive treatment regime is usually related to patient and graft survival, since the change in treatment is on the causal pathway to graft failure. Some analyses have used a standard Cox proportional hazards survival model with treatment changes included as a time-dependent covariate [1,2]. However, assumptions of this model require that changes in the time-dependent covariate be unrelated in this way to the survival outcome. Ignoring this assumption may lead to biased results [3]. In transplantation, one drug may be discontinued and an alternative chosen due to graft rejection or the declining health of the patient for some other reason and so drug failures are typically indeed related to the survival outcomes. The primary objective of this paper is to develop a method for analyzing the efficacy of immunosuppressive treatment after liver transplant that accounts for potential dependencies between drug failures and survival. Since the change in drug may be directly related to health status, we propose that treating it as a time-to-event process in a joint outcome model with time-to-graft-failure will provide less biased results. Modelling drug failures as a time-to-event process has not been considered in analyses of survival after organ transplant in registry data.

The joint modelling of two time-to-event processes can be used for drug studies where the duration on and survival of a patient on a particular drug is of interest, especially in situations where the termination of the initial treatment drug is driven by the deteriorating health status of the patient. There is an association between failures of immunosuppressive treatment and graft survival, since treatment failures are likely triggered by another event that increases the risk of death or graft failure (e.g., organ rejection, or cancer occurrence). We suggest that jointly modelling graft survival and time-to-drug-failure in a joint survival outcome model with frailty terms linking the two outcomes is preferable in this situation. We consider an individual frailty to account for unobserved heterogeneity between patients. In addition, we are able to add a longitudinal component to thejoint model with the variable creatinine, which is collected every year after transplant. Creatinine is important both as a general health indicator, as well as an indicator of

*Address correspondence to this author at the Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, N6A 5B7, Canada; Tel: (519) 661-3607; Fax: (519) 661-3813;
E-mail: erenouf@uwo.ca

E-ISSN: 1929-6029/16

treatment drug toxicity which may also lead to treatment switching.

Much of the research in modelling two time-to-event processes started with the joint analysis of recurrent events and a terminating event. These models allow for a recurrent and a terminal event that are not independent. An important early paper in this proposed a Cox proportional hazards model with shared frailty for recurrent events and a terminal event in an MCMC approach [4]. The unobserved frailty in the model measures the latent health status of the patient and it is related to both the recurrent event and the terminal event. A shared frailty joint model accommodating multivariate longitudinal and bivariate time-to-event data with extension to a multivariate survival component was proposed in the setting of cure fractions [5]. Other important work in the area of multiple time-to-event processes includes [6] where the authors propose a joint frailty model using a non-parametric likelihood method. Both [4] and [6] offer a thorough review of the history of research in this area from the 1990s to the mid 2000s. Another important paper uses two additive shared frailties to model trial and treatment heterogeneity in a meta-analysis [7]. Liu and Huang [8] is one of the few papers to consider repeated longitudinal events and two or more time-to-event processes, using a shared random effects model. The hazard of the terminal event (death) depends on both the longitudinal random effects (CD4 counts) and the frailty term from the recurrent event (infection). More recently, Musoro [9] proposed as hared frailty model for multiple longitudinal outcomes and multiple repeated events (infections) without a terminal event using an MCMC approach for inference.

The use of a joint modelling approach for modelling two event processes which did not includea recurrent event, but one that must come before the other, was first undertaken by Elashoff [10] in a competing risks setting. Methods for multiple failure times in the setting of competing risks and semi-competing risks data have become very popular, with further papers from Elashoff [11] followed by Williamson [12] and many others. Most recently, illness-death models have been applied to semi-competing risks data [13-15].

In this paper we utilize a joint model with a longitudinal component in log (creatinine), and abivariate survival model comprised of a time-to-drug-failure process and a time-to-graft-survival process. This is a novel approach to drug efficacy analysis in transplantation. While [8] considered a longitudinal outcome with a recurrent process and a terminal event, a model with two survival events and a longitudinal component has not been previously considered. We hypothesize that higher creatinine levels are correlated with a greater risk of drug failure and also with greater risk of failure of the transplanted liver. We also hypothesize that having a change from initial drug therapy is associated with a greater risk of graft failure. In Section 2.1 we describe the development of our two time-to-event process joint model. In Section 3 we motivate the need for this type of model in the analysis of graft survival after liver transplantation. We end with a final section containing a discussion and future projects.

2. STATISTICAL METHODS

2.1. Joint Longitudinal and Survival Sub-Models

The longitudinal component related to modelling the trajectory of log(creatinine) is a longitudinal mixed effects model that is linked to both survival models via random effects. We examine graft failure up to three years post transplant. Let y_{ij} represent the longitudinal marker log(creatinine) for subject i, i = 1, ..., n, at time-point m_{ij}, j = 1, ...n_i. The mixed effects model is

$$y_{ij} = \beta_0 + \beta_1 m_{ij} + \boldsymbol{x}'_{ij}\boldsymbol{\beta} + b_{0i} + b_{1i} m_{ij} + \in_{ij} \; j = 1, \ldots, n_i; i = 1, \ldots, n$$

where β_0 and β_1 are the intercept and slope; x_{ij} are a set of covariates with respective vectors of regression parameters β. The random effects b_{0i}, b_{1i} are independent and normally distributed. Let $b_i = (b_{0i}, b_{1i})$. The variance-covariance matrix of b_i, D = diag (ν_1, ν_2). Then we assume the measurement error $\varepsilon_{ij} \sim N(0, \sigma_2)$ is independent of the b_i.

The longitudinal process influences two time-to-event processes: time-to-drug-failure and time-to-graft-failure. For the ith individual, let t_{ki} represent lifetime, k = 1,2, with t_{1i} representing time-to-drug-failure with shape parameter α_1, and t_{2i} representing time-to-graft-failure with shape parameter α_2. The scale parameter λ_{ki} can vary across individuals and events types. We re-parameterize λ_{ki} as

$$\log(\lambda_{ki}) = \boldsymbol{z}'_{\boldsymbol{ki}}\boldsymbol{\gamma}_{\boldsymbol{k}} + W_{ki} + \phi_k c_i k = 1,2$$

where z_{ki} and γ_k are the covariates and corresponding regression coefficients and where $\varphi_1 = \varphi$. The W_{ki} link the random effects from the longitudinal model to the time-to-drug-failure model, where $W_{1i} = \xi_1 \hat{b}_{0i} + \xi_2 \hat{b}_{1i}$, and the time-to-graft-failure model $(W_{1i} = \rho_1 \hat{b}_{0i} + \rho_2 \hat{b}_{1i})$.

The ζ and ρ are association parameters measuring the relationship between the vector of random effects bi and the time-to-event processes. The c_i are subject-specific frailty terms which link the two responses. In the graft failure component, φ measures the association between c_i and risk of graft failure. We assume that the repeated measures of creatinine are correlated through the random effects b_i, and the hazard of the terminal event depends on the longitudinal component and the time-to-drug-failure component through b_i and c_i respectively. The frailty term c_i is assumed independent of b_i.

The data we use is taken from the Scientific Registry for Transplant Recipients (SRTR). The SRTR data system includes data on all donor, wait-listed candidates, and transplant recipients in the U.S., submitted by the members of the Organ Procurement and Transplantation Network (OPTN). The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors. A detailed description of the database is available in [16].

3. APPLICATION: A JOINT MODEL FOR THREE YEAR GRAFT SURVIVAL AFTER LIVER TRANSPLANT USING REGISTRY DATA

The goal of our investigation is to apply a new method for assessing effectiveness of immuno-suppression after liver transplantation, and to identify risk factors for graft failure while taking into account the initial immunosuppression and time on a particular drug or drug combination. Through joint modelling of the time-to-drug-failure and time-to-graft-failure processes, we can account for the association between time on initial drug therapy and graft survival. Time-to-drug-failure can be extended to a recurrent event format following the methods of [6,8] or others, however we provide an example with the time to first occurrence only, to provide a simplified model that establishes whether there is an association between creatinine, time on initial drug, and graft survival. We hypothesize that higher creatinine values (i.e. larger random effects) and a shorter time-on-initial-drug (i.e. larger frailty) increase the risk of graft failure. The time-to-drug-failure frailty serves as a surrogate indicator for many possible health events since someone who is doing well after transplant is less likely to experience a change inimmunosuppressive regime.

Estimation was carried out with a Bayesian approach and a Markov chain Monte Carlo (MCMC) algorithm in R [17] and Just Another Gibbs Sampler [18] to obtain estimates of the posterior distributions, and with the priors specified in JAGS as non-informative where the components of γ_1 and γ_2 are ~ N(0,10000), α_k ~ Unif(0,1), and ρ_1, ρ_2 and ζ_1, ζ_2 are N(0,10000). The c_i ~ N(0,1/σ_c), with σ_c ~ Unif(0,10000). The prior distributions for the random effects bi were specified as N(0, D) with the diagonal components of D ~ Unif(0,10000). We ran three chains for 500,000 iterations with 400,000 burn-in, which took approximately 30 hours. Convergence was judged by the Brooks-Gelman-Rubin (BGR) convergence diagnostic [19]. We used the deviance information criterion (DIC) for model comparison. Covariates were removed from the model if the credible interval contained zero.

We analyzed those patients age 16 and older, receiving a first cadaveric liver transplant in the United States between January 1, 2000 and December 31, 2002. There were 10,015 subjects, with 1,757 (17.5%) events (graft failure, defined as death or retransplant) in the first three years post-transplant. Immuno-suppressive therapy is recorded at baseline and discharge from hospital after transplant, then at six months, and yearly. Exact dates for treatment failures are not given, so the data are interval censored. To simplify the analysis we took the midpoint of the interval as the date of treatment change. If a subject also died during the interval where a treatment change was recorded, we took the midpoint of the start of the interval and the date of death as the day of treatment change. We do not use the data collected on the date of death or retransplant, since this would introduce bias as only those who die have a measurement at this unspecified time-point. If we could not determine baseline treatment, the subject was removed from the analysis. We chose to analyze baseline immuno-suppression by combination (rather than as any exposure or not, as in [2, 20]), since this accounts for any interaction or synergistic effects between SRL and the calcineurin inhibitors (CYA or TAC). This also allows us to see the effects of SRL uncontaminated by other drugs.

4. RESULTS

The SRTR collects immunosuppressive therapy for only the first five years after transplant. Most of the changes from initial therapy occur within the first year. Of the 10,015 subjects without missing data, 2,468 (25%) changed their initial therapy within the first three years post-transplant. Table **1** shows the number of

subjects on each drug combination and how many changed therapies by year 3. This table shows the great disparity in treatment switching among the various regimes.

Table 1: Summary of the Number of Individuals (n) on Initial Drug Therapies and the Number and Percent that Changed (n Changed (%)) from Initial Therapy within Three Years Post-Transplant

Initial treatment	n	n changed (%)
TAC	8218	1554 (18.9%)
SRL + TAC	533	405 (76.0%)
SRL	149	26 (47.7%)
CYA	1026	416 (40.5%)
SRL + CYA	89	67 (75.3%)

Treating drug failure as a time-to-event process is important to acknowledge the large amount of drug switching taking place in every group except TAC and to account for it in the analysis. We only examine primary immunosuppressive therapy in this model. For the sake of simplicity, induction therapy such as thymoglobulin or anti-CD25 antibody, and adjuvant immunosuppression such as steroids, azathioprine, or mycophenolate mofetil, are not considered at this time. Figure **1** shows Kaplan-Meier curves for graft survival by initial drug therapy. Those started on CYA + SRL, SRL alone, or TAC alone enjoy the best survival. Figure **2** shows Kaplan-Meier curves for time to drug failure for each therapy. Here we see that SRL alone or in combination (SRL + TAC, SRL + CYA) has the shortest time to drug failure. Of the 2,468 who terminated initial therapy, 80% did so in the first year post-transplant.

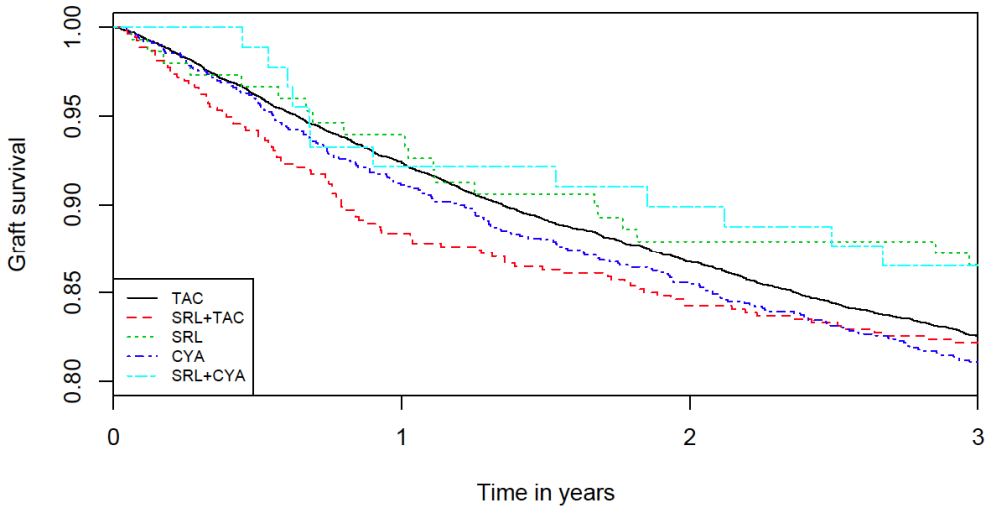

Figure 1: Graft survival by initial immunosuppression.

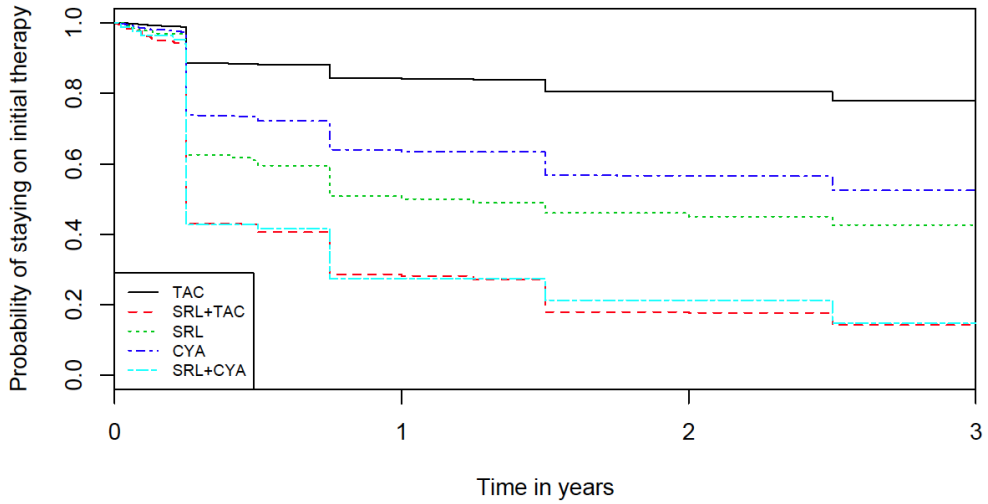

Figure 2: Time-to-drug-failure, by initial immunosuppression.

We analyzed HCV positive and HCV negative subjects separately on account of non-proportional hazards in this variable. There were 4,513 subjects (952 graft failure events) who were HCV positive and 5,502 subjects (805 graft failure events) who were HCV negative. Recurrence of HCV is universal after transplant, however the speed at which it progresses depends on different factors such as viral load, donor age and other risks that are not completely understood. There is growing evidence that this includes immunosuppressive regime [21, 22]. Table **2** shows the number of subjects on each immunosuppressive regime in each cohort. Other variables considered included recipient and donor age in decades, recipient and donor gender, gender mismatch between recipient and donor, donor BMI, whether a non-heart-beating donor was used, hepatocellular carcinoma (HCC), previous malignancy (pre-transplant), recipient race, recipient blood type, whether a split liver was received, and whether the recipient was infulminant hepatic failure at the time of transplant. Donor BMI was missing in three subjects so the average was substituted. We also tested for covariate interactions with HCC status. We incorporated into our analysis longitudinal values of creatinine up to three years post-transplant.

Table 2: Summary of Initial Drug Therapies by HCV Status

HCV positive cohort	
Initial treatment	**n (%)**
TAC	3741(82.9%)
SRL + TAC	223 (4.9%)
SRL	69 (1.5%)
CYA	440 (9.7%)
SRL + CYA	40 (0.9%)
HCV negative cohort	
Initial treatment	**n (%)**
TAC	4477 (81.3%)
SRL + TAC	310 (5.6%)
SRL	80 (1.5%)
CYA	586 (10.7%)
SRL + CYA	49 (0.9%)

Missing values were allowed in creatinine since the longitudinal trajectory can still be estimated. There were a maximum of 5 longitudinal values (day of transplant, 6 months, year 1, year 2, year 3). The mean number of creatinine observations per subject was 3.3 and the median was 3. We found a lower DIC when we did not include the slope random effect from the longitudinal component in both time-to-event processes, and so the models for both cohorts included only an intercept random effect in the longitudinal component of the model. This could be because the value of the overall slope β_1 and its estimated random effects are very small (0.006) in the longitudinal model, or possibly because in our three year analysis there are not enough repeated measurements of creatinine to improve the model when this random effect is included. The final model for each cohort was chosen after considering DIC as well as Cox Snell residuals for the graft survival component.

Results from the best fitting joint model for the HCV positive cohort are shown in Table **3**. Most of the covariates listed in Table **3** have a significant effect on time-to-graft-failure or time-to-drug-failure. The covariate gender was kept in the longitudinal model because it improved the DIC even though the credible interval contained zero. The random effect b_{0i} was an important linkage term between the longitudinal and the graft failure sub-models. In the longitudinal component, we saw a lower log(creatinine) over time in female subjects, while a higher log (creatinine) was seen in older subjects, those in fulminant failure at time of transplant, those who were diabetic at transplant, and those with a higher donor BMI. The only covariate significant in the time-to-drug-failure sub-model was treatment. Being on any initial treatment other than TAC significantly shortens the time to change of initial immunosuppressive therapy. The greatly increased risk of drug failure from initial immunosuppressive regime fits with the degree of treatment switching seen. The linkage parameter ζ was not significantly associated with to time-to-drug-failure.

In the time-to-graft-failure component, we found that none of the immunosuppressive treatments are significantly different from the reference drug TAC in terms of impact on graft survival, after taking into account covariates and the time-to-drug-failure frailty effect. The model is structured so that drug effects are tested in comparison to TAC, the most commonly used treatment. Risk of graft failure is increased for those who are of African American race compared to all other races, for those who had any previous malignancy, and for increased donor age. We found no significant effect for HCC status in the model, even with the variable for previous malignancy removed, and no interaction between initial drug therapy and HCC status.

Table 3: Posterior Means for the Log Hazard, Hazard, Standard Error and Quantiles from the HCV Positive Cohort Joint Model. (ns=not significant; AA = African American; F = female)

Parameter	Mean	Hazard	Std err	CI (lower)	CI (upper)
Longitudinal sub-model:					
intercept	-0.056		0.026	-0.105	-0.006
slope	0.006		0.0002	0.006	0.006
age in decades	0.004		0.0005	0.003	0.005
fulminant failure	0.361		0.116	0.138	0.589
gender (F)	-0.116		0.010	-0.137	-0.096 (ns)
diabetes	0.068		0.018	0.032	0.102
race (AA)	0.032		0.006	0.022	0.044
baseline trt: TAC (ref)	---		---	---	---
SRL + TAC	0.012		0.021	-0.028	0.055 (ns)
SRL	0.021		0.036	-0.047	0.089 (ns)
CYA	0.040		0.015	0.010	0.069
CYA + SRL	-0.038		0.047	-0.128	0.054 (ns)
ν_1(variance of b_{0i})	0.064			0.060	0.067
Time-to-drug-failure sub-model:					
intercept	-9.761		0.097	-9.952	-9.580
baseline trt: TAC alone (ref)	---		---	---	---
SRL + TAC	7.522		0.401	6.755	8.310
SRL	3.705		0.609	2.545	4.895
CYA	3.022		0.251	2.527	3.514
CYA + SRL	9.523		1.102	7.436	11.715
σ_c (variance of c_i)	14.929			13.682	16.374
Time-to-graft-failure sub-model:					
α (shape parameter)	0.933		0.028	0.879	0.986
intercept	-8.681		0.215	-9.088	-8.263
donor age (decades)	0.016	1.016	0.002	0.012	0.020
race (AA)	0.182	1.200	0.036	0.114	0.248
gender (F)	0.128		0.075	-0.021	0.272 (ns)
previous malignancy	0.451	1.568	0.107	0.239	0.660
baseline trt: TAC (ref)	--		--	--	--
SRL + TAC	-0.101		0.158	-0.426	0.190 (ns)
SRL	-0.462		0.321	-1.146	0.142 (ns)
CYA	0.040		0.108	-0.184	0.243 (ns)
CYA + SRL	-0.210		0.373	-0.977	0.493 (ns)
$\rho_{\square}$	0.447	1.564	0.160	0.131	0.770
ϕ	0.046	1.047	0.011	0.023	0.068

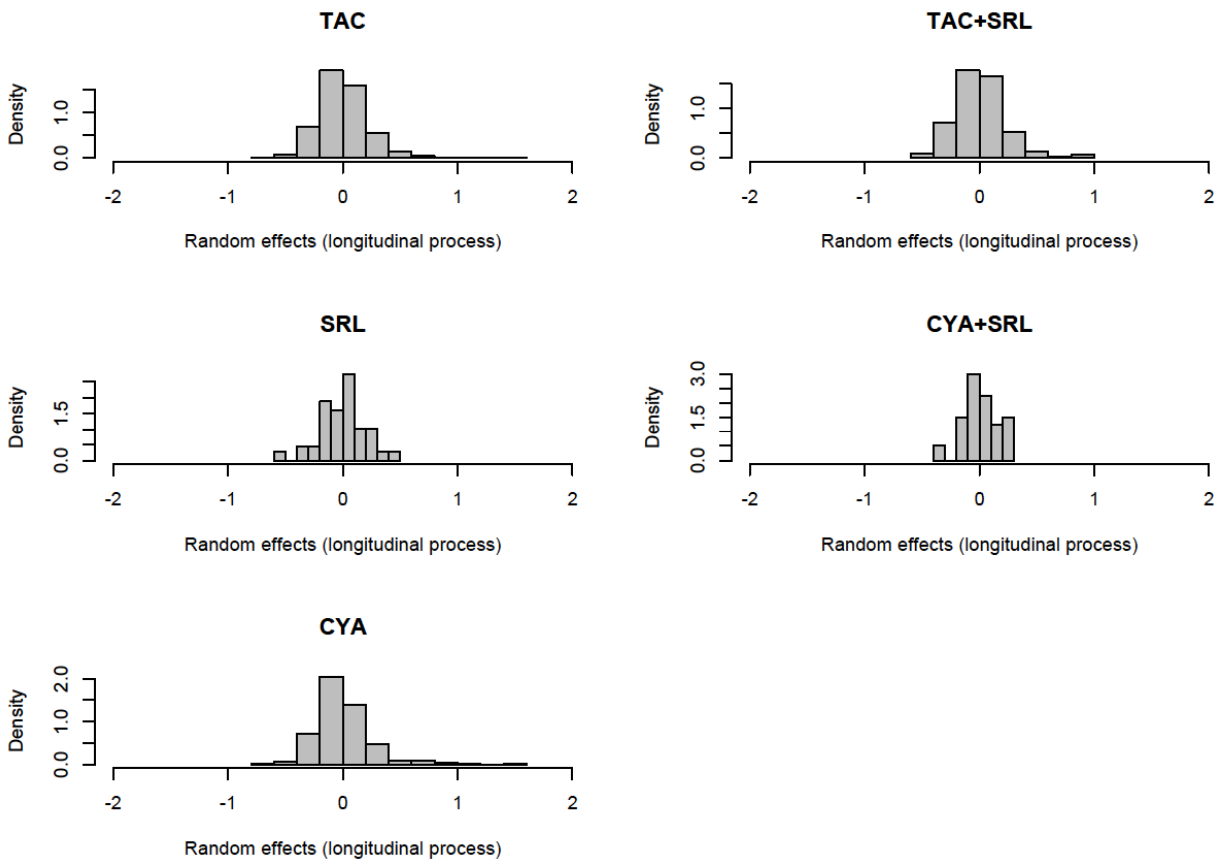

Figure 3: HCV positive cohort: Histograms of subject-specific intercept random effects from the longitudinal model, by initial treatment. Positive random effects are associated with increased risk of graft failure.

The subject-specific intercept random effect b_{0i} from the longitudinal component has a significant link with graft survival through the association parameter (ρ_1), with a hazard ratio of 1.567 for each one unit change in the random effect. The random effects themselves range from -0.69 to 1.43, and so the hazard can be as beneficial as 0.73 for those with lower than average creatinine levels, and as high as 1.90. When grouped by drug the random effects do not show any distinct patterns (see Figure **3**). The individual log frailty c_i from the time-to-drug-failure component has a significant link through the association parameter φ to the graft survival model in the HCV positive cohort, where it increases survival time for those with negative frailties, and decreases survival time for those with positive log frailties. Negative log frailties are associated with a lower risk of a drug failure event. Therefore, failure of initial therapy is associated with shorter graft survival. This is medically sensible since a change in treatment is often precipitated by an adverse event such as rejection.

Histograms of individual log frailties grouped by whether a drug failure occurred are shown in Figure **4**, and plots grouped by initial treatment are shown in Figure **5**. The value of the frailty can be quite large, with a range of approximately -6 to +7, so the estimated coefficient for the individual frailty in the survival model, while small at 0.046 (hazard: exp(0.046 × log frailty)), can have a large effect depending on the value of the subject specific frailty, with a hazard ratio ranging from a beneficial 0.76, to an increased hazard of 1.38. Note however, that there is quite a wide range of individual frailties for all drug therapies (see Figure **5**). The variance of the subject-specific frailty is 10.5, which shows there is a great deal of unobserved heterogeneity between subjects. To put the results into context, Table **4** shows these effects for each treatment combination, comparing the mean, minimum and maximum frailties of subjects who did not change therapy to those who did. For the group on initial therapy of TAC alone who did experience a drug failure event, the mean frailty shows the highest risk of graft failure (mean hazard: 1.25), suggesting that subjects who do not do well on the 'gold standard' TAC are at greatest risk of graft failure.

Results from the best fitting joint model for the HCV negative cohort are shown in Table **5**. The random effect for the intercept b0iwas seen as an important

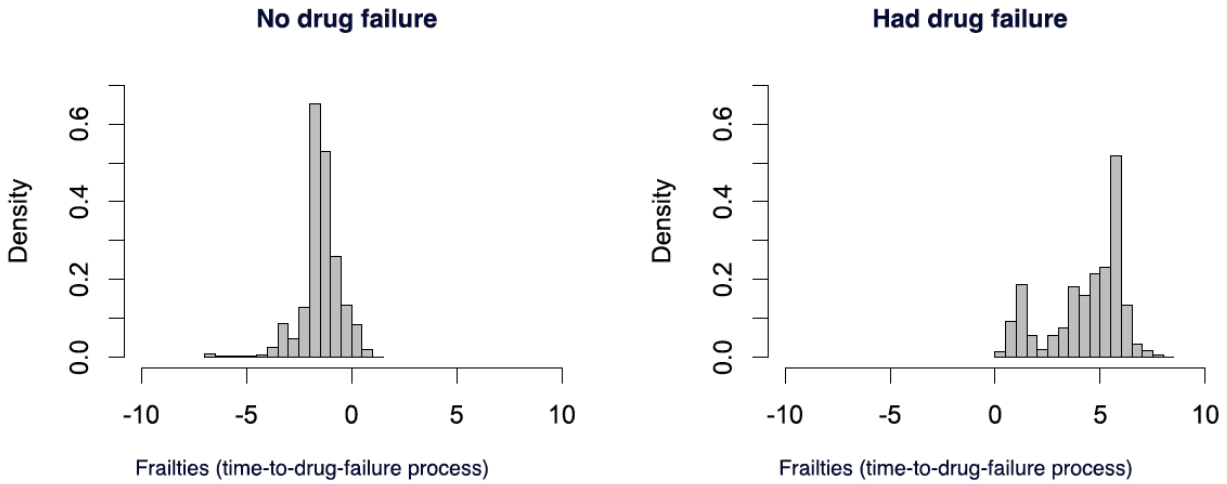

Figure 4: HCV positive cohort: Histograms of individual log frailties from time-to-drug-failure model, by whether initial treatment was changed. Positive frailties are associated with increased risk of graft failure.

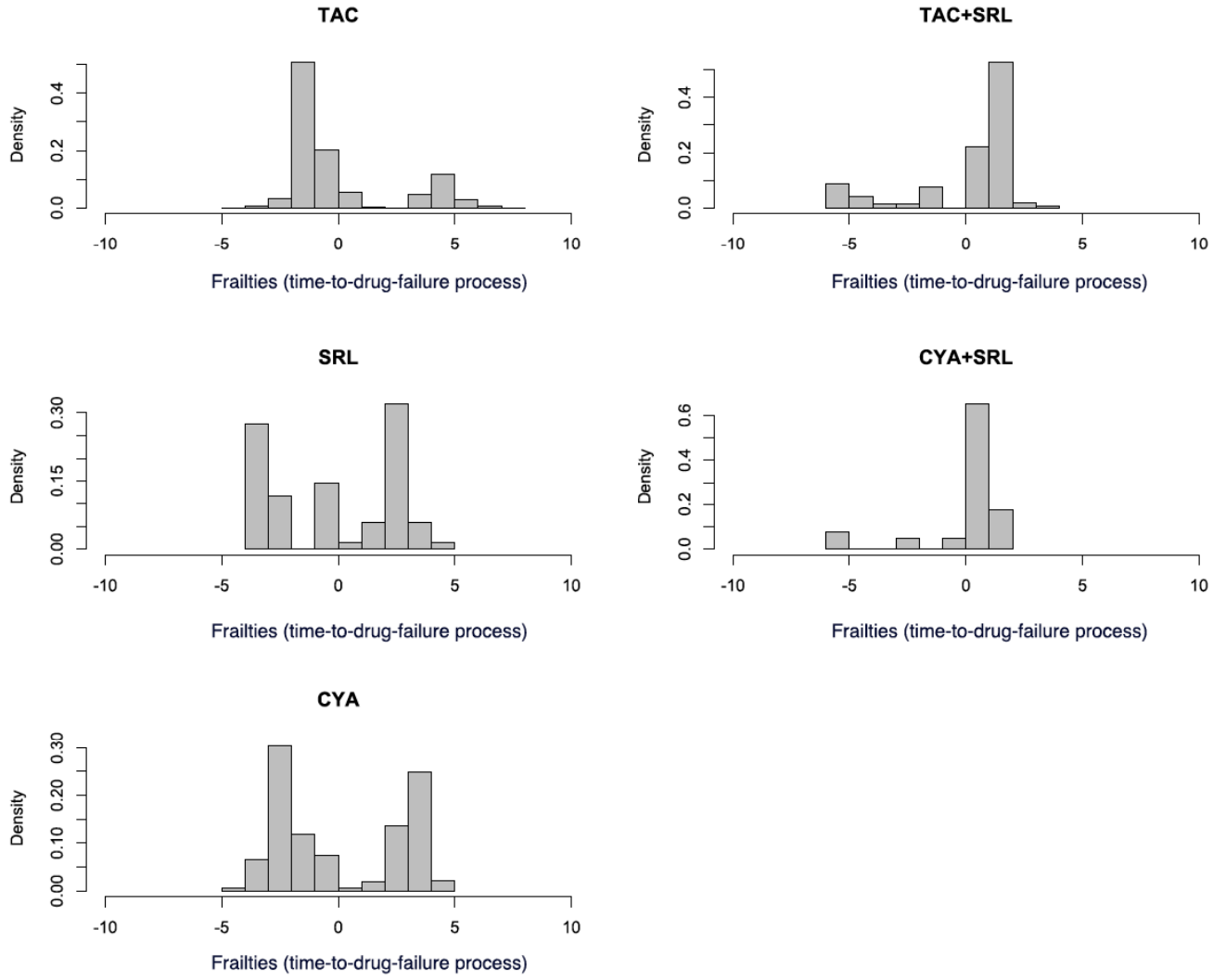

Figure 5: HCV positive cohort: Histograms of individual log frailties from time-to-drug-failure model, by initial immunosuppression. Positive frailties are associated with increased risk of graft failure.

Table 4: Summary of Frailty Effects on Graft Failure by Initial Drug Therapy for the HCV Positive Cohort

Initial treatment	No drug failure: Mean (min, max)	Hazard (min, max)	Had drug failure: Mean (min, max)	Hazard (min, max)
TAC	-1.1 (-4.3, 1.3)	0.8, 1.1	4.9 (3.4, 7.1)	1.2, 1.4
SRL + TAC	-3.9 (-6.0, -1.3)	0.8, 0.9	1.2 (0.2, 3.5)	1.0, 1.2
SRL	-2.2 (-3.9, 0.1)	0.8, 1.0	2.7 (1.7, 4.5)	1.1, 1.2
CYA	-2.2 (-4.6, 0.5)	0.8, 1.0	3.0 (1.9, 4.9)	1.1, 1.3
SRL + CYA	-4.5 (-5.7, -2.7)	0.8, 0.9	0.6 (-0.02, 2.0)	1.0, 1.1

Table 5: Posterior Means for the Log Hazard, Hazard, Standard Error and Quantiles from the HCV Negative Cohort Joint Model

Parameter	Mean	Hazard	Std err	CI (lower)	CI (upper)
Longitudinal sub-model:					
intercept	-0.128		0.019	-0.165	-0.091
slope	0.005		0.0002	0.005	0.005
age in decades	0.007		0.0004	0.006	0.008
fulminant failure	0.104		0.018	0.068	0.140
gender (F)	-0.154		0.009	-0.173	-0.137
diabetes	0.052		0.016	0.021	0.084
donor BMI	0.0002		0.00007	0.0001	0.0003
baseline trt: TAC alone (ref)	---		---	---	---
SRL + TAC	-0.041		0.018	-0.078	-0.006
SRL	-0.005		0.036	-0.074	0.063 (ns)
CYA	0.015		0.014	-0.011	0.104 (ns)
CYA + SRL	0.014		0.045	-0.073	0.104 (ns)
ν_1(variance of b_{0i})	0.075			0.071	0.079
Time-to-drug-failure sub-model:					
intercept	-9.949		0.081	-10.110	-9.792
baseline trt: TAC alone (ref)	---		---	---	---
SRL + TAC	7.249		0.336	6.594	7.916
SRL	4.112		0.565	3.029	7.916
CYA	2.643		0.220	2.206	3.075
CYA + SRL	5.844		0.735	4.437	7.356
σ_c (variance of c_i)	15.152			14.085	16.129
α(shape parameter)	0.848		0.029	0.789	0.906
Time-to-graft-failure sub-model:					
intercept	-8.134		0.226	-8.584	-7.683
HCC	0.527	1.694	0.135	0.253	0.790
donor age (decades)	0.010	1.011	0.002	0.006	0.014
gender	-0.212	0.809	0.077	-0.355	-0.060
baseline trt: TAC (ref)	--		--	--	--
SRL + TAC	0.233		0.145	-0.067	0.516 (ns)
SRL	-0.199		0.321	-0.836	0.380 (ns)
CYA	0.167		0.112	-0.059	0.388 (ns)
CYA + SRL	-0.707		0.516	-1.777	0.211 (ns)
$\rho_{\square}$	0.632	1.881	0.153	0.337	0.929
ϕ	0.064	1.066	0.012	0.039	0.088

linkage term between the longitudinal and the graft failure sub-models (but again, not in the time-to-drug-failure sub-model). We saw a lower log(creatinine) over time in female subjects and in subjects on SRL + TAC, while a higher log(creatinine) was seen in older subjects, those in fulminant failure at time of transplant, those who were diabetic at transplant, and those with a higher donor BMI. This last result, while the effect is small, is nevertheless interesting and could be investigated further. It is possible that a donor with a higher BMI has more fatty tissue in the liver, resulting in more difficult postoperative recovery and higher creatinine. Fatty liver can also lead to more inflammation and therefore higher doses of immunosuppression. It is also interesting that this effect was not seen in the HCV positive cohort, possibly because HCV positive subjects are often counselled against accepting a marginal donor organ with increased risk due to fatty liver.

Similar to the HCV positive cohort, the only covariates significant in the time-to-drug-failuresub-model were baseline treatment. Being on any initial treatment other than tacrolimus significantly shortens the time to failure of initial immunosuppressive therapy. In the graft survival component, no immunosuppressive treatment was significantly different from the reference drug TAC in terms of impact on graft survival. Risk of

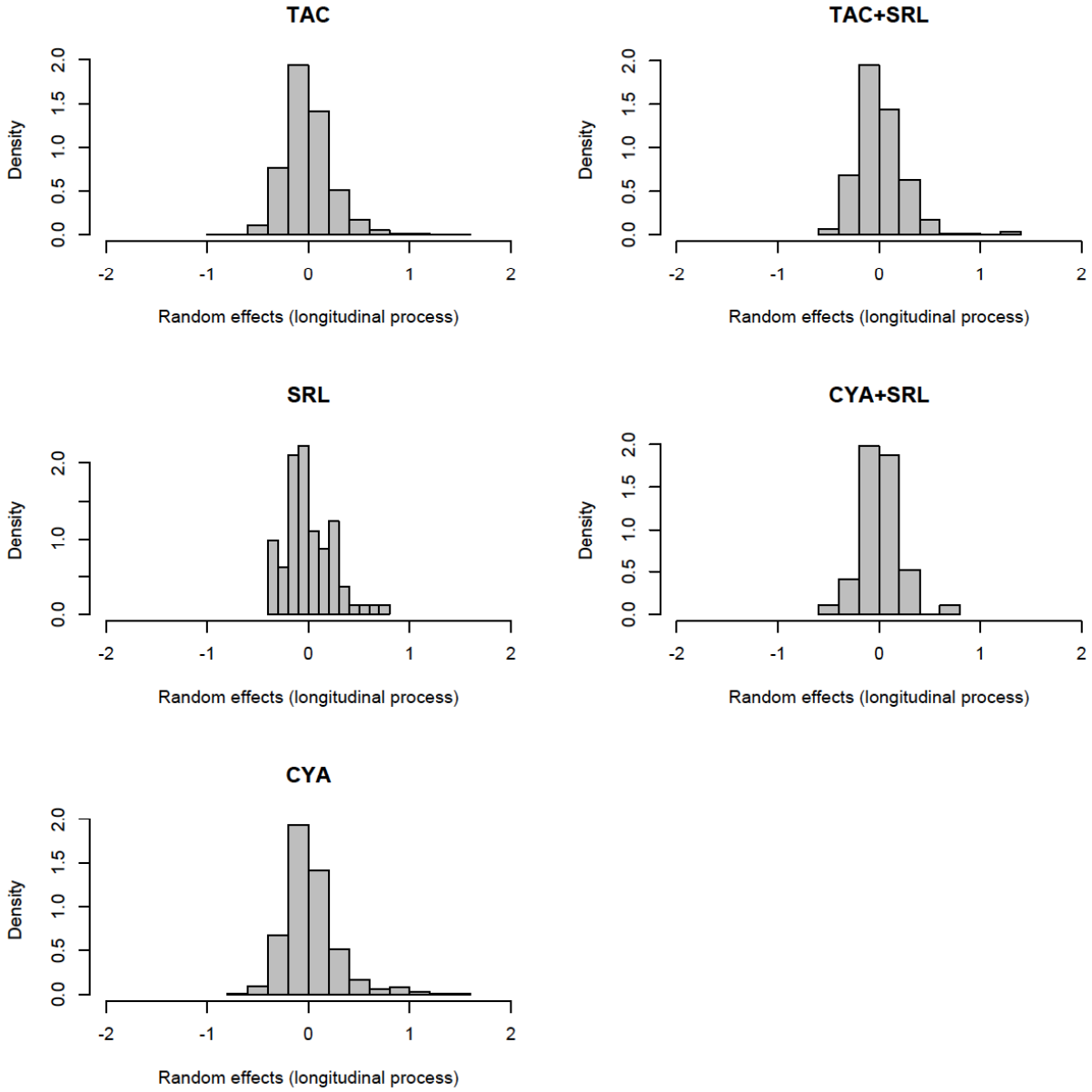

Figure 6: HCV negative cohort: Histograms of subject-specific intercept random effects from the longitudinal model, by initial treatment. Positive random effects are associated with increased risk of graft failure.

graft failure is increased for those who have HCC, for those who are African American, and for increased donor age. Gender was kept in the model because it improved DIC. We also tested all variables for interaction with HCC status and did not find any significant interaction.

Another interesting feature of this analysis is the strong effect of the longitudinal random effect for log(creatinine) in the graft survival sub-model. A one unit increase in log(creatinine) has a hazard ratio of 1.881 (compared to 1.567 in the HCV positive cohort). The subject-specific randomeffect ranges from -0.971 to 1.469, so the hazard can be reduced for negative random effects (hazardratio 0.541) or increased to as much as 2.530 for the largest random effect. It is possible that this HCV negative cohort, having a wider variety of indications for liver transplant compared to the more homogeneous HCV positive cohort, comprise a group with diseases that involve more renal decompensation. When grouped by drug (see Figure **6**), the intercept random effects are similar to the HCV positive cohort, i.e. no distinct pattern by drug (compare to Figure **3**).

The individual frailty c_i from the time-to-drug-failure component also has a significant linkage through the association parameter φ in the time-to-graft-failure model for the HCV negative cohort, where it increases survival time for those with negative frailties, and decreases survival time for those with positive frailties. This result is slightly larger than in the HCV positive cohort, with a higher risk seen in the negative cohort. Histograms of individual log frailties grouped by whether a drug change occurred are shown in Figure **7**, and plots grouped by initial treatment are shown in Figure **8**. The value of the frailty also had a wider range than the HCV positive cohort (-6.5 to +9). Figure **8** shows the wide range of individual frailties for all drug therapies. The variance of the subject-specific frailty was similar to the positive cohort, showing there is a great deal of unobserved heterogeneity between subjects.

Table **6** shows the frailty effects for each treatment combination, comparing the frailties of subjects who did not change therapy to those who did in the HCV negative cohort. Again we seethe greatest mean frailty in those who had baseline drug failure on TAC. The overall picture emerging here is that subjects who must discontinue baseline immunosuppressive therapy indeed have a poorer outcome, as expected, due to events precipitating drug failure. The calcineurin-sparing combinations of SRL + CYA and SRL + TAC stand out (in both cohorts) for having the largest negative mean frailty in those who did not experience drug failure, and the smallest positive mean frailty in those who did have a drug failure event. This translates to a lower risk of graft failure regardless of failures in treatment. The effect of initial drug regime on graft survival must be considered in the context of the time-to-drug-failure, in order to understand the overall risk to

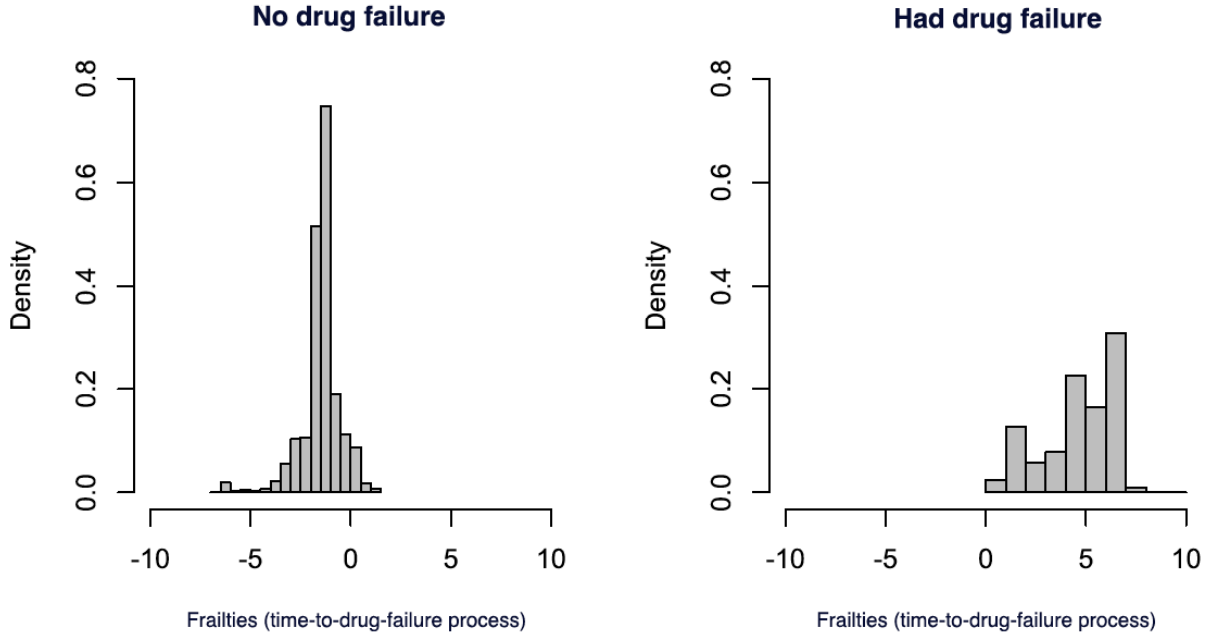

Figure 7: HCV negative cohort: Histograms of individual log frailties from time-to-drug-failure model, by whether drug failure occurred. Positive frailties are associated with increased risk of graft failure.

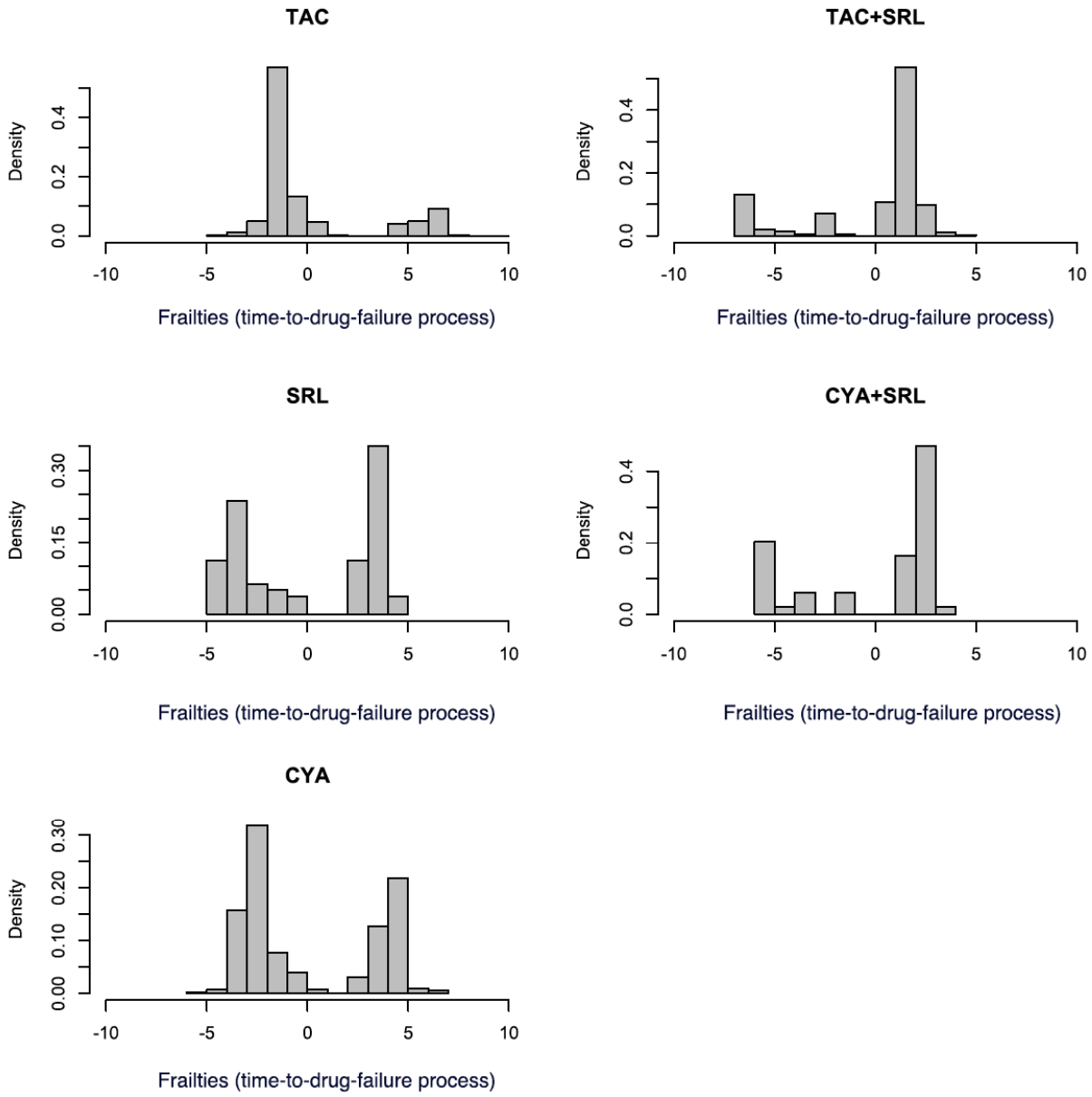

Figure 8: HCV negative cohort: Histograms of individual log frailties from time-to-drug-failure model, by initial immunosuppression. Positive frailties are associated with increased risk of graft failure.

Table 6: Summary of Frailty Effects on Graft Survival by Initial Drug Therapy for the HCV Negative Cohort

Initial trt	No drug failure: Mean (min, max)	Hazard (min, max)	Had drug failure: Mean (min, max)	Hazard (min, max)
TAC	-1.3 (-5.0, 1.4)	0.7, 1.1	5.7 (4.4, 9.1)	1.3, 1.8
SRL + TAC	-4.8 (-6.5, -1.9)	0.7, 0.9	1.6 (0.6, 4.3)	1.0, 1.3
SRL	-3.3 (-4.9, -0.7)	0.7, 1.0	3.3 (2.0, 4.7)	1.1, 1.4
CYA	-2.6 (-5.4, 0.5)	0.7, 1.0	4.0 (2.7, 6.8)	1.2, 1.5
SRL + CYA	-4.3 (-5.5, -1.3)	0.7, 0.9	2.3 (1.3, 3.4)	1.1, 1.2

each subject. Stable subjects who do not need to change baseline treatment have the best outcome, regardless of initial regime. For those who do have to switch treatments, those at highest risk seem to be the

Table 7: Results from a Standard Cox Proportional Hazards Model for the HCV Positive Cohort

Parameter	Coefficient	Hazard	Std Error	p
donor age (decades)	0.016	1.016	0.002	< 0.001
log(creatinine) day 0	0.220	1.246	0.070	0.002
gender (F)	0.141	1.152	0.074	0.057
previous malignancy	0.385	1.470	0.111	< 0.001
HCC	0.197	1.217	0.112	0.078
race (AA)	0.163	1.177	0.034	< 0.001

small group who must switch from TAC alone to some other treatment. This suggests an interesting new area of research into the dynamics of drug failure on TAC compared to any other treatment, exploring what risks are specific to this group.

We compared our model to a model using the often employed standard Cox proportional hazards model, with initial treatment as a baseline covariate. We also included log(creatinine) from the day of transplant as a covariate (86 were missing and so the mean was substituted for these cases). Using a stepwise procedure, the covariates that remain significant in the model are similar to both of the joint models presented above, although not identical. The significant covariates from this model for the HCV positive cohort are presented in Table **7**. There were no significant differences in graft survival by baseline treatment for the HCV positive cohort. The hazards for the significant covariates are very similar to the joint model, and the standard errors are almost the same. The Cox proportional hazards model has a smaller effect for log(creatinine) compared to the random effects in the joint model, and the values for log(creatinine) are smaller (range: -2.3 to 2.9) than the values for the random effects and so this translates to a smaller effect overall compared to the intercept random effect for creatinine in the joint model. The joint model also has the added hazard from covariate φ and its association with the frailty c_i in the time-to-drug-failure model.

In addition, we compared our model to a Cox proportional hazards model that treats the use of SRL as a time-dependent covariate. Significant findings are shown in Table **8**. We treated SRL as a time-dependent indicator variable where it takes a value of 1 if the subject is started on SRL (incombination or alone) at transplant, and changes to 0 when the subject changed treatment. Here we see that the effect for SRL is very significant (p = 0.029) and the hazard is 1.262 for a subject starting on SRL, and this is a result that we did not see in either of the previous models. This shows the biased results that are obtained when a time-dependent covariate is used in a Cox proportional hazards model when a change in the time-dependent covariate is also related to outcome. Some of this bias could stem from the dramatic differences in treatment switching. Subjects whos witch treatment early are those with adverse events who therefore experience more graft failure. Results from other covariates such as donor age, log(creatinine), gender and previous malignancy are similar to the previously illustrated models.

Table 8: Results from a Cox Proportional Hazards Model with Time-Dependent SRL Treatment Covariate for the HCV Positive Cohort

Parameter	Coefficient	Hazard	Std Error	p
any SRL (time-dependent)	0.233	1.262	0.107	0.029
donor age (decades)	0.010	1.016	0.002	< 0.001
log(creatinine) day 0	0.155	1.168	0.060	0.010
gender (female)	0.126	1.134	0.064	0.048
race (African American)	0.164	1.178	0.029	< 0.001
any CYA at baseline	0.634	1.885	0.279	0.023
previous malignancy	0.388	1.474	0.094	<0.001

Table 9: Results from a Standard Cox Proportional Hazards Model for the HCV Negative Cohort

Parameter	Coefficient	Hazard	Std Error	p
donor age (decades)	0.009	1.009	0.002	< 0.001
log(creatinine) day 0	0.158	1.171	0.065	0.015
gender (F)	-0.172	0.842	0.075	0.022
race (AA)	0.128	1.137	0.042	0.002
previous malignancy	0.724	2.063	0.133	< 0.001

The comparisons to the Cox proportional hazards model for the HCV negative cohort show similar findings. In the HCV negative cohort, when we apply the standard Cox proportional hazards model to graft survival we get the significant results seen in Table **9**. Again, baseline treatment is not significant factor for graft survival in any combination. The effects for donor age and gender are similar to the joint model.

In the Cox model with time dependent treatment effect for SRL, we again see a greatly increased risk for any SRL exposure, with significant results shown in Table **10**. This analysis shows that using two time-to-event processes to analyze drug failure and graft failure, along with a longitudinal component in creatinine, is a valuable approach. The model has captured the increased risk to graft failure that is present with a sharp change in the biomarker, or with an adverse event that precipitates drug failure.

5. DISCUSSION

Analyzing treatment changes as a time-to-event process is a preferred approach in observational data analysis since it avoids discarding data. It also avoids the violation of model assumptions such as when treating drug as a time-dependent covariate. Analysis of the two time-to-event outcomes of drug failure and graft survival using a joint model can account for dependence between the two processes without making strong assumptions. We allow the important information contained in the time-to-drug-failure component to influence the hazard of the time-to-graft failure component. Transplant registry data, with 100% enrolment and follow up until death, is a valuable and readily available data source that can provide insight into factors affecting health outcomes after liver transplant. The time-to-drug-failure process acts as a surrogate for time to any adverse event such as infection, rejection, cancer occurrence or other, that may be unreliably collected in the registry data. This two outcome joint model describes the data structure well. We believe that use of a straightforward joint survival outcome model is appropriate for two reasons: first, clinical practice tells us that the association is strong between the drug failure and graft failure processes, and second, less complex joint modelling techniques are more likely to be adopted in practice in

Table 10: Results from a Cox Proportional Hazards Model with Time-Dependent SRL Treatment Covariate for the HCV Negative Cohort

Parameter	Coefficient	Hazard	Std Error	p
any SRL (time-dependent)	0.404	1.498	0.095	< 0.001
donor age (decades)	0.008	1.008	0.002	< 0.001
log(creatinine) day 0	0.166	1.180	0.055	0.003
gender (F)	-0.177	0.838	0.064	0.006
recipient age (decades)	0.005	1.005	0.003	0.057
HCC	0.226	1.253	0.119	0.059
previous malignancy	0.758	2.133	0.112	< 0.001
race (AA)	0.102	1.108	0.036	0.005
blood type AB	-0.309	0.734	0.149	0.038
diabetic	0.239	1.270	0.099	0.016
CYA at baseline	0.276	1.317	0.087	0.002

this field. Joint modelling is not commonplace in analysis of SRTR data for transplantation, yet despite readily available and appropriate data, it has only been used in one application of which we are aware [4].

We did not find any significant differences between baseline treatments when compared to the 'gold standard' treatment TAC. However in the context of our joint outcome model, treatment effects should be interpreted on a subject-specific basis conditional on the frailty. We found great variability in the frailties. This is an indication that treatments alone are not good predictors of survival - generally, one treatment does provide better on average survival than another, yet there is great variability between treatments. When we condition on the frailty, we will have a better prediction of survival and better understanding of survival by treatment patterns. Rather than discard information, we have presented a model that includes all available data in a way that makes the most scientific sense. The longitudinal component takes into account factors affecting creatinine level over time. It is a proxy for choice of initial treatment, mimicking the physician decision process by taking into account all covariates affecting the evolution of creatinine over time. Degree of renal impairment is an important factor due to nephrotoxicity of the immunosuppression treatment. Our model offers more insight into the medical process and makes scientific sense. The subject-specific frailties from the time-to-drug-failure model account for latent variables that have a significant effect on graft survival, thus accounting for more of the variance in the data and enhancing our understanding of the whole process while improving the fit of the model. The information contained in the frailty covariate φ adds valuable dimension to the model that cannot be accounted for in traditional models.

A limitation of this data set is the lack of exact dates for treatment failures, and the use of the interval midpoint is not ideal, however randomized clinical trial data (with exact dates) is not readily available for public analysis. Statistical bias related to the use of the midpoint is discussed in [23, 24]. To add recurrent events to the drug failure model would involve the use of doubly interval censored data that are also correlated. Furthermore, data on subjects who experience treatment failure more than once during a data collection interval is not collected with sufficient granularity in the SRTR to allow for a joint outcome analysis with recurrent events in drug failure. An application of joint modelling techniques to detailed randomized clinical trial data would provide greater understanding of the problem, since clinical trials normally collect exact dates. With collection of exact dates, many other time-to-event processes could be examined for their association with graft survival, such as time to post-transplant diabetes, or time to cancer occurrence or recurrence.

Another limitation is that important information such as treatment dosage amounts and drug trough levels are missing from the registry data. Trough levels measure the amount of drug exposure per patient, which can vary on a subject-specific basis even when subjects are given the same amount of drug. This is another interesting avenue worth pursuing with transplant data in joint frailty models. The authors acknowledge the limitations of this three year retrospective analysis using registry data. All models in this paper would benefit from more data in the SRL arms. In this analysis we suspect that the use of only three years of data in the longitudinal component is not enough to see a significant effect from the random effect from the ζ_k in the time-to-drug-failure component. However our goal was met which showed that joint modelling provides added value to analysis of survival after liver transplantation.

The data reported here have been supplied by the Minneapolis Medical Research Foundation (MMRF) as the contractor for the Scientific Registry of Transplant Recipients (SRTR). The interpretation and reporting of these data are the responsibility of the author(s) and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government.

REFERENCES

[1] Hadley S, Samore MH, Lewis WD, Jenkins RL, Karchmer AW, Hammer SM. Major infectious complications after orthotopic liver transplantation and comparison of outcomes in patients receiving cyclosporine or FK506 as primary immunosuppression. Transplantation 1995; 59(6): 851-858. http://dx.doi.org/10.1097/00007890-199503000-00009

[2] Watt KD, Dierkhising R, Heimbach JK, Charlton MR. Impact of sirolimus and tacrolimus on mortality and graft loss in liver transplant recipients with or without hepatitis C virus: an analysis of the Scientific Registry of Transplant Recipients database. Liver Transplantation 2012; 18(9): 1029-1036. http://dx.doi.org/10.1002/lt.23479

[3] Fisher LD, Lin DY. Time-dependent covariates in the cox proportional-hazards regression model. Annual Review of Public Health 1999; 20(1): 145-157. http://dx.doi.org/10.1146/annurev.publhealth.20.1.145

[4] Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. Biometrics 2004; 60(3): 747-756. http://dx.doi.org/10.1111/j.0006-341X.2004.00225.x

[5] Chi Y, Ibrahim JG. Joint models for multivariate longitudinal and multivariate survival data. Biometrics 2006; 62(2): 432-445.

[6] Rondeau V, Mathoulin-Pelissier S, Jacqmin-Gadda H, Brute V, Soubeyran P. Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. Biostatistics 2007; 8(4): 708-721. http://dx.doi.org/10.1093/biostatistics/kxl043

[7] Rondeau V, Michiels S, Liquet B, Pignon J-P. Investigating trial and treatment heterogeneity in an individual patient data meta-analysis of survival data by means of the penalized maximum likelihood approach. Statistics in Medicine 2008; 27(11): 1894-1910. http://dx.doi.org/10.1002/sim.3161

[8] Liu L, Huang X. Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. Journal of the Royal Statistical Society: Series C (Applied Statistics) 2009; 58(1): 65-81. http://dx.doi.org/10.1111/j.1467-9876.2008.00641.x

[9] Musoro JZ, Geskus RB, Zwinderman AH. A joint model for repeated events of different types and multiple longitudinal outcomes with application to a follow-up study of patients after kidney transplant. Biometrical Journal 2015; 57(2): 185-200. http://dx.doi.org/10.1002/bimj.201300167

[10] Elashoff RM, Li, G, Li N. An approach to joint analysis of longitudinal measurements and competing risks failure time data. Statistics in Medicine 2007; 26(14): 2813. http://dx.doi.org/10.1002/sim.2749

[11] Elashoff RM, Li, G, Li N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. Biometrics 2007; 64(3): 762-771. http://dx.doi.org/10.1111/j.1541-0420.2007.00952.x

[12] Williamson PR, Kolamunnage-Dona R, Philipson P, Marson AG. Joint modelling of longitudinal and competing risks data. Statistics in Medicine 2008; 27(30): 6426-6438. http://dx.doi.org/10.1002/sim.3451

[13] Xu J, Kalbfleisch JD, Tai B. Statistical analysis of illness-death processes and semicompeting risks data. Biometrics 2010; 66(3): 716-725. http://dx.doi.org/10.1111/j.1541-0420.2009.01340.x

[14] Han B, Yu M, Dignam JJ, Rathouz PJ. Bayesian approach for flexible modeling of semicompeting risks data. Statistics in Medicine 2014; 33(29): 5111-5125. http://dx.doi.org/10.1002/sim.6313

[15] Rouanet A, Joly P, Dirties JF, Proust-Lima C, Jacqmin-Gadda H. Joint latent class model for longitudinal data and interval-censored semi-competing events: Application to dementia. Biometrics 2015; 59: 1-23.

[16] Dickinson DM, Ellison MD, Webb RL. Data sources and structure. American Journal of Transplantation 2003; 3(s4): 13-28. http://dx.doi.org/10.1034/j.1600-6143.3.s4.3.x

[17] R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2010.

[18] Plummer M *et al*. Jags: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd international workshop on distributed statistical computing. 2003; 124: 125-132.

[19] Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 1998; 7(4): 434-455.

[20] Toso C, Merani S, Bigam DL, Shapiro AM, Kneteman NM. Sirolimus-based immunosuppression is associated with increased survival after liver transplantation for hepatocellular carcinoma. Hepatology 2010; 51(4): 1237-1243. http://dx.doi.org/10.1002/hep.23437

[21] McKenna GJ, Trotter JF, Klintmalm E, Onaca N, Ruiz R, Jennings LW, Neri N, O'Leary JG, Davis GL, Levy MF, *et al*. Limiting hepatitis C virus progression in liver transplant recipients using sirolimus-based immunosuppression. American Journal of Transplantation 2011; 11(11): 2379-2387. http://dx.doi.org/10.1111/j.1600-6143.2011.03767.x

[22] Howell J, Angus P, Gow P. Hepatitis C recurrence: the Achilles heel of liver transplantation. Transplant Infectious Disease 2014; 16(1): 1-16. http://dx.doi.org/10.1111/tid.12173

[23] Odell PM, Anderson KM, and Ralph B D'Agostino. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. Biometrics 1992; 951-959. http://dx.doi.org/10.2307/2532360

[24] Rucker G, Messerer D. Remission duration: An example of interval-censored observations. Statistics in Medicine 1988; 7(11): 1139-1145. http://dx.doi.org/10.1002/sim.4780071106

Received on 13-07-2016 Accepted on 29-07-2016 Published on 16-08-2016

http://dx.doi.org/10.6000/1929-6029.2016.05.03.6