Supplementary Materials for "Time-to-event analysis when the event is defined on a finite time interval"

Catherine Lee
Division of Research,
Kaiser Permanente, Oakland, California, U.S.A.
catherine.lee@kp.org

Stephanie J. Lee
Clinical Research Division,
Fred Hutchinson Cancer Research Center, Seattle, WA

Sebastien Haneuse
Department of Biostatistics,
Harvard T. H. Chan School of Public Health, Boston, Massachusetts, U.S.A.

Introduction

We present additional results that are beyond what could be presented in the main manuscript. This document is organized as follows:

Appendix A: Description of the CIBMTR HCT data referenced in Section 2 of the main manuscript

Appendix B: Components of marginal log-likelihood function referenced in Section 4.2 of the main manuscript

Appendix C: Expression for joint density of (T_1, T_2) referenced in Section 3.4 of the main manuscript

Appendix D: Equivalence between standard transition probability expressions from the multi-state model literature and absolute risk profile contributions calculated via integration over the joint density, referenced in Section 3.5 of the main manuscript

Appendix E: Simulation results corresponding to the 'baseline' scenario enumerated in Section 5.1 and additional data scenarios described in Section 5.2 of the main manuscript

E.1 baseline scenario

E.2 reducing sample size from n=5,000 to n=1,000

E.3 administrative censoring via a censoring variable simulated from an Exponential distribution with mean 100

E.4 administrative censoring solely of the terminal event at 365 days

E.5 increasing the non-susceptibility fraction from 46% to 76%

E.6 modification of α_{23} and κ_{23} so that the hazard functions for the truncated and standard Weibull distributions are comparable on the interval $(0, \tau)$

E.7 baseline scenario without frailty term, i.e. Gamma frailty variance equal to zero

Appendix F: Additional results from the analysis of stem cell transplantation data in Section 6 of the main manuscript

Appendix G: Details regarding the included `R` code

A Description of CIBMTR HCT data

Table A.1: Description of baseline characteristics of study population comprised of patients who underwent first HCT between 1999-2011 and distribution of observed outcomes (acute GHVD and/or death) by the end of 365 days post-transplantation.

			Observed outcome category, %			
	N	%	Both acute GVHD & death	Acute GVHD & censored for death	Death without acute GVHD	Censored for both
Total subjects	9651		11.0	6.7	28.7	53.7
Gender						
Male	5366	55.6	11.7	6.9	28.3	53.1
Female	4285	44.4	10.1	6.4	29.2	54.4
Age, years						
<10	653	6.8	5.7	7.0	21.6	65.7
10-19	1162	12.0	8.5	8.3	21.8	61.4
20-29	1572	16.3	11.3	8.1	25.9	54.7
30-39	1581	16.4	11.6	6.3	26.5	55.6
40-49	2095	21.7	11.4	6.8	29.6	52.2
50-59	2008	20.8	12.6	5.6	35.2	46.5
60+	580	6.0	11.9	3.1	38.1	46.9
Disease type						
AML	4919	51.0	10.1	6.1	31.8	52.0
ALL	2071	21.5	10.4	6.6	29.3	53.7
CML	1525	15.8	12.9	8.3	18.0	60.9
MDS	1136	11.8	13.1	7.1	28.5	51.2
Disease status						
Early	4873	50.5	9.0	7.3	21.1	62.5
Intermediate	2316	24.0	11.2	7.0	28.4	53.5
Advanced	2462	25.5	14.6	5.1	43.9	36.4
HLA compatibility						
Identical sibling	3941	40.8	7.9	5.7	26.0	60.4
8/8	4100	42.5	11.8	7.3	29.7	51.1
7/8	1610	16.7	16.3	7.5	32.7	43.5

B Likelihood contributions and score equations

B.1 Marginal likelihood contributions

Let $\eta_\omega = X_\omega \beta_\omega$ and $\Lambda_{\omega,0}(y) = \int_0^y \lambda_{\omega,0}(s)\ ds$, for $\omega \in \{14, 24, 23, 34\}$.

$$\begin{aligned}
f_1(y_1, y_2; X) &= \int_0^\infty f_1(y_1, y_2; X, \gamma) f_\gamma(\gamma)\ d\gamma \\
&= \int_0^\infty \Big[(1-\pi)\cdot\gamma\cdot\lambda_{23,0}(y_1)e^{\eta_{23}}\cdot\exp\left\{-\gamma\left(\Lambda_{23,0}(y_1)e^{\eta_{23}} + \Lambda_{24,0}(y_1)e^{\eta_{24}}\right)\right\}\cdot \\
&\quad \gamma\cdot\lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\cdot\exp\left\{-\gamma\Lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\right\}\cdot \\
&\quad \left[\Gamma\left(1/\theta\right)\right]^{-1}\theta^{-\frac{1}{\theta}}\gamma^{\frac{1}{\theta}-1}\exp\left\{-\gamma/\theta\right\}\Big]\ d\gamma \\
&= \frac{(1-\pi)\cdot\lambda_{23,0}(y_1)e^{\eta_{23}}\cdot\lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\cdot(1+\theta)}{\left[1+\theta\left\{\Lambda_{23,0}(y_1)e^{\eta_{23}} + \Lambda_{24,0}(y_1)e^{\eta_{24}} + \Lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\right\}\right]^{1/\theta+2}}
\end{aligned}$$

$$\begin{aligned}
f_2(y_1, y_2; X) &= \int_0^\infty f_2(y_1, y_2; X, \gamma) f_\gamma(\gamma)\ d\gamma \\
&= \int_0^\infty \Big[(1-\pi)\cdot\gamma\cdot\lambda_{23,0}(y_1)e^{\eta_{23}}\cdot\exp\left\{-\gamma\left(\Lambda_{23,0}(y_1)e^{\eta_{23}} + \Lambda_{24,0}(y_1)e^{\eta_{24}}\right)\right\}\cdot \\
&\quad \exp\left\{-\gamma\Lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\right\}\cdot\left[\Gamma\left(1/\theta\right)\right]^{-1}\theta^{-\frac{1}{\theta}}\gamma^{\frac{1}{\theta}-1}\exp\left\{-\gamma/\theta\right\}\Big]\ d\gamma \\
&= (1-\pi)\cdot\lambda_{23,0}(y_1)e^{\eta_{23}}\cdot \\
&\quad \left[1+\theta\left\{\Lambda_{23,0}(y_1)e^{\eta_{23}} + \Lambda_{24,0}(y_1)e^{\eta_{24}} + \Lambda_{34,0}(y_2-y_1)e^{\eta_{34}}\right\}\right]^{-1/\theta-1}
\end{aligned}$$

$$\begin{aligned}
f_3(y_1, y_2; X) &= \int_0^\infty f_3(y_1, y_2; X, \gamma) f_\gamma(\gamma)\ d\gamma \\
&= \int_0^\infty \Big[\pi\cdot\gamma\cdot\lambda_{14,0}(y_2)e^{\eta_{14}}\cdot\exp\left\{-\gamma\Lambda_{14,0}(y_2)e^{\eta_{14}}\right\}\cdot \\
&\quad \left[\Gamma\left(1/\theta\right)\right]^{-1}\theta^{-\frac{1}{\theta}}\gamma^{\frac{1}{\theta}-1}\exp\left\{-\gamma/\theta\right\}\Big]\ d\gamma \\
&= \pi\cdot\lambda_{14,0}(y_2)e^{\eta_{14}}\cdot\left[\Gamma\left(1/\theta\right)\right]^{-1}\theta^{-\frac{1}{\theta}}\cdot \\
&\quad \int_0^\infty \gamma^{1/\theta}\cdot\exp\left[-\gamma\left\{\frac{1}{\theta} + \Lambda_{14,0}(y_2)e^{\eta_{14}}\right\}\right]\ d\gamma \\
&= \pi\cdot\lambda_{14,0}(y_2)e^{\eta_{14}}\left\{1+\theta\cdot\Lambda_{14,0}(y_2)e^{\eta_{14}}\right\}^{-1/\theta-1}
\end{aligned}$$

$$\begin{aligned}
f_4(y_1, y_2; X) &= \int_0^\infty f_4(y_1, y_2; X, \gamma) f_\gamma(\gamma)\ d\gamma \\
&= \int_0^\infty \Big[\pi\cdot\gamma\cdot\lambda_{14,0}(y_2)e^{\eta_{14}}\cdot\exp\left\{-\gamma\Lambda_{14,0}(y_2)e^{\eta_{14}}\right\} + \\
&\quad (1-\pi)\cdot\gamma\cdot\lambda_{24,0}(y_2)e^{\eta_{24}}\cdot\exp\left\{-\gamma\left(\Lambda_{23,0}(y_2)e^{\eta_{23}} + \Lambda_{24,0}(y_2)e^{\eta_{24}}\right)\right\}\Big]\cdot f_\gamma(\gamma)\ d\gamma \\
&= \int_0^\infty \Big[\pi\cdot\gamma\cdot\lambda_{14,0}(y_2)e^{\eta_{14}}\cdot\exp\left\{-\gamma\Lambda_{14,0}(y_2)e^{\eta_{14}}\right\} + \\
&\quad (1-\pi)\cdot\gamma\cdot\lambda_{24,0}(y_2)e^{\eta_{24}}\cdot\exp\left\{-\gamma\left(\Lambda_{23,0}(y_2)e^{\eta_{23}} + \Lambda_{24,0}(y_2)e^{\eta_{24}}\right)\right\}\Big]\cdot \\
&\quad \left[\Gamma\left(1/\theta\right)\right]^{-1}\theta^{-\frac{1}{\theta}}\gamma^{\frac{1}{\theta}-1}\exp\left\{-\gamma/\theta\right\}\ d\gamma
\end{aligned}$$

$$
\begin{aligned}
&= \pi \cdot \lambda_{14,0}(y_2) e^{\eta_{14}} \left\{1 + \theta \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\}^{-1/\theta - 1} + \\
&\quad (1-\pi) \cdot \lambda_{24,0}(y_2) e^{\eta_{24}} \left[1 + \theta \left\{\Lambda_{23,0}(y_2) e^{\eta_{23}} + \Lambda_{24,0}(y_2) e^{\eta_{24}}\right\}\right]^{-1/\theta - 1}
\end{aligned}
$$

$$
\begin{aligned}
f_5(y_1, y_2; X) &= \int_0^\infty f_5(y_1, y_2; X, \gamma) \cdot f_\gamma(\gamma) \ d\gamma \\
&= \int_0^\infty \pi \exp\left\{-\gamma \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\} \cdot \left[\Gamma\left(1/\theta\right)\right]^{-1} \theta^{-\frac{1}{\theta}} \gamma^{\frac{1}{\theta} - 1} \exp\left\{-\gamma/\theta\right\} \ d\gamma \\
&= \pi \left[\Gamma\left(1/\theta\right)\right]^{-1} \theta^{-\frac{1}{\theta}} \int_0^\infty \gamma^{\frac{1}{\theta} - 1} \exp\left[-\gamma \left\{\frac{1}{\theta} + \Lambda_{14,0}(y_2) e^{\eta_{14}}(t)\right\}\right] \ d\gamma \\
&= \pi \left\{1 + \theta \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\}^{-1/\theta}
\end{aligned}
$$

$$
\begin{aligned}
f_6(y_1, y_2; X) &= \int_0^\infty f_6(y_1, y_2; X, \gamma) f_\gamma(\gamma) \ d\gamma \\
&= \int_0^\infty \left[\pi \exp\left\{-\gamma \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\} + \right. \\
&\quad \left. (1-\pi) \cdot \exp\left\{-\gamma \left(\Lambda_{23,0}(y_2) e^{\eta_{23}} + \Lambda_{24,0}(y_2) e^{\eta_{24}}\right)\right\}\right] \cdot f_\gamma(\gamma) \ d\gamma \\
&= \int_0^\infty \left[\pi \exp\left\{-\gamma \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\} + \right. \\
&\quad \left. (1-\pi) \cdot \exp\left\{-\gamma \left(\Lambda_{23,0}(y_2) e^{\eta_{23}} + \Lambda_{24,0}(y_2) e^{\eta_{24}}\right)\right\}\right] \cdot \\
&\quad \left[\Gamma\left(1/\theta\right)\right]^{-1} \theta^{-\frac{1}{\theta}} \gamma^{\frac{1}{\theta} - 1} \exp\left\{-\gamma/\theta\right\} \ d\gamma \\
&= \pi \left\{1 + \theta \cdot \Lambda_{14,0}(y_2) e^{\eta_{14}}\right\}^{-1/\theta} + (1-\pi) \left[1 + \theta \left\{\Lambda_{23,0}(y_2) e^{\eta_{23}} + \Lambda_{24,0}(y_2) e^{\eta_{24}}\right\}\right]^{-1/\theta}
\end{aligned}
$$

C Joint density

As in the illness-death model with shared frailty, the transition hazards (6)-(9) determine the joint distribution of T_1 and T_2. We denote the joint density on the observable region $\mathcal{U} = (0 < T_1 < T_2) \cap (0 < T_1 < \tau)$, by $g_{\mathcal{U}}(t_1, t_2)$, for $t_1 > t_2$. We assign the remaining probability mass along the line $T_1 = \tau$, with density denoted $g_{T_1=\tau,T_2}(t_2)$, representing subjects who never experience the non-terminal event T_1. This implies that the probability mass for non-susceptible patients ($L = 1$) is distributed along the line $T_1 = \tau$. For patients who are susceptible ($L = 0$), the set of possible transitions coincides with the standard illness-death model in Xu et al. (2010) [1] and the corresponding density expressions align with the joint density, which we denote by f and can be found in the Supplementary Materials of Lee et al (2016) [2].

We derive the joint density corresponding to the model in Section 3 for the semi-Markov model, where $\Lambda_\omega(t) = \int_0^t \lambda_\omega(s)\, ds$.

$$
\begin{aligned}
g(t_1, t_2|\gamma) &= g(T_1 = t_1, T_2 = t_2|\gamma) \\
&= P(L = 1) \cdot g(t_1, t_2|L = 1, \gamma) + P(L = 0) \cdot g(t_1, t_2|L = 0, \gamma) \\
&= \pi \cdot g(t_1, t_2|L = 1, \gamma) + (1 - \pi) \cdot g(t_1, t_2|L = 0, \gamma) \\
&= \pi \cdot \lambda_{14}(t_2|\gamma) \cdot \exp\{-\Lambda_{14}(t_2)\} \cdot 1(T_1 = \tau) + (1 - \pi) \cdot f(t_1, t_2|\gamma) \\
&= \pi \cdot \lambda_{14}(t_2|\gamma) \cdot \exp\{-\Lambda_{14}(t_2)\} \cdot 1(T_1 = \tau) + \\
&\quad\; (1 - \pi) \cdot [\lambda_{23}(t_1|\gamma)\lambda_{34}(t_2 - t_1|t_1, \gamma) \cdot \\
&\qquad \exp\left\{-\Lambda_{23}(t_1|\gamma) - \Lambda_{24}(t_1|\gamma) - \int_0^{t_2 - t_1} \lambda_{34}(v|t_1, \gamma)\, dv\right\} \cdot 1(t_1 < t_2) + \\
&\qquad \lambda_{24}(t_2|\gamma) \exp\{-\Lambda_{23}(t_2|\gamma) - \Lambda_{34}(t_2|\gamma)\} \cdot 1(T_1 = \tau)] .
\end{aligned}
$$

Collecting like terms, we see that the probability density along the line $T_1 = \tau$ is given by

$$
g_{T_1=\tau,T_2}(t_2|\gamma) = \pi \cdot \lambda_{14}(t_2|\gamma) \cdot \exp\{-\Lambda_{14}(t_2)\} + (1 - \pi) \cdot \lambda_{24}(t_2|\gamma) \exp\{-\Lambda_{23}(t_2|\gamma) - \Lambda_{34}(t_2|\gamma)\}
$$

and the probability density over the region $\mathcal{U}$ is given by

$$
g_{\mathcal{U}}(t_1, t_2|\gamma) = (1 - \pi) \cdot \lambda_{23}(t_1|\gamma)\lambda_{34}(t_2 - t_1|t_1, \gamma) \exp\left\{-\Lambda_{23}(t_1|\gamma) - \Lambda_{24}(t_1|\gamma) - \int_0^{t_2 - t_1} \lambda_{34}(v|t_1, \gamma)\, dv\right\} .
$$

D Equivalence of absolute risk profiles and transition probabilities

Showing the equivalence of the transition probabilities presented in Putter et al. [3], P_{12}^0, P_{13}^2, P_{13}^1 and P_{11}, and the absolute risk profiles presented in Section 3.5, which are calculated via integration of the joint density function. Note: The expressions in Putter et al. [3] are presented for the Markov model.

Let $\Lambda_\omega(t) = \int_0^t \lambda_\omega(s)\ ds$.

$$
\begin{aligned}
p_{\text{aGVHD only}}(t|\gamma) &= \int_t^\infty \int_0^t g_{\mathcal{U}}(u, v|\gamma)\ du\ dv \\
&= \int_t^\infty \int_0^t \lambda_1(u|\gamma) \cdot \lambda_3(v|\gamma) \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) - \Lambda_3(v|\gamma) + \Lambda_3(u|\gamma)\}\ du\ dv \\
&= \left[\int_0^t \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) + \Lambda_3(u|\gamma)\}\ du\right] \left[\int_t^\infty \lambda_3(v|\gamma) \exp\{-\Lambda_3(v|\gamma)\}\ dv\right] \\
&= \int_0^t \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) + \Lambda_3(u|\gamma)\} \cdot \exp\{-\Lambda_3(t|\gamma)\}\ du \\
&= P_{12}(0, t|\gamma) \\
&= \\
p_{\text{aGVHD and death}}(t|\gamma) &= \int_0^t \int_u^t g_{\mathcal{U}}(u, v|\gamma)\ dv\ du \\
&= \int_0^t \int_u^t \lambda_1(u|\gamma) \cdot \lambda_3(v|\gamma) \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) - \Lambda_3(v|\gamma) + \Lambda_3(u|\gamma)\}\ dv\ du \\
&= \int_0^t \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma)\} \cdot \left[\int_u^t \lambda_3(v|\gamma) \cdot \exp\{-\Lambda_3(v|\gamma) + \Lambda_3(u|\gamma)\}\ dv\right]\ du \\
&= P_{13}^2(0, t|\gamma) \\
&= \\
p_{\text{death only}}(t|\gamma) &= \int_0^t g_{T_1=\tau, T_2}(v|\gamma)\ dv \\
&= \int_0^t \lambda_2(v|\gamma) \cdot \exp\{-\Lambda_1(v|\gamma) - \Lambda_2(v|\gamma)\}\ dv \\
&= P_{13}^1(0, t|\gamma) \\
&= \\
p_{\text{alive without aGVHD}}(t|\gamma) &= \int_t^\infty \int_u^\infty g_{\mathcal{U}}(u, v|\gamma)\ dv\ du + \int_t^\infty g_{T_1=\infty, T_2}(v|\gamma)\ dv \\
&= \int_t^\infty \int_u^\infty \lambda_1(u|\gamma) \cdot \lambda_3(v|\gamma) \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) - \Lambda_3(v|\gamma) + \Lambda_3(u|\gamma)\}\ dv\ du + \\
&\qquad \int_t^\infty \lambda_2(v|\gamma) \cdot \exp\{-\Lambda_1(v|\gamma) - \Lambda_2(v|\gamma)\}\ dv \\
&= \int_t^\infty \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) + \Lambda_3(u|\gamma)\} \left[\int_u^\infty \lambda_3(v|\gamma) \exp\{-\Lambda_3(v|\gamma)\}\ dv\right]\ du + \\
&\qquad \int_t^\infty \lambda_2(v|\gamma) \cdot \exp\{-\Lambda_1(v|\gamma) - \Lambda_2(v|\gamma)\}\ dv \\
&= \int_t^\infty \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma) + \Lambda_3(u|\gamma)\} \exp\{-\Lambda_3(u|\gamma)\}\ du +
\end{aligned}
$$

$$
\begin{aligned}
&\int_t^\infty \lambda_2(v|\gamma) \cdot \exp\{-\Lambda_1(v|\gamma) - \Lambda_2(v|\gamma)\}\ dv \\
&= \int_t^\infty \lambda_1(u|\gamma) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma)\}\ du + \int_t^\infty \lambda_2(v|\gamma) \cdot \exp\{-\Lambda_1(v|\gamma) - \Lambda_2(v|\gamma)\}\ dv \\
&= \int_t^\infty (\lambda_1(u|\gamma) + \lambda_2(u|\gamma)) \cdot \exp\{-\Lambda_1(u|\gamma) - \Lambda_2(u|\gamma)\}\ du \\
&= \exp\{-\Lambda_1(t|\gamma) - \Lambda_2(t|\gamma)\} \\
&= P_{11}(0, t|\gamma)
\end{aligned}
$$

E Additional simulation results

E.1 Baseline scenario

Table E.1: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets generated under the 'baseline' scenario with n=5,000.

	FI-ID$^\gamma$		FI-ID		Conlon$^\gamma$		Conlon		ID$^\gamma$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction														
$\beta_{s,0}$	0.05	0.05	0.05	0.05	0.75	0.05	0.05	0.05			0.05	0.05		
$\beta_{s,x}$	0.06	0.06	0.06	0.06	0.56	0.06	0.06	0.07			0.07	0.07		
Non-susceptible: $1 \to 4$														
$\log(\kappa_{14})$	0.29	0.27	0.20	0.18	1.56	0.37	0.22	0.20						
$\log(\alpha_{14})$	0.03	0.03	0.03	0.02	0.13	0.04	0.03	0.02						
$\beta_{14,x}$	0.05	0.05	0.05	0.04	1.32	0.06	0.05	0.05						
Susceptible: $2 \to 4$														
$\log(\kappa_{24})$	0.15	0.14	0.14	0.14	0.74	0.13	0.13	0.14	0.10	0.10				
$\log(\alpha_{24})$	0.03	0.03	0.03	0.03	0.43	0.03	0.03	0.03	0.04	0.03				
$\beta_{24,x}$	0.08	0.08	0.09	0.08	0.60	0.08	0.08	0.08	0.06	0.05			0.04	0.04
Susceptible: $2 \to 3$														
$\log(\kappa_{23})$	0.12	0.11	0.15	0.14	0.76	0.05	0.05	0.05	0.04	0.04	0.05	0.05		
$\log(\alpha_{23})$	0.03	0.03	0.03	0.03	0.64	0.02	0.02	0.02	0.04	0.03	0.02	0.02		
$\beta_{23,x}$	0.06	0.06	0.06	0.06	2.90	0.06	0.06	0.06	0.06	0.05	0.06	0.06	0.05	0.05
Susceptible: $3 \to 4$														
$\log(\kappa_{34})$	0.14	0.14	0.12	0.11	0.78	0.15	0.12	0.11	0.18	0.17				
$\log(\alpha_{34})$	0.02	0.02	0.02	0.02	0.24	0.02	0.02	0.02	0.03	0.02				
$\beta_{34,x}$	0.06	0.05	0.05	0.05	1.23	0.05	0.05	0.05	0.06	0.06			0.05	0.05
log-frailty variance														
$\log(\theta)$	0.12	0.12			4.35	0.17			0.16	0.12				

Table E.2: Coverage probabilities, based on 2,000 simulated datasets generated under the 'baseline' scenario with n=5,000.

	FI-ID$^{\gamma}$	FI-ID	Conlon$^{\gamma}$	Conlon
Non-susceptible fraction				
$\beta_{s,0}$	0.94	0.44	0.46	0.95
$\beta_{s,x}$	0.95	0.95	0.91	0.90
Non-susceptible: $1 \to 4$				
$\log(\kappa_{14})$	0.95	0.00	0.72	0.00
$\log(\alpha_{14})$	0.94	0.00	0.76	0.00
$\beta_{14,x}$	0.95	0.75	0.87	0.92
Susceptible: $2 \to 4$				
$\log(\kappa_{24})$	0.94	0.85	0.93	0.67
$\log(\alpha_{24})$	0.95	0.09	0.86	0.08
$\beta_{24,x}$	0.94	0.74	0.95	0.90
Susceptible: $2 \to 3$				
$\log(\kappa_{23})$	0.95	0.92	0.00	0.00
$\log(\alpha_{23})$	0.95	0.54	0.03	0.30
$\beta_{23,x}$	0.95	0.94	0.95	0.96
Susceptible: $3 \to 4$				
$\log(\kappa_{34})$	0.95	0.00	0.95	0.00
$\log(\alpha_{34})$	0.95	0.00	0.96	0.00
$\beta_{34,x}$	0.94	0.92	0.94	0.92
log-frailty variance				
$\log(\theta)$	0.95		0.96	

E.2 Reducing sample size from n=5,000 to n=1,000

Table E.3: Mean point estimates, based on 2,000 simulated datasets generated under the 'baseline' scenario with n=1,000.

	Truth	FI-ID$^{\gamma}$	FI-ID	Conlon$^{\gamma}$	Conlon	ID$^{\gamma}$	Cure	MSM
Non-susceptible fraction								
$\beta_{s,0}$	-0.41	-0.41	-0.32	-0.53	-0.41		0.14	
$\beta_{s,x}$	0.50	0.50	0.50	0.53	0.54		0.51	
Non-susceptible: $1 \to 4$								
$\log(\kappa_{14})$	-8.52	-8.55	-7.13	-9.13	-7.49			
$\log(\alpha_{14})$	0.34	0.34	0.15	0.36	0.19			
$\beta_{14,x}$	0.25	0.25	0.19	0.29	0.23			
Susceptible: $2 \to 4$								
$\log(\kappa_{24})$	-5.30	-5.33	-5.23	-5.26	-5.12	-5.37		
$\log(\alpha_{24})$	0.34	0.34	0.23	0.30	0.25	0.09		
$\beta_{24,x}$	0.50	0.50	0.41	0.52	0.45	-0.08		0.08
Susceptible: $2 \to 3$								
$\log(\kappa_{23})$	-2.16	-2.21	-2.41	-1.80	-1.77	-1.98	-1.50	
$\log(\alpha_{23})$	-0.69	-0.69	-0.75	-0.63	-0.63	-1.10	-0.58	
$\beta_{23,x}$	0.25	0.25	0.24	0.25	0.24	-0.31	0.38	-0.26
Susceptible: $3 \to 4$								
$\log(\kappa_{34})$	-6.21	-6.25	-5.70	-6.26	-5.70	-7.04		
$\log(\alpha_{34})$	0.26	0.27	0.16	0.27	0.16	0.37		
$\beta_{34,x}$	0.15	0.15	0.14	0.16	0.14	0.11		0.15
log-frailty variance								
$\log(\theta)$	-1.71	-1.79		-2.06		-0.52		

Table E.4: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets generated under the 'baseline' scenario with n=1,000.

	FI-ID$^{\gamma}$		FI-ID		Conlon$^{\gamma}$		Conlon		ID$^{\gamma}$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction														
$\beta_{s,0}$	0.10	0.10	0.10	0.11	0.55	0.12	0.11	0.11			0.11	0.12		
$\beta_{s,x}$	0.14	0.14	0.14	0.14	0.26	0.15	0.14	0.15			0.14	0.15		
Non-susceptible: $1 \to 4$														
$\log(\kappa_{14})$	0.63	0.62	0.47	0.42	1.65	0.84	0.52	0.46						
$\log(\alpha_{14})$	0.07	0.07	0.06	0.05	1.19	0.09	0.06	0.05						
$\beta_{14,x}$	0.12	0.12	0.10	0.10	0.71	0.13	0.11	0.10						
Susceptible: $2 \to 4$														
$\log(\kappa_{24})$	0.32	0.31	0.31	0.32	0.64	0.30	0.30	0.31	0.24	0.23				
$\log(\alpha_{24})$	0.07	0.07	0.07	0.08	0.36	0.06	0.06	0.06	0.08	0.06				
$\beta_{24,x}$	0.18	0.18	0.21	0.18	0.54	0.18	0.19	0.18	0.12	0.12			0.08	0.08
Susceptible: $2 \to 3$														
$\log(\kappa_{23})$	0.31	0.30	0.68	0.85	0.58	0.11	0.10	0.11	0.08	0.10	0.12	0.12		
$\log(\alpha_{23})$	0.07	0.07	0.07	0.07	0.61	0.05	0.05	0.05	0.09	0.07	0.05	0.05		
$\beta_{23,x}$	0.13	0.13	0.13	0.13	0.63	0.13	0.12	0.13	0.12	0.12	0.14	0.15	0.10	0.11
Susceptible: $3 \to 4$														
$\log(\kappa_{34})$	0.31	0.31	0.27	0.26	1.11	0.33	0.27	0.26	0.42	0.37				
$\log(\alpha_{34})$	0.05	0.05	0.04	0.04	0.10	0.05	0.04	0.04	0.06	0.05				
$\beta_{34,x}$	0.12	0.12	0.11	0.11	0.61	0.12	0.11	0.11	0.14	0.14			0.11	0.11
log-frailty variance														
$\log(\theta)$	0.37	0.35			3.99	0.66			0.38	0.29				

Table E.5: Coverage probabilities, based on 2,000 simulated datasets generated under the 'baseline' scenario with n=1,000.

	FI-ID$^{\gamma}$	FI-ID	Conlon$^{\gamma}$	Conlon
Non-susceptible fraction				
$\beta_{s,0}$	0.95	0.87	0.87	0.96
$\beta_{s,x}$	0.96	0.97	0.95	0.95
Non-susceptible: $1 \to 4$				
$\log(\kappa_{14})$	0.94	0.12	0.93	0.40
$\log(\alpha_{14})$	0.94	0.08	0.92	0.27
$\beta_{14,x}$	0.94	0.89	0.94	0.93
Susceptible: $2 \to 4$				
$\log(\kappa_{24})$	0.95	0.94	0.95	0.90
$\log(\alpha_{24})$	0.96	0.75	0.95	0.71
$\beta_{24,x}$	0.95	0.89	0.95	0.94
Susceptible: $2 \to 3$				
$\log(\kappa_{23})$	0.95	0.98	0.10	0.05
$\log(\alpha_{23})$	0.96	0.92	0.57	0.78
$\beta_{23,x}$	0.96	0.94	0.95	0.96
Susceptible: $3 \to 4$				
$\log(\kappa_{34})$	0.95	0.47	0.95	0.47
$\log(\alpha_{34})$	0.95	0.30	0.95	0.30
$\beta_{34,x}$	0.95	0.93	0.95	0.93
log-frailty variance				
$\log(\theta)$	0.97		0.97	

E.3 Administrative censoring via a censoring variable simulated from an Exponential distribution with mean 100

Table E.6: Mean point estimates, based on 2,000 simulated datasets of size n=5,000 generated with administrative censoring via a censoring variable simulated from an Exponential distribution with mean 100.

	Truth	FI-ID$^\gamma$	FI-ID	Conlon$^\gamma$	Conlon	ID$^\gamma$	Cure	MSM
Non-susceptible fraction								
$\beta_{s,0}$	-0.41	-0.40	-0.31	-0.60	-0.52		0.12	
$\beta_{s,x}$	0.50	0.50	0.48	0.57	0.56		0.51	
Non-susceptible: $1 \to 4$								
$\log(\kappa_{14})$	-8.52	-8.52	-7.26	-10.46	-9.33			
$\log(\alpha_{14})$	0.34	0.33	0.17	0.52	0.40			
$\beta_{14,x}$	0.25	0.26	0.17	0.47	0.40			
Susceptible: $2 \to 4$								
$\log(\kappa_{24})$	-5.30	-5.32	-5.25	-5.23	-5.14	-5.20		
$\log(\alpha_{24})$	0.34	0.34	0.26	0.29	0.25	0.26		
$\beta_{24,x}$	0.50	0.50	0.43	0.52	0.49	-0.05		0.08
Susceptible: $2 \to 3$								
$\log(\kappa_{23})$	-2.16	-2.17	-2.27	-1.83	-1.81	-2.08	-1.50	
$\log(\alpha_{23})$	-0.69	-0.69	-0.74	-0.64	-0.67	-0.64	-0.59	
$\beta_{23,x}$	0.25	0.25	0.23	0.24	0.24	-0.18	0.37	-0.26
Susceptible: $3 \to 4$								
$\log(\kappa_{34})$	-6.21	-6.21	-5.95	-6.19	-5.95	-7.38		
$\log(\alpha_{34})$	0.26	0.26	0.21	0.25	0.21	0.40		
$\beta_{34,x}$	0.15	0.15	0.14	0.15	0.14	0.03		0.14
log-frailty variance								
$\log(\theta)$	-1.71	-1.75		-2.02		0.81		

Table E.7: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets of size n=5,000 generated with administrative censoring via a censoring variable simulated from an Exponential distribution with mean 100.

	FI-ID$^\gamma$		FI-ID		Conlon$^\gamma$		Conlon		ID$^\gamma$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction														
$\beta_{s,0}$	0.07	0.06	0.06	0.06	0.15	0.08	0.07	0.07			0.07	0.07		
$\beta_{s,x}$	0.08	0.08	0.08	0.08	0.12	0.08	0.08	0.08			0.08	0.08		
Non-susceptible: $1 \to 4$														
$\log(\kappa_{14})$	0.91	0.87	0.67	0.69	1.17	1.22	0.98	0.99						
$\log(\alpha_{14})$	0.11	0.11	0.10	0.10	0.12	0.12	0.11	0.11						
$\beta_{14,x}$	0.16	0.16	0.14	0.14	0.53	0.21	0.18	0.18						
Susceptible: $2 \to 4$														
$\log(\kappa_{24})$	0.15	0.15	0.15	0.16	0.14	0.14	0.14	0.14	0.11	0.12				
$\log(\alpha_{24})$	0.04	0.03	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03				
$\beta_{24,x}$	0.10	0.09	0.10	0.09	0.10	0.10	0.09	0.09	0.09	0.09			0.04	0.04
Susceptible: $2 \to 3$														
$\log(\kappa_{23})$	0.14	0.13	0.18	0.17	0.06	0.05	0.05	0.05	0.05	0.05	0.06	0.06		
$\log(\alpha_{23})$	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03	0.03	0.03	0.03		
$\beta_{23,x}$	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.05	0.05
Susceptible: $3 \to 4$														
$\log(\kappa_{34})$	0.18	0.18	0.16	0.17	0.19	0.19	0.16	0.17	0.21	0.20				
$\log(\alpha_{34})$	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03				
$\beta_{34,x}$	0.08	0.08	0.08	0.07	0.08	0.08	0.08	0.07	0.10	0.10			0.05	0.05
log-frailty variance														
$\log(\theta)$	0.23	0.23			0.67	0.67			0.07	0.07				

Table E.8: Coverage probabilities, based on 2,000 simulated datasets of size n=5,000 generated with administrative censoring via a censoring variable simulated from an Exponential distribution with mean 100.

	FI-ID$^\gamma$	FI-ID	Conlon$^\gamma$	Conlon
Non-susceptible fraction				
$\beta_{s,0}$	0.95	0.63	0.34	0.66
$\beta_{s,x}$	0.95	0.94	0.86	0.89
Non-susceptible: $1 \to 4$				
$\log(\kappa_{14})$	0.93	0.54	0.67	0.89
$\log(\alpha_{14})$	0.94	0.62	0.66	0.89
$\beta_{14,x}$	0.95	0.90	0.82	0.89
Susceptible: $2 \to 4$				
$\log(\kappa_{24})$	0.95	0.92	0.92	0.78
$\log(\alpha_{24})$	0.95	0.45	0.72	0.17
$\beta_{24,x}$	0.95	0.87	0.94	0.95
Susceptible: $2 \to 3$				
$\log(\kappa_{23})$	0.95	0.97	0.00	0.00
$\log(\alpha_{23})$	0.95	0.71	0.50	0.83
$\beta_{23,x}$	0.94	0.93	0.95	0.94
Susceptible: $3 \to 4$				
$\log(\kappa_{34})$	0.96	0.63	0.95	0.63
$\log(\alpha_{34})$	0.96	0.56	0.95	0.56
$\beta_{34,x}$	0.95	0.94	0.95	0.94
log-frailty variance				
$\log(\theta)$	0.97		0.99	

E.4 Administrative censoring solely of the terminal event at 365 days

Table E.9: Mean point estimates, based on 2,000 simulated datasets of size n=5,000 with administrative censoring solely of the terminal event at 365 days post-transplantation. For the proposed finite interval illness-death model with frailty using B-spline parameterized baseline hazard functions (FI-ID$^{\gamma}{}_{BS}$), we made the following specifications: for all transitions we placed two internal knots at the 33rd and 67th percentiles of the observed event times; we used quadratic B-splines for the 1$\rightarrow$4 and 3$\rightarrow$4 transitions and cubic B-splines for the 2$\rightarrow$4 and 2$\rightarrow$3 transitions.

	Truth	FI-ID$^{\gamma}{}_{WB}$	FI-ID$^{\gamma}{}_{BS}$	FI-ID	Conlon$^{\gamma}$	Conlon	ID$^{\gamma}$	Cure	MSM
Non-susceptible fraction									
$\beta_{s,0}$	-0.41	-0.41	-0.38	-0.34	-0.54	-0.48		0.14	
$\beta_{s,x}$	0.50	0.50	0.49	0.50	0.54	0.53		0.50	
Non-susceptible: $1 \rightarrow 4$									
$\log(\kappa_{14})$	-8.52	-8.51		-7.53	-9.66	-8.78			
$\log(\alpha_{14})$	0.34	0.33		0.21	0.45	0.35			
$\beta_{14,x}$	0.25	0.25	0.23	0.21	0.33	0.29			
Susceptible: $2 \rightarrow 4$									
$\log(\kappa_{24})$	-5.30	-5.30		-5.16	-5.24	-5.10	-5.50		
$\log(\alpha_{24})$	0.34	0.34		0.23	0.31	0.25	0.32		
$\beta_{24,x}$	0.50	0.50	0.48	0.41	0.49	0.46	-0.03		-0.05
Susceptible: $2 \rightarrow 3$									
$\log(\kappa_{23})$	-2.16	-2.17		-2.38	-1.80	-1.79	-2.02	-1.50	
$\log(\alpha_{23})$	-0.69	-0.69		-0.76	-0.62	-0.66	-0.71	-0.58	
$\beta_{23,x}$	0.25	0.25	0.25	0.23	0.22	0.22	-0.23	0.38	-0.26
Susceptible: $3 \rightarrow 4$									
$\log(\kappa_{34})$	-6.21	-6.22		-5.79	-6.18	-5.79	-7.87		
$\log(\alpha_{34})$	0.26	0.26		0.18	0.25	0.18	0.48		
$\beta_{34,x}$	0.15	0.15	0.15	0.14	0.15	0.14	0.01		0.14
log-frailty variance									
$\log(\theta)$	-1.71	-1.73	-1.78		-1.89		0.80		

Table E.10: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets of size n=5,000 with administrative censoring solely of the terminal event at 365 days post-transplantation. For the proposed finite interval illness-death model with frailty using B-spline parameterized baseline hazard functions (FI-ID$^{\gamma}{}_{BS}$), we made the following specifications: for all transitions we placed two internal knots at the 33rd and 67th percentiles of the observed event times; we used quadratic B-splines for the 1→4 and 3→4 transitions and cubic B-splines for the 2→4 and 2→3 transitions.

	FI-ID$^{\gamma}{}_{WB}$		FI-ID$^{\gamma}{}_{BS}$		FI-ID		Conlon$^{\gamma}$		Conlon		ID$^{\gamma}$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction																
$\beta_{s,0}$	0.05	0.05	0.08	0.06	0.05	0.05	0.05	0.05	0.05	0.05			0.05	0.05		
$\beta_{s,x}$	0.06	0.06	0.07	0.07	0.06	0.06	0.06	0.06	0.06	0.06			0.07	0.07		
Non-susceptible: $1 \to 4$																
$\log(\kappa_{14})$	0.43	0.42			0.37	0.37	0.52	0.55	0.44	0.47						
$\log(\alpha_{14})$	0.05	0.05			0.05	0.05	0.05	0.06	0.05	0.05						
$\beta_{14,x}$	0.07	0.07	0.07	0.07	0.06	0.06	0.07	0.07	0.07	0.07						
Susceptible: $2 \to 4$																
$\log(\kappa_{24})$	0.15	0.14			0.13	0.13	0.13	0.13	0.13	0.13	0.11	0.11				
$\log(\alpha_{24})$	0.03	0.03			0.03	0.03	0.03	0.03	0.02	0.03	0.02	0.02				
$\beta_{24,x}$	0.08	0.08	0.09	0.08	0.09	0.08	0.08	0.08	0.08	0.08	0.08	0.08			0.04	0.04
Susceptible: $2 \to 3$																
$\log(\kappa_{23})$	0.12	0.11			0.17	0.16	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05		
$\log(\alpha_{23})$	0.03	0.03			0.03	0.03	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02		
$\beta_{23,x}$	0.06	0.06	0.07	0.06	0.06	0.06	0.06	0.06	0.05	0.06	0.07	0.07	0.06	0.06	0.05	0.05
Susceptible: $3 \to 4$																
$\log(\kappa_{34})$	0.14	0.14			0.12	0.12	0.14	0.15	0.12	0.12	0.17	0.16				
$\log(\alpha_{34})$	0.02	0.02			0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02				
$\beta_{34,x}$	0.06	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.08	0.08			0.05	0.05
log-frailty variance																
$\log(\theta)$	0.14	0.14	0.22	0.21			0.23	0.24			0.05	0.05				

Table E.11: Coverage probabilities, based on 2,000 simulated datasets of size n=5,000 with administrative censoring solely of the terminal event at 365 days post-transplantation.

	FI-ID$^{\gamma}{}_{WB}$	FI-ID$^{\gamma}{}_{BS}$	FI-ID	Conlon$^{\gamma}$	Conlon
Non-susceptible fraction					
$\beta_{s,0}$	0.95	0.85	0.69	0.32	0.74
$\beta_{s,x}$	0.95	0.93	0.95	0.91	0.91
Non-susceptible: $1 \rightarrow 4$					
$\log(\kappa_{14})$	0.95		0.24	0.46	0.93
$\log(\alpha_{14})$	0.94		0.23	0.49	0.95
$\beta_{14,x}$	0.95	0.94	0.87	0.80	0.92
Susceptible: $2 \rightarrow 4$					
$\log(\kappa_{24})$	0.94		0.81	0.92	0.63
$\log(\alpha_{24})$	0.95		0.08	0.82	0.06
$\beta_{24,x}$	0.95	0.95	0.76	0.94	0.91
Susceptible: $2 \rightarrow 3$					
$\log(\kappa_{23})$	0.95		0.88	0.00	0.00
$\log(\alpha_{23})$	0.95		0.46	0.12	0.62
$\beta_{23,x}$	0.95	0.91	0.93	0.93	0.91
Susceptible: $3 \rightarrow 4$					
$\log(\kappa_{34})$	0.95		0.07	0.95	0.07
$\log(\alpha_{34})$	0.95		0.01	0.94	0.01
$\beta_{34,x}$	0.94	0.95	0.93	0.94	0.93
log-frailty variance					
$\log(\theta)$	0.96	0.96		0.96	

E.5 Increasing the non-susceptibility fraction from 46% to 76%

Table E.12: Mean point estimates, based on 2,000 simulated datasets of size n=5,000 generated by increasing the non-susceptibility fraction from 46% to 76%.

	Truth	FI-ID$^\gamma$	FI-ID	Conlon$^\gamma$	Conlon	ID$^\gamma$	Cure	MSM
Non-susceptible fraction								
$\beta_{s,0}$	0.84	0.84	0.96	0.74	0.91		1.15	
$\beta_{s,x}$	0.50	0.50	0.54	0.53	0.56		0.61	
Non-susceptible: $1 \to 4$								
$\log(\kappa_{14})$	-8.52	-8.50	-7.23	-8.86	-7.34			
$\log(\alpha_{14})$	0.34	0.33	0.16	0.37	0.18			
$\beta_{14,x}$	0.25	0.25	0.21	0.28	0.22			
Susceptible: $2 \to 4$								
$\log(\kappa_{24})$	-5.30	-5.31	-5.31	-5.33	-5.19	-6.44		
$\log(\alpha_{24})$	0.34	0.34	0.21	0.32	0.25	0.08		
$\beta_{24,x}$	0.50	0.50	0.33	0.52	0.41	0.14		0.15
Susceptible: $2 \to 3$								
$\log(\kappa_{23})$	-2.16	-2.18	-2.09	-1.83	-1.73	-2.76	-1.45	
$\log(\alpha_{23})$	-0.69	-0.69	-0.72	-0.62	-0.63	-1.45	-0.58	
$\beta_{23,x}$	0.25	0.25	0.30	0.23	0.30	-0.39	0.37	-0.40
Susceptible: $3 \to 4$								
$\log(\kappa_{34})$	-6.21	-6.22	-5.67	-6.26	-5.67	-5.91		
$\log(\alpha_{34})$	0.26	0.26	0.15	0.27	0.15	0.20		
$\beta_{34,x}$	0.15	0.15	0.14	0.15	0.14	0.14		0.14
log-frailty variance								
$\log(\theta)$	-1.71	-1.74		-1.68		-2.97		

Table E.13: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets of size n=5,000 generated by increasing the non-susceptibility fraction from 46% to 76%.

	FI-ID$^\gamma$		FI-ID		Conlon$^\gamma$		Conlon		ID$^\gamma$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction														
$\beta_{s,0}$	0.05	0.05	0.05	0.05	0.06	0.06	0.06	0.06			1.08	0.05		
$\beta_{s,x}$	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08			0.89	0.08		
Non-susceptible: $1 \to 4$														
$\log(\kappa_{14})$	0.24	0.24	0.16	0.14	0.33	0.33	0.17	0.15						
$\log(\alpha_{14})$	0.03	0.03	0.02	0.02	0.04	0.04	0.02	0.02						
$\beta_{14,x}$	0.04	0.04	0.04	0.03	0.04	0.27	0.04	0.03						
Susceptible: $2 \to 4$														
$\log(\kappa_{24})$	0.22	0.22	0.25	0.26	0.26	0.80	0.24	0.24	0.11	0.10				
$\log(\alpha_{24})$	0.05	0.05	0.06	0.06	0.13	0.63	0.04	0.05	0.02	0.02				
$\beta_{24,x}$	0.12	0.12	0.19	0.16	0.14	0.31	0.15	0.15	0.03	0.03			0.03	0.03
Susceptible: $2 \to 3$														
$\log(\kappa_{23})$	0.18	0.17	0.18	0.17	0.08	0.73	0.08	0.07	0.05	0.06	0.21	0.07		
$\log(\alpha_{23})$	0.04	0.05	0.04	0.05	0.04	0.03	0.03	0.03	0.02	0.03	0.06	0.03		
$\beta_{23,x}$	0.10	0.10	0.10	0.10	0.10	0.75	0.10	0.10	0.07	0.07	0.22	0.09	0.07	0.07
Susceptible: $3 \to 4$														
$\log(\kappa_{34})$	0.19	0.19	0.17	0.17	0.20	0.86	0.17	0.17	0.19	0.19				
$\log(\alpha_{34})$	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03				
$\beta_{34,x}$	0.08	0.08	0.08	0.07	0.09	0.88	0.08	0.07	0.08	0.07			0.07	0.07
log-frailty variance														
$\log(\theta)$	0.16	0.16			0.92	0.19			1.02	0.54				

Table E.14: Coverage probabilities, based on 2,000 simulated datasets of size n=5,000 generated by increasing the non-susceptibility fraction from 46% to 76%.

	FI-ID$^{\gamma}$	FI-ID	Conlon$^{\gamma}$	Conlon
Non-susceptible fraction				
$\beta_{s,0}$	0.95	0.37	0.63	0.81
$\beta_{s,x}$	0.95	0.93	0.93	0.90
Non-susceptible: $1 \rightarrow 4$				
$\log(\kappa_{14})$	0.95	0.00	0.85	0.00
$\log(\alpha_{14})$	0.95	0.00	0.85	0.00
$\beta_{14,x}$	0.95	0.76	0.92	0.83
Susceptible: $2 \rightarrow 4$				
$\log(\kappa_{24})$	0.94	0.96	0.95	0.92
$\log(\alpha_{24})$	0.95	0.47	0.95	0.54
$\beta_{24,x}$	0.95	0.79	0.95	0.90
Susceptible: $2 \rightarrow 3$				
$\log(\kappa_{23})$	0.95	0.87	0.01	0.00
$\log(\alpha_{23})$	0.95	0.93	0.43	0.49
$\beta_{23,x}$	0.95	0.91	0.95	0.94
Susceptible: $3 \rightarrow 4$				
$\log(\kappa_{34})$	0.95	0.12	0.94	0.12
$\log(\alpha_{34})$	0.95	0.02	0.94	0.02
$\beta_{34,x}$	0.95	0.93	0.95	0.93
log-frailty variance				
$\log(\theta)$	0.96		0.91	

E.6 Truncated and standard Weibull distributions align on $(0, \tau)$

Table E.15: Mean point estimates, based on 2,000 simulated datasets of size n=5,000 under the modification of α_{23} and κ_{23} so that the hazard functions for the truncated and standard Weibull distributions are comparable on the interval $(0, \tau)$.

	Truth	FI-ID$^\gamma$	FI-ID	Conlon$^\gamma$	Conlon	ID$^\gamma$	Cure	MSM
Non-susceptible fraction								
$\beta_{s,0}$	-0.41	-0.41	-0.35	-0.38	-0.34		0.51	
$\beta_{s,x}$	0.50	0.50	0.51	0.50	0.51		0.66	
Non-susceptible: $1 \to 4$								
$\log(\kappa_{14})$	-8.52	-8.51	-7.28	-8.22	-7.21			
$\log(\alpha_{14})$	0.34	0.34	0.17	0.30	0.16			
$\beta_{14,x}$	0.25	0.25	0.21	0.24	0.20			
Susceptible: $2 \to 4$								
$\log(\kappa_{24})$	-5.30	-5.30	-5.08	-5.29	-5.04	-6.05		
$\log(\alpha_{24})$	0.34	0.34	0.24	0.35	0.27	0.40		
$\beta_{24,x}$	0.50	0.50	0.45	0.49	0.46	0.02		0.10
Susceptible: $2 \to 3$								
$\log(\kappa_{23})$	-7.13	-7.14	-6.70	-7.17	-6.78	-4.49	-6.18	
$\log(\alpha_{23})$	0.64	0.64	0.55	0.65	0.57	-0.05	0.54	
$\beta_{23,x}$	0.25	0.25	0.22	0.25	0.23	-0.45	0.38	-0.46
Susceptible: $3 \to 4$								
$\log(\kappa_{34})$	-6.21	-6.23	-5.70	-6.16	-5.70	-7.73		
$\log(\alpha_{34})$	0.26	0.26	0.15	0.25	0.15	0.48		
$\beta_{34,x}$	0.15	0.15	0.13	0.15	0.13	-0.01		0.13
log-frailty variance								
$\log(\theta)$	-1.71	-1.73		-1.90		0.58		

Table E.16: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets of size n=5,000 under the modification of α_{23} and κ_{23} so that the hazard functions for the truncated and standard Weibull distributions are comparable on the interval $(0, \tau)$.

	FI-ID$^\gamma$		FI-ID		Conlon$^\gamma$		Conlon		ID$^\gamma$		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction														
$\beta_{s,0}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05			0.05	0.05		
$\beta_{s,x}$	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06			0.07	0.07		
Non-susceptible: $1 \to 4$														
$\log(\kappa_{14})$	0.27	0.27	0.20	0.19	0.26	0.26	0.20	0.19						
$\log(\alpha_{14})$	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.02						
$\beta_{14,x}$	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.04						
Susceptible: $2 \to 4$														
$\log(\kappa_{24})$	0.13	0.12	0.10	0.11	0.12	0.12	0.11	0.11	0.25	0.16				
$\log(\alpha_{24})$	0.03	0.03	0.03	0.03	0.02	0.02	0.02	0.02	0.06	0.04				
$\beta_{24,x}$	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.08	0.07			0.03	0.03
Susceptible: $2 \to 3$														
$\log(\kappa_{23})$	0.18	0.17	0.15	0.16	0.17	0.17	0.15	0.15	0.17	0.13	0.15	0.15		
$\log(\alpha_{23})$	0.03	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.07	0.04	0.02	0.02		
$\beta_{23,x}$	0.07	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.07	0.06	0.06	0.06
Susceptible: $3 \to 4$														
$\log(\kappa_{34})$	0.16	0.16	0.14	0.14	0.16	0.16	0.14	0.14	0.21	0.19				
$\log(\alpha_{34})$	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.03	0.02				
$\beta_{34,x}$	0.07	0.07	0.07	0.06	0.07	0.07	0.07	0.06	0.10	0.09			0.06	0.06
log-frailty variance														
$\log(\theta)$	0.13	0.13			0.15	0.16			0.11	0.07				

Table E.17: Coverage probabilities, based on 2,000 simulated datasets of size n=5,000 under the modification of α_{23} and κ_{23} so that the hazard functions for the truncated and standard Weibull distributions are comparable on the interval $(0, \tau)$.

	FI-ID$^\gamma$	FI-ID	Conlon$^\gamma$	Conlon
Non-susceptible fraction				
$\beta_{s,0}$	0.95	0.78	0.91	0.70
$\beta_{s,x}$	0.95	0.95	0.95	0.95
Non-susceptible: $1 \to 4$				
$\log(\kappa_{14})$	0.95	0.00	0.79	0.00
$\log(\alpha_{14})$	0.95	0.00	0.79	0.00
$\beta_{14,x}$	0.95	0.82	0.94	0.80
Susceptible: $2 \to 4$				
$\log(\kappa_{24})$	0.94	0.46	0.95	0.35
$\log(\alpha_{24})$	0.94	0.02	0.92	0.12
$\beta_{24,x}$	0.95	0.81	0.95	0.88
Susceptible: $2 \to 3$				
$\log(\kappa_{23})$	0.95	0.22	0.94	0.37
$\log(\alpha_{23})$	0.95	0.04	0.94	0.15
$\beta_{23,x}$	0.96	0.92	0.96	0.94
Susceptible: $3 \to 4$				
$\log(\kappa_{34})$	0.95	0.06	0.94	0.06
$\log(\alpha_{34})$	0.95	0.00	0.93	0.00
$\beta_{34,x}$	0.95	0.91	0.95	0.91
log-frailty variance				
$\log(\theta)$	0.95		0.84	

E.7 Baseline scenario without frailty, i.e. Gamma frailty variance equal to zero

Table E.18: Mean point estimates, based on 2,000 simulated datasets of size n=5,000 under the modification of no frailty.

	Truth	FI-ID	Conlon	ID	Cure	MSM
Non-susceptible fraction						
$\beta_{s,0}$	-0.41	-0.41	-0.46		0.18	
$\beta_{s,x}$	0.50	0.50	0.52		0.45	
Non-susceptible: $1 \rightarrow 4$						
$\log(\kappa_{14})$	-8.52	-8.53	-8.84			
$\log(\alpha_{14})$	0.34	0.34	0.37			
$\beta_{14,x}$	0.25	0.25	0.27			
Susceptible: $2 \rightarrow 4$						
$\log(\kappa_{24})$	-5.30	-5.31	-5.26	-4.92		
$\log(\alpha_{24})$	0.34	0.34	0.33	-0.13		
$\beta_{24,x}$	0.50	0.50	0.50	0.06		0.10
Susceptible: $2 \rightarrow 3$						
$\log(\kappa_{23})$	-2.16	-2.16	-1.77	-1.97	-1.47	
$\log(\alpha_{23})$	-0.69	-0.69	-0.60	-1.32	-0.51	
$\beta_{23,x}$	0.25	0.25	0.23	-0.28	0.35	-0.25
Susceptible: $3 \rightarrow 4$						
$\log(\kappa_{34})$	-6.21	-6.22	-6.22	-6.22		
$\log(\alpha_{34})$	0.26	0.26	0.26	0.26		
$\beta_{34,x}$	0.15	0.15	0.15	0.15		0.17

Table E.19: Empirical and analytical standard errors calculated using the standard deviations of the sampling distributions and means of the estimated analytical standard errors, based on 2,000 simulated datasets of size n=5,000 under the modification of no frailty.

	FI-ID		Conlon		ID		Cure		MSM	
	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$	SD	$\widehat{\text{SE}}$
Non-susceptible fraction										
$\beta_{s,0}$	0.05	0.04	0.05	0.05			0.05	0.05		
$\beta_{s,x}$	0.06	0.06	0.06	0.06			0.06	0.06		
Non-susceptible: $1 \rightarrow 4$										
$\log(\kappa_{14})$	0.21	0.20	0.22	0.22						
$\log(\alpha_{14})$	0.02	0.02	0.02	0.02						
$\beta_{14,x}$	0.04	0.04	0.05	0.04						
Susceptible: $2 \rightarrow 4$										
$\log(\kappa_{24})$	0.13	0.14	0.13	0.13	0.08	0.08				
$\log(\alpha_{24})$	0.03	0.03	0.02	0.03	0.01	0.01				
$\beta_{24,x}$	0.08	0.08	0.08	0.08	0.03	0.04			0.04	0.04
Susceptible: $2 \rightarrow 3$										
$\log(\kappa_{23})$	0.10	0.10	0.05	0.05	0.03	0.04	0.05	0.05		
$\log(\alpha_{23})$	0.03	0.03	0.02	0.02	0.01	0.02	0.02	0.02		
$\beta_{23,x}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.05
Susceptible: $3 \rightarrow 4$										
$\log(\kappa_{34})$	0.12	0.12	0.12	0.12	0.12	0.12				
$\log(\alpha_{34})$	0.02	0.02	0.02	0.02	0.02	0.02				
$\beta_{34,x}$	0.05	0.05	0.05	0.05	0.05	0.05			0.05	0.05

Table E.20: Coverage probabilities, based on 2,000 simulated datasets of size n=5,000 under the modification of no frailty.

	FI-ID	Conlon
Non-susceptible fraction		
$\beta_{s,0}$	0.95	0.78
$\beta_{s,x}$	0.95	0.94
Non-susceptible: $1 \rightarrow 4$		
$\log(\kappa_{14})$	0.94	0.71
$\log(\alpha_{14})$	0.94	0.71
$\beta_{14,x}$	0.95	0.93
Susceptible: $2 \rightarrow 4$		
$\log(\kappa_{24})$	0.96	0.95
$\log(\alpha_{24})$	0.95	0.95
$\beta_{24,x}$	0.95	0.95
Susceptible: $2 \rightarrow 3$		
$\log(\kappa_{23})$	0.95	0
$\log(\alpha_{23})$	0.96	0.01
$\beta_{23,x}$	0.95	0.93
Susceptible: $3 \rightarrow 4$		
$\log(\kappa_{34})$	0.96	0.96
$\log(\alpha_{34})$	0.96	0.96
$\beta_{34,x}$	0.94	0.94

F Additional results from the analysis of stem cell transplantation data

Table F.1: Distribution of events for CIBMTR HCT data (n=9,651)

	n	%
Acute GVHD, death within 100 days of transplantation	502	5.2
Acute GVHD, death 100 days post-transplantation	556	5.8
Acute GVHD, censored before 100 days	1	0.0
Acute GVHD, censored after 100 days post-transplantation	643	6.7
Death within 100 days without acute GVHD	983	10.2
Death after 100 days, without acute GVHD	1786	18.5
Censored after 100 days, without either	5159	53.5
Censored within 100 days, without either	21	0.2

Table F.2: Estimates and standard errors for Weibull parameters from analyses of the CIBMTR HCT data (n=9,651) including: the proposed finite interval illness-death model with frailty (FI-ID$^\gamma$); the cure fraction illness-death model of Conlon et al. (2014) with an additional frailty (Conlon$^\gamma$); the illness-death model with frailty (ID$^\gamma$); and the cure fraction model for acute GVHD.

	FI-ID$^\gamma$		Conlon$^\gamma$		ID$^\gamma$		Cure	
	Est	SE	Est	SE	Est	SE	Est	SE
Weibull parameters								
$\log(\kappa_{14})$	-7.98	0.18	-7.56	0.17				
$\log(\alpha_{14})$	0.14	0.03	0.08	0.03				
$\log(\kappa_{24})$	-16.16	0.03	-11.76	0.95	-7.43	0.17		
$\log(\alpha_{24})$	0.80	0.02	0.73	0.08	0.32	0.03		
$\log(\kappa_{23})$	-6.83	0.18	-6.58	0.18	-14.14	0.27	-5.93	0.13
$\log(\alpha_{23})$	0.66	0.02	0.62	0.03	1.05	0.02	0.48	0.02
$\log(\kappa_{34})$	-5.44	0.16	-5.33	0.15	-7.91	0.19		
$\log(\alpha_{34})$	-0.10	0.03	-0.13	0.03	0.07	0.03		

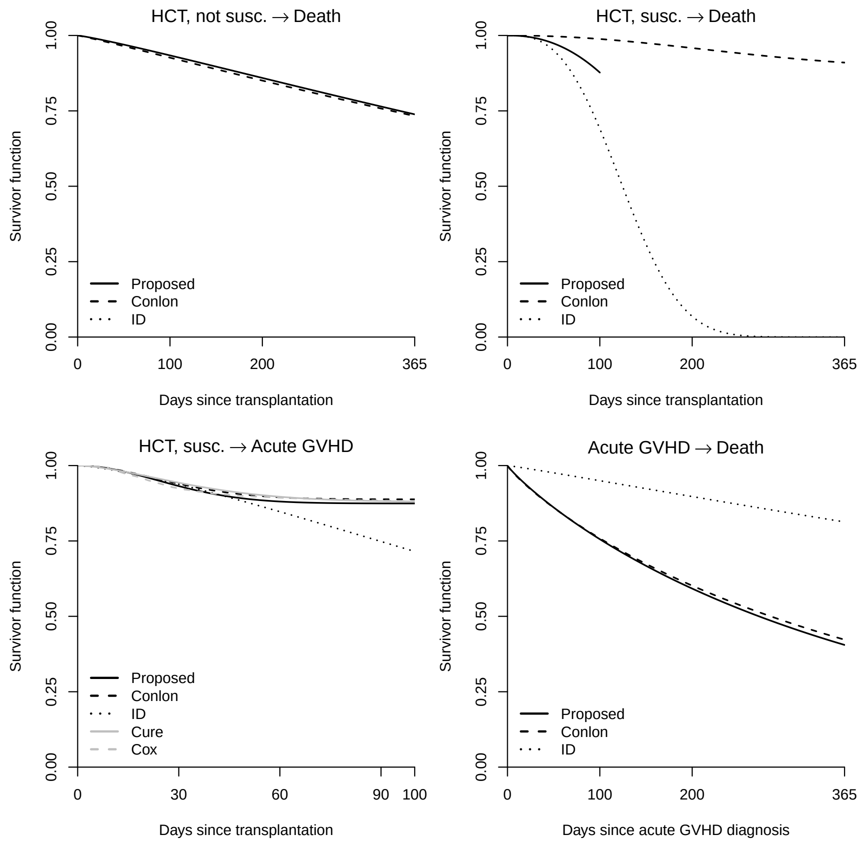


Figure F.1: Estimated baseline survivor functions from the CIBMTR HCT data for four transitions, based on: the proposed finite interval illness-death model with frailty (FI-ID$^\gamma$); the cure fraction illness-death model of Conlon et al. (2014) with an additional frailty (Conlon$^\gamma$); the illness-death model with frailty (ID$^\gamma$); the cure fraction model for acute GVHD; and the Cox model for acute GVHD. The marginal baseline survivor functions (with respect to the frailty, γ) are shown for models FI-ID$^\gamma$ and Conlon$^\gamma$ for the 'HCT, susc.→ Death' and 'HCT, susc.→ Acute GVHD' transitions.

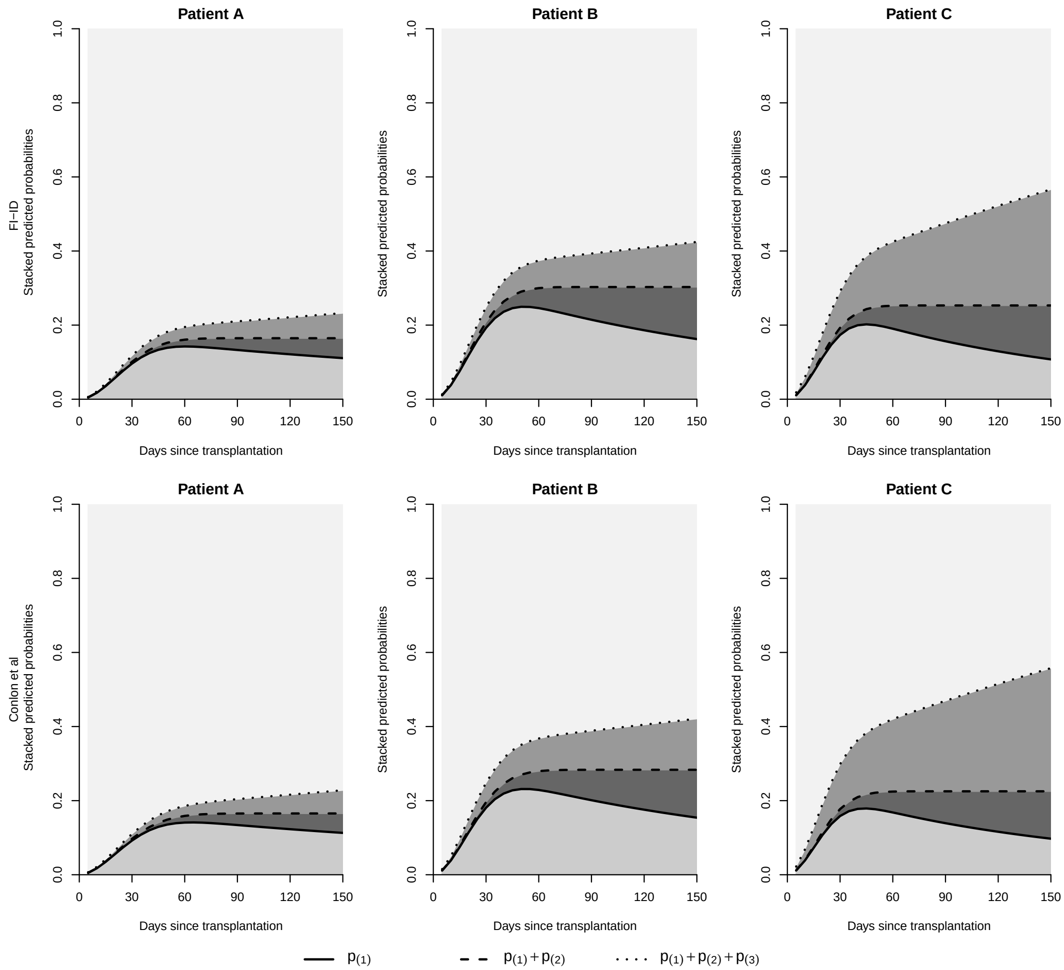
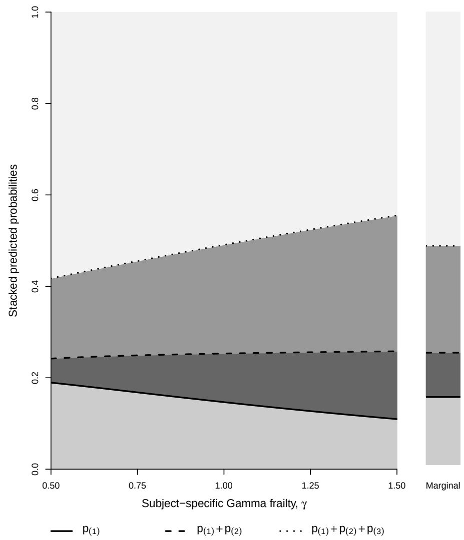
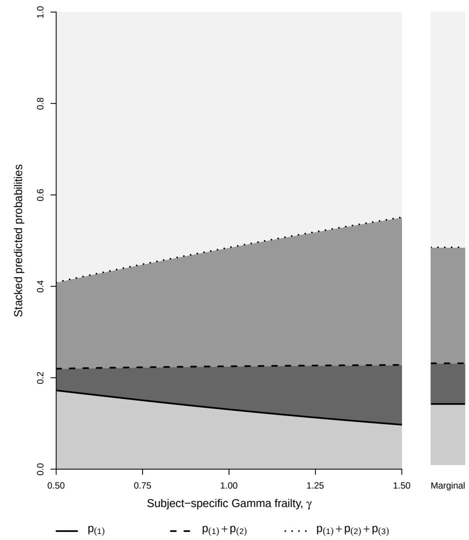


Figure F.2: Estimated absolute risk profiles calculated for the first 150 post-transplantation corresponding to three patients (A, B, C) based on proposed finite-interval illness-death model with frailty (FI-ID$^{\gamma}$) in the first row, and the cure fraction illness-death model of Conlon et al. (2014) with an additional frailty (Conlon$^{\gamma}$) in the second row.



(a) FI-ID$^{\gamma}$



(b) Conlon$^{\gamma}$

Figure F.3: For model: (a) proposed finite-interval illness-death model with frailty (FI-ID$^{\gamma}$), and (b) Conlon et al.(2014) with an additional frailty (Conlon$^{\gamma}$). Within each plot, the left panel presents the estimated absolute risk profiles (conditional on frailty, γ) for Patient C at 100 days post-transplantation for varying values of the gamma frailty, γ, and the right panel presents the estimated marginal absolute risk profiles for Patient C at 100 days post-transplantation.

G Sample R code

All statistical analyses were conducted in R [4]. In our implementation of the proposed finite interval illness-death model with both Weibull or truncated Weibull baseline hazard functions as well as B-spline parameterizations, separately, we make use of the `optim` function in the `stats` package for quasi-Newton non-linear optimization of the log-likelihood. To obtain starting values for the non-linear optimization, for the proposed model with Weibull and truncated Weibull baseline hazard, we utilized: the `weibreg` function in the `eha` package for univariate Weibull regression [5]; and `FreqID_HReg` function in the `SemiCompRisks` package [6] to fit an illness-death model with shared frailty. For the proposed model with B-spline parameterized baseline hazard functions, we used: the `bSpline` function in the `splines2` package[7]; the `coxph` function in the `survival` package [8]; and the `numdiff` function in the `pracma` [9] package for numerical differentiation.

We provide R script and functions to run the proposed finite interval illness-death model with Weibull and truncated Weibull and B-spline baseline hazard parameterizations; please see files: `sim-script-SemiCompRisksv3_1.R` and `sim-functions-SemiCompRisksv3_1.R`.

References

[1] Xu J, Kalbfleisch J and Tai B. Statistical analysis of illness–death processes and semicompeting risks data. *Biometrics* 2010; 66(3): 716–725.

[2] Lee KH, Dominici F, Schrag D et al. Hierarchical models for semicompeting risks data with application to quality of end-of-life care for pancreatic cancer. *Journal of the American Statistical Association* 2016; 111(515): 1075–1095.

[3] Putter H, Fiocco M and Geskus R. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; 26(11): 2389–2430.

[4] R Core Team. R: A language and environment for statistical computing, 2017. URL `https://www.R-project.org/`.

[5] Brostrom G. eha: Event history analysis, 2016. URL `http://CRAN.R-project.org/package=eha`. R package version 2.4-4.

[6] Lee K, Lee C, Alvares D et al. SemiCompRisks: Hierarchical models for parametric and semi-parametric analyses of semi-competing risks data, 2019. URL `http://CRAN.R-project.org/package=SemiCompRisks`. R package version 3.3.

[7] Wenjie W and Jun Y. *splines2: Regression Spline Functions and Classes Too*, 2017. URL `https://CRAN.R-project.org/package=splines2`. R package version 0.2.5.

[8] Therneau T. *A Package for Survival Analysis in S*, 2015. URL `https://CRAN.R-project.org/package=survival`. R package version 2.38.

[9] Borchers H. *pracma: Practical Numerical Math Functions*, 2019. URL `https://CRAN.R-project.org/package=pracma`. R package version 2.2.5.