SEMIPARAMETRIC REGRESSION MODEL FOR RECURRENT BACTERIAL INFECTIONS AFTER HEMATOPOIETIC STEM CELL TRANSPLANTATION

Chi Hyun Lee[1], Chiung-Yu Huang[2], Todd E. DeFor[3], Claudio G. Brunstein[3]

Daniel J. Weisdorf[3] and Xianghua Luo[3]

[1] *University of Massachusetts, Amherst,*

[2] *University of California, San Francisco and* [3] *University of Minnesota*

Supplementary Material

Web Appendix A. Proof under the Random Censoring Assumption

A.1. Regularity conditions

We assume the following regularity conditions:

(C1) The true parameter $\boldsymbol{\beta}$ is in the interior of the parameter space $\mathbb{R}^{2p}$.

(C2) $\mathbf{A}$ is a $p \times 1$ vector of covariates that is bounded.

(C3) Σ is nonsingular.

A.2. Uniqueness and consistency of $\hat{\boldsymbol{\beta}}$

We begin by reformulating the estimating functions (2.2) and (2.4) as

$$\mathbf{D}_0(\mathbf{b}_0) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1,\mathbf{a}_2,\mathbf{b}_0)(\mathbf{a}_2-\mathbf{a}_1)\frac{O_{L_0}(t,s)}{\hat{G}_0(t\wedge L_0)}\hat{F}_0(\mathrm{d}t,\mathrm{d}s,\mathrm{d}\mathbf{a}_1;\mathbf{a}_2,\mathbf{b}_0)\hat{H}(\mathrm{d}\mathbf{a}_2), \tag{S.1}$$

$$\mathbf{D}_1^*(\mathbf{b}) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1,\mathbf{a}_2,\mathbf{b}_1)(\mathbf{a}_2-\mathbf{a}_1)\frac{O_{L_1}(t,s)}{\hat{G}_1(t\wedge L_1)}\hat{F}_1^*(\mathrm{d}t,\mathrm{d}s,\mathrm{d}\mathbf{a}_1;\mathbf{a}_2,\mathbf{b})\hat{H}(\mathrm{d}\mathbf{a}_2), \tag{S.2}$$

where $\hat{F}_0$ is the empirical estimator of the subdistribution function $F_0(t,s,\mathbf{a}_1;\mathbf{a}_2,\mathbf{b}_0) = \Pr[Z_{i0}\le t, \exp\{(\mathbf{a}_2-\mathbf{A}_i)^T\mathbf{b}_0\}X_i \le s, \mathbf{A}_i\le \mathbf{a}_1, \Delta_{i0}=1]$, $\hat{F}_1^*(t,s,\mathbf{a}_1;\mathbf{a}_2,\mathbf{b})$ is

$$\frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=1}^{m_i^*} I\left[Z_{ij}\le t, \exp\{(\mathbf{a}_2-\mathbf{A}_i)^T\mathbf{b}_0\}X_i + \exp\{(\mathbf{a}_2-\mathbf{A}_i)^T\mathbf{b}_1\}Y_{ij}\le s, \mathbf{A}_i\le\mathbf{a}_1, \Delta_{ij}=1\right],$$

the weighted average version of the empirical estimator of $F_1(t,s,\mathbf{a}_1;\,\mathbf{a}_2,\mathbf{b}) = \Pr[Z_{ij}\le t, \exp\{(\mathbf{a}_2-\mathbf{A}_i)^T\mathbf{b}_0\}X_i + \exp\{(\mathbf{a}_2-\mathbf{A}_i)^T\mathbf{b}_1\}Y_{ij}\le s, \mathbf{A}_i\le\mathbf{a}_1, \Delta_{ij}=1]$ and $\hat{H}$ is the empirical distribution function of $H(\mathbf{a}_2) = \Pr(\mathbf{A}_i\le\mathbf{a}_2)$. The Kaplan–Meier estimators $\hat{G}_0(t)$ and $\hat{G}_1(t)$ can be expressed as continuous and compactly differentiable functions (Gill and Johansen, 1990). Empirical estimators $\hat{F}_0$, $\hat{F}_1^*$, and $\hat{H}$ are also continuous and compactly differentiable functionals, and it follows that $\mathbf{D}_0$ and $\mathbf{D}_1^*$ are continuous and compactly differentiable. Since estimating function (S.1) is monotone in $\mathbf{b}_0$, the solution to $\mathbf{D}_0(\mathbf{b}_0) = 0$ is unique. Equation (S.2) is also monotone in $\mathbf{b}_1$ given $\mathbf{b}_0$, and $\mathbf{D}_1^*\{(\hat{\boldsymbol{\beta}}_0{}^T, \mathbf{b}_1^T)^T\} = 0$ has a unique solution.

Note that $\hat{G}_0$, $\hat{G}_1$, $\hat{F}_0$, $\hat{F}_1^*$, and $\hat{H}$ are uniformly consistent estimators. The consistency of $\hat{\boldsymbol{\beta}}_0$ corresponding to the time from transplant to the first infection has been established by Huang (2002). Given that $\hat{\boldsymbol{\beta}}_0$ is consistent for $\boldsymbol{\beta}_0$, $\mathbf{D}_1^{*T}(\mathbf{b})(\mathbf{b}_1 - \boldsymbol{\beta}_1)$ converges almost surely and uniformly in $\mathbf{b}$ to

$$\mathrm{E}\left[\mathrm{E}\{\, w(\mathbf{A}_i, \mathbf{A}_{i'}, \mathbf{b}_1)\mathbf{A}_{ii'}^T(\mathbf{b}_1 - \boldsymbol{\beta}_1)O_{L_1}\left[Z_{ij}^0, Z_{ii'j}^0\{(\boldsymbol{\beta}_0^T, \mathbf{b}_1^T)^T\}\right] \big|\, \mathbf{A}_i, \mathbf{A}_{i'}\}\right], \quad \text{(S.3)}$$

which equals $\mathrm{E}\left[\mathrm{E}\{w(\mathbf{A}_i, \mathbf{A}_{i'}, \mathbf{b}_1)\mathbf{A}_{ii'}^T(\mathbf{b}_1 - \boldsymbol{\beta}_1)O_{L_1}[Z_{ij}^0, \exp(A_{ii'}^T\boldsymbol{\beta}_0)X_i^0 + \exp\{\mathbf{A}_{ii'}^T (\mathbf{b}_1 - \boldsymbol{\beta}_1)\}\exp(A_{ii'}^T\boldsymbol{\beta}_1)Y_{ij}^0]|\mathbf{A}_i, \mathbf{A}_{i'}\}\right]$. Expression (S.3) is equal to 0 only when $\mathbf{b}_1 = \boldsymbol{\beta}_1$, which implies strong consistency of $\hat{\boldsymbol{\beta}}_1$ for $\boldsymbol{\beta}_1$. Thus, given the consistency of $\hat{\boldsymbol{\beta}}_0$, the consistency of the estimator $\hat{\boldsymbol{\beta}}$ follows.

A.3. Asymptotic normality of $\mathbf{D}(\boldsymbol{\beta})$

Define $\mathbf{D}(\mathbf{b}) \equiv \{\mathbf{D}_0^T(\mathbf{b}_0), \mathbf{D}_1^{*T}(\mathbf{b})\}^T$. By the functional delta method and the influence function approach, $n^{1/2}\mathbf{D}(\boldsymbol{\beta})$ is asymptotically normal with mean zero and variance Ω. Following the proof in Huang (2002), we derive the sensitivity curves of $n^{1/2}\mathbf{D}_0(\boldsymbol{\beta}_0)$ and $n^{1/2}\mathbf{D}_1^*(\boldsymbol{\beta})$ as follows, for i:

$$\xi_{i0}(\boldsymbol{\beta}_0) = n^{-3/2}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_0)\mathbf{A}_{ii'}\left[\frac{\Delta_{i0}O_{L_0}\{Z_{i0}, Z_{ii'0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i0}\wedge L_0)} - \frac{\Delta_{i'0}O_{L_0}\{Z_{i'0}, Z_{i'i0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i'0}\wedge L_0)}\right] + n^{-3/2}\int_0^{L_0}\frac{Q_0(t,\boldsymbol{\beta}_0)\hat{G}_0(t-)}{Y_0(t)\hat{G}_0(t)}\mathrm{d}\hat{M}_{i0}(t), \tag{S.4}$$

$$\xi_{i1}^*(\boldsymbol{\beta}) = n^{-3/2}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_1)\mathbf{A}_{ii'}\left[\frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\frac{\Delta_{ij}O_{L_1}\{Z_{ij}, Z_{ii'j}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{ij}\wedge L_1)} - \frac{1}{m_{i'}^*}\sum_{l=1}^{m_{i'}^*}\frac{\Delta_{i'l}O_{L_1}\{Z_{i'l}, Z_{i'il}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{i'l}\wedge L_1)}\right] + n^{-3/2}\int_0^{L_1}\frac{Q_1^*(t,\boldsymbol{\beta})\hat{G}_1(t-)}{Y_1^*(t)\hat{G}_1(t)}\mathrm{d}\hat{M}_{i1}^*(t), \tag{S.5}$$

in which,

$$Q_0(t,\boldsymbol{\beta}_0) = \sum_{i=1}^{n}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_0)\mathbf{A}_{ii'}\left[\frac{\Delta_{i0}O_{L_0}\{Z_{i0}, Z_{ii'0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i0}\wedge L_0)}I(Z_{i0} > t)\right],$$

$$Q_1^*(t,\boldsymbol{\beta}) = \sum_{i=1}^{n}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_1)\mathbf{A}_{ii'}\left[\frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\frac{\Delta_{ij}O_{L_1}\{Z_{ij}, Z_{ii'j}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{ij}\wedge L_1)}I(Z_{ij} > t)\right],$$

$Y_0(t) = \sum_{i=1}^{n} I(Z_{i0} \geq t)$, $\hat{M}_{i0}(t) = I(Z_{i0} \leq t, \Delta_{i0} = 0) - \int_0^t I(Z_{i0} \geq s)\mathrm{d}\hat{\Lambda}_0(s)$,

$$Y_1^*(t) = \sum_{i=1}^{n} \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} I(Z_{ij} \geq t), \text{ and}$$
$$\hat{M}_{i1}^*(t) = \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} I(Z_{ij} \leq t, \Delta_{ij} = 0) - \int_0^t \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} I(Z_{ij} \geq s)\mathrm{d}\hat{\Lambda}_1(s),$$

and $\hat{\Lambda}_k$ is the Nelson-Aalen estimator corresponding to $\hat{G}_k$ for $k = 0, 1$. Note that the last terms on the right-hand side of equations (S.4) and (S.5) are derived based on the martingale representation of $\hat{G}_0(t)$ and $\hat{G}_1(t)$. The variance Ω can be estimated by $\hat{\Omega} = \sum_{i=1}^{n} \{\xi_{i0}^T(\hat{\boldsymbol{\beta}}_0), \xi_{i1}^{*T}(\hat{\boldsymbol{\beta}})\}^T \{\xi_{i0}^T(\hat{\boldsymbol{\beta}}_0), \xi_{i1}^{*T}(\hat{\boldsymbol{\beta}})\}$, which is shown to be a consistent estimator by the Glivenko-Cantelli theorem of Pollard (1984).

A.4. Asymptotic linearity of $\mathbf{D}(\mathbf{b})$ at $\mathbf{b} = \boldsymbol{\beta}$

For $\mathbf{b}_0$ and $\mathbf{b}$ converging to $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}$, we can show that $\mathbf{D}_0(\mathbf{b}_0) = \tilde{\mathbf{D}}_0(\mathbf{b}_0) + o_p(||\mathbf{b}_0 - \boldsymbol{\beta}_0|| + n^{-1/2})$ and $\mathbf{D}_1^*(\mathbf{b}) = \tilde{\mathbf{D}}_1^*(\mathbf{b}) + o_p(||\mathbf{b} - \boldsymbol{\beta}|| + n^{-1/2})$, respectively, where

$$\tilde{\mathbf{D}}_0(\mathbf{b}_0) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_0)(\mathbf{a}_2 - \mathbf{a}_1)\frac{O_{L_0}(t,s)}{\hat{G}_0(t \wedge L_0)}\hat{F}_0(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b}_0)\hat{H}(\mathrm{d}\mathbf{a}_2)$$
$$\tilde{\mathbf{D}}_1^*(\mathbf{b}) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_1)(\mathbf{a}_2 - \mathbf{a}_1)\frac{O_{L_1}(t,s)}{\hat{G}_1(t \wedge L_1)}\hat{F}_1^*(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b})\hat{H}(\mathrm{d}\mathbf{a}_2).$$

Since there exist points in $\mathbf{b}$ where $\tilde{\mathbf{D}}(\mathbf{b}) = \{\tilde{\mathbf{D}}_0^T(\mathbf{b}_0), \tilde{\mathbf{D}}_1^{*T}(\mathbf{b})\}^T$ is nondifferentiable, the first-order Taylor expansion cannot be directly used. Instead, we use the generalized law of mean (Huang, 2000). Let Σ be the limit of the left and right partial derivative of $\tilde{\mathbf{D}}(\mathbf{b})$. For $\mathbf{b}$ converging to $\boldsymbol{\beta}$, we obtain that

$$\begin{aligned}\mathbf{D}(\mathbf{b}) &= \tilde{\mathbf{D}}(\mathbf{b}) + o_p(||\mathbf{b}-\boldsymbol{\beta}|| + n^{-1/2})\\ &= \mathbf{D}(\boldsymbol{\beta}) + \Sigma(\mathbf{b}-\boldsymbol{\beta}) + o_p(||\mathbf{b}-\boldsymbol{\beta}|| + n^{-1/2}).\end{aligned}$$

Thus, $\mathbf{D}(\mathbf{b})$ is asymptotically linear at $\mathbf{b} = \boldsymbol{\beta}$.

A.5. Asymptotic normality of $\hat{\boldsymbol{\beta}}$

It follows that $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is asymptotically normal with mean zero and variance $\Sigma^{-1}\Omega(\Sigma^{-1})^T$, which can be consistently estimated by $\hat{\Sigma}^{-1}\hat{\Omega}(\hat{\Sigma}^{-1})^T$, where $\hat{\Sigma} = \partial\tilde{\mathbf{D}}(\hat{\boldsymbol{\beta}})/\partial\mathbf{b}$ is consistent for Σ.

A.6. Efficiency of $\hat{\boldsymbol{\beta}}$

To examine the efficiency gain of using the proposed method over Huang's method, we rewrite the estimating function in an empirical average form as follows,

$$\mathbf{D}(\boldsymbol{\beta}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\boldsymbol{\phi}(X_i, Y_{ij}, \Delta_{i0}, \Delta_{ij}, \mathbf{A}_i; \boldsymbol{\beta}) + o_p(n^{-1/2}),$$

where $\boldsymbol{\phi}(X_i, Y_{ij}, \Delta_{i0}, \Delta_{ij}, \mathbf{A}_i; \boldsymbol{\beta}) = \{\boldsymbol{\phi}_0^T(X_i, \Delta_{i0}, \mathbf{A}_i; \boldsymbol{\beta}_0), \boldsymbol{\phi}_1^T(X_i, Y_{ij}, \Delta_{ij}, \mathbf{A}_i; \boldsymbol{\beta})\}^T$, in which

$$\boldsymbol{\phi}_0(X_i, \Delta_{i0}, \mathbf{A}_i; \boldsymbol{\beta}_0) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_0)(\mathbf{a}_2 - \mathbf{a}_1) \frac{O_{L_0}(t,s)}{G_0(t \wedge L_0)} \times F_0(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \boldsymbol{\beta}_0) H(\mathrm{d}\mathbf{a}_2)$$

$$\boldsymbol{\phi}_1(X_i, Y_{ij}, \Delta_{ij}, \mathbf{A}_i; \boldsymbol{\beta}) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_1)(\mathbf{a}_2 - \mathbf{a}_1) \frac{O_{L_1}(t,s)}{G_1(t \wedge L_1)} \times F_1(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \boldsymbol{\beta}) H(\mathrm{d}\mathbf{a}_2).$$

For simplicity of notation, we denote $\boldsymbol{\phi}_{ij}(\boldsymbol{\beta}) = \boldsymbol{\phi}(X_i, Y_{ij}, \Delta_{i0}, \Delta_{ij}, \mathbf{A}_i; \boldsymbol{\beta})$. The asymptotic variance of $n^{1/2}\mathbf{D}(\boldsymbol{\beta})$ is

$$\Omega = \mathrm{E}\left\{\boldsymbol{\phi}_{ij}(\boldsymbol{\beta})^{\otimes 2}\right\} - \mathrm{E}\left[\frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\left\{\boldsymbol{\phi}_{ij}(\boldsymbol{\beta}) - \frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\boldsymbol{\phi}_{ij}(\boldsymbol{\beta})\right\}^{\otimes 2}\right].$$

We note that the asymptotic variance of $n^{1/2}(\bar{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is $\Sigma^{-1}\mathrm{E}\{\boldsymbol{\phi}_{ij}(\boldsymbol{\beta})^{\otimes 2}\}(\Sigma^{-1})^T$, which is greater or equal to the asymptotic variance of $n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$, $\Sigma^{-1}\Omega\Sigma^{-1}$. The proposed estimator $\hat{\boldsymbol{\beta}}$ is more efficient than the estimator $\bar{\boldsymbol{\beta}}$ from Huang's method when there exists $m_i^* \geq 2$ for any subject i, $i = 1, \ldots, n$.

Web Appendix B. Proof under the Conditional Independent Censoring Assumption

Here we only provide detailed proofs of the asymptotic properties for $\tilde{\boldsymbol{\beta}}$

under the conditional independent censoring assumption when the covariate-specific Kaplan–Meier estimator is used for the estimation of $G(t \mid \mathbf{A})$. Similar techniques can be used for establishing the asymptotic properties when a semi-parametric regression model is used. For the arguments below, we need the regularity conditions (C1) and (C2) in Web Appendix A.1 and an additional condition, namely,

(C4) Σ^c is nonsingular.

B.1. Uniqueness and consistency of $\tilde{\beta}$

We rewrite the estimating functions (2.5) and (2.6) as

$$\mathbf{D}_0^c(\mathbf{b}_0) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_0)(\mathbf{a}_2 - \mathbf{a}_1)\frac{O_{L_0}(t,s)}{\hat{G}_0(t \mid \mathbf{a}_1)}\hat{F}_0(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b}_0)\hat{H}(\mathrm{d}\mathbf{a}_2), \tag{S.6}$$

$$\mathbf{D}_1^{c*}(\mathbf{b}) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \mathbf{b}_1)(\mathbf{a}_2 - \mathbf{a}_1)\frac{O_{L_1}(t,s)}{\hat{G}_1(t \mid \mathbf{a}_1)}\hat{F}_1^*(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b})\hat{H}(\mathrm{d}\mathbf{a}_2). \tag{S.7}$$

The covariate-specific Kaplan–Meier estimators $\hat{G}_0(t \mid \mathbf{A})$ and $\hat{G}_1(t \mid \mathbf{A})$ are continuous and compactly differentiable as well as $\hat{F}_0$, $\hat{F}_1^*$, and $\hat{H}$. Thus, it follows that $\mathbf{D}_0^c$ and $\mathbf{D}_1^{c*}$ are continuous and compactly differentiable functionals. Due to the monotonicity of the estimating functions (S.6) in $\mathbf{b}_0$ and (S.7) in $\mathbf{b}_1$ given $\mathbf{b}_0$, the solutions to $D_0^c(\mathbf{b}_0) = 0$ and $D_1^{c*}(\mathbf{b}) = 0$ are unique.

Given the uniform consistency of the Kaplan–Meier estimators and that of the empirical functions, the consistency of $\tilde{\boldsymbol{\beta}}_0$ and $\tilde{\boldsymbol{\beta}}_1$ can be shown in a manner similar to that in Web Appendix A.2.

B.2. Asymptotic normality of $\mathbf{D}^c(\boldsymbol{\beta})$

Let $\mathbf{D}^c(\mathbf{b}) \equiv \{\mathbf{D}_0^{cT}(\mathbf{b}_0), \mathbf{D}_1^{c*T}(\mathbf{b})\}^T$. By the functional delta method and the influence function approach, we show that $n^{1/2}\mathbf{D}^c(\boldsymbol{\beta})$ is asymptotically normal with mean zero and variance Ω^c. The proof is in the same line as Web Appendix A.3. For i, we derive

$$\psi_{i0}(\boldsymbol{\beta}_0) = n^{-3/2}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_0)\mathbf{A}_{ii'}\left[\frac{\Delta_{i0}O_{L_0}\{Z_{i0}, Z_{ii'0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i0}\wedge L_0 \mid \mathbf{A}_i)} - \frac{\Delta_{i'0}O_{L_0}\{Z_{i'0}, Z_{i'i0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i'0}\wedge L_0 \mid \mathbf{A}_{i'})}\right] + n^{-3/2}\int_0^{L_0}\frac{Q_0^c(t,\boldsymbol{\beta}_0)\hat{G}_0(t- \mid \mathbf{A}_i)}{Y_0(t)\hat{G}_0(t \mid \mathbf{A}_i)}\mathrm{d}\hat{M}_{i0}^c(t), \tag{S.8}$$

$$\psi_{i1}^*(\boldsymbol{\beta}) = n^{-3/2}\sum_{i'=1}^{n} w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_1)\mathbf{A}_{ii'}\left[\frac{1}{m_i^*}\sum_{j=1}^{m_i^*}\frac{\Delta_{ij}O_{L_1}\{Z_{ij}, Z_{ii'j}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{ij}\wedge L_1 \mid \mathbf{A}_i)} - \frac{1}{m_{i'}^*}\sum_{l=1}^{m_{i'}^*}\frac{\Delta_{i'l}O_{L_1}\{Z_{i'l}, Z_{i'il}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{i'l}\wedge L_1 \mid \mathbf{A}_{i'})}\right] + n^{-3/2}\int_0^{L_1}\frac{Q_1^{c*}(t,\boldsymbol{\beta})\hat{G}_1(t- \mid \mathbf{A}_i)}{Y_1^*(t)\hat{G}_1(t \mid \mathbf{A}_i)}\mathrm{d}\hat{M}_{i1}^{c*}(t), \tag{S.9}$$

in which,

$$Q_0^c(t,\boldsymbol{\beta}_0) = \sum_{i=1}^n \sum_{i'=1}^n w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_0)\mathbf{A}_{ii'} \left[\frac{\Delta_{i0} O_{L_0}\{Z_{i0}, Z_{ii'0}(\boldsymbol{\beta}_0)\}}{\hat{G}_0(Z_{i0} \wedge L_0 \mid \mathbf{A}_i)} I(Z_{i0} > t) \right],$$

$$Q_1^{c*}(t,\boldsymbol{\beta}) = \sum_{i=1}^n \sum_{i'=1}^n w(\mathbf{A}_i, \mathbf{A}_{i'}, \boldsymbol{\beta}_1)\mathbf{A}_{ii'} \left[\frac{1}{m_i^*} \sum_{j=1}^{m_i^*} \frac{\Delta_{ij} O_{L_1}\{Z_{ij}, Z_{ii'j}(\boldsymbol{\beta})\}}{\hat{G}_1(Z_{ij} \wedge L_1 \mid \mathbf{A}_i)} I(Z_{ij} > t) \right],$$

$$\hat{M}_{i0}^c(t) = I(Z_{i0} \le t, \Delta_{i0} = 0) - \int_0^t I(Z_{i0} \ge s) \mathrm{d}\hat{\Lambda}_0(s \mid \mathbf{A}_i),$$

$$\hat{M}_{i1}^{c*}(t) = \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} I(Z_{ij} \le t, \Delta_{ij} = 0) - \int_0^t \frac{1}{m_i^*} \sum_{j=1}^{m_i^*} I(Z_{ij} \ge s) \mathrm{d}\hat{\Lambda}_1(s \mid \mathbf{A}_i)$$

and $\hat{\Lambda}_k(t \mid \mathbf{A})$ is the Nelson–Aalen estimator corresponding to $\hat{G}_k(t \mid \mathbf{A})$ for $k = 0, 1$. The last terms on the right-hand side of equations (S.8) and (S.9) result from the large sample properties of $\hat{G}_0(t \mid \mathbf{A})$ and $\hat{G}_1(t \mid \mathbf{A})$. The variance Ω^c can be consistently estimated by $\hat{\Omega}^c = \sum_{i=1}^n \{\psi_{i0}^T(\tilde{\boldsymbol{\beta}}_0), \psi_{i1}^{*T}(\tilde{\boldsymbol{\beta}})\}^T \{\psi_{i0}^T(\tilde{\boldsymbol{\beta}}_0), \psi_{i1}^{*T}(\tilde{\boldsymbol{\beta}})\}$.

B.3. Asymptotic linearity of $\mathbf{D}^c(\mathbf{b})$ at $\mathbf{b} = \boldsymbol{\beta}$

We define

$$\tilde{\mathbf{D}}_0^c(\mathbf{b}_0) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_0)(\mathbf{a}_2 - \mathbf{a}_1) \frac{O_{L_0}(t,s)}{\hat{G}_0(t \mid \mathbf{a}_1)} \hat{F}_0(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b}_0) \hat{H}(\mathrm{d}\mathbf{a}_2)$$

$$\tilde{\mathbf{D}}_1^{c*}(\mathbf{b}) = \int_{t,s,\mathbf{a}_1,\mathbf{a}_2} w(\mathbf{a}_1, \mathbf{a}_2, \boldsymbol{\beta}_1)(\mathbf{a}_2 - \mathbf{a}_1) \frac{O_{L_1}(t,s)}{\hat{G}_1(t \mid \mathbf{a}_1)} \hat{F}_1^*(\mathrm{d}t, \mathrm{d}s, \mathrm{d}\mathbf{a}_1; \mathbf{a}_2, \mathbf{b}) \hat{H}(\mathrm{d}\mathbf{a}_2).$$

Let Σ^c be the limit of the left and right partial derivative of $\tilde{\mathbf{D}}^c(\mathbf{b})$. We can prove the linearity of $\mathbf{D}^c(\mathbf{b})$ in a similar way as Web Appendix A.4. Thus, we omit the details and present the main result. By the generalized law of mean, for $\mathbf{b}$ converging to $\boldsymbol{\beta}$, we obtain that

$$\mathbf{D}^c(\mathbf{b}) = \mathbf{D}^c(\boldsymbol{\beta}) + \Sigma^c(\mathbf{b} - \boldsymbol{\beta}) + o_p(||\mathbf{b} - \boldsymbol{\beta}|| + n^{-1/2}).$$

Thus, $\mathbf{D}^c(\mathbf{b})$ is asymptotically linear at $\mathbf{b} = \boldsymbol{\beta}$.

B.4. Asymptotic normality of $\tilde{\boldsymbol{\beta}}$

The asymptotic normality and linearity of $\mathbf{D}^c(\boldsymbol{\beta})$ yield that $n^{1/2}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$ is asymptotically normal with mean zero and variance $(\Sigma^c)^{-1}\Omega^c\{(\Sigma^c)^{-1}\}^T$, which can be consistently estimated by $(\hat{\Sigma}^c)^{-1}\hat{\Omega}^c\{(\hat{\Sigma}^c)^{-1}\}^T$, where $\hat{\Sigma}^c = \partial\tilde{\mathbf{D}}^c(\tilde{\boldsymbol{\beta}})/\partial\mathbf{b}$ is a consistent estimator of Σ^c.

Web Table S1. Summary of baseline characteristics

Table S1: Summary of patient- and transplant-related characteristics.

	No. Patients (%) / Median (Range)		
Variables	All Patients	Children (Age < 18)	Adults (Age ≥ 18)
N	516	155	361
Age at TX	36.9 (0.5−71.4)	9.4 (0.5−17.9)	47.4 (18.1−71.4)
Gender			
Male	304 (59)	100 (65)	204 (57)
Female	212 (41)	55 (35)	157 (43)
Diagnosis			
ALL	131 (25)	67 (43)	64 (18)
AML	217 (42)	63 (41)	154 (43)
CML	19 (4)	1 (1)	18 (5)
Hodgkin's Lymphoma	7 (1)	1 (1)	6 (2)
Multiple Myeloma	1 (0)	0 (0)	1 (0)
Myelodysplastic Syndrome	45 (9)	9 (6)	36 (10)
Myeloproliferative Neoplasm	10 (2)	0 (0)	10 (3)
Neuroblastoma	1 (0)	1 (1)	0 (0)
Non-Hodgkin's Lymphoma	59 (11)	6 (4)	53 (15)
Other Leukemia	21 (4)	7 (5)	14 (4)
Other Malignancy	5 (1)	0 (0)	5 (1)
CMV Serostatus			
Positive	301 (58)	100 (65)	201 (56)
Negative	215 (41)	55 (35)	160 (44)
Type of Transplant			
Double Cord	374 (72)	60 (39)	314 (87)
Single Cord	142 (28)	95 (61)	47 (13)
Conditioning Regimen			
Myeloablative	281 (54)	150 (97)	131 (36)
Non-Myeloablative w ATG	67 (13)	0 (0)	67 (19)
Non-Myeloablative wo ATG	168 (33)	5 (3)	163 (45)
HLA Locus Matching Score			
4/6	262 (51)	44 (28)	218 (60)
5/6	202 (39)	86 (55)	116 (32)
6/6	52 (10)	25 (16)	27 (7)
GVHD Prophylaxis			
CSA/MMF/MTX	449 (87)	104 (67)	344 (95)
Other	67 (13)	51 (33)	16 (4)
CD34+ graft infused ($\times 10^6$/kg)	0.49 (0.06−27.53)	0.58 (0.06−8.42)	0.47 (0.07−27.53)
Low	130 (25)	35 (23)	95 (26)
High	386 (75)	120 (77)	266 (74)
TNC dose infused ($\times 10^8$/kg)	0.38 (0.11−4.89)	0.48 (0.15−2.27)	0.36 (0.11−4.89)
Low	139 (27)	29 (19)	110 (30)
High	377 (73)	126 (81)	251 (70)

Abbreviations: TX=transplant; ALL=acute lymphoblastic leukemia; AML=acute myeloblastic leukemia; CML=chronic myeloid leukemia; CMV=cytomegalovirus; ATG=anti-thymocyte globulin; HLA=human leukocyte antigen; GVHD=graft-versus-host disease; CSA=cyclosporin; MMF=mychophenolate mofetil; MTX=methotrexate; TNC=total nucleated cell; High: dose $> 1^{st}$ quartile; low: dose $\leq 1^{st}$ quartile.

References

Gill, R. D. and Johansen, S. (1990). A survey of product-integration with a view towards application in survival analysis. *Ann. Statist.* **18**, 1501-1555.

Huang, Y. (2000). Two-sample multistate accelerated sojourn times model. *J. Amer. Statist. Assoc.* **95**, 619-627.

Huang, Y. (2002). Censored regression with the multistate accelerated sojourn times model. *J. Roy. Stat. Soc. Ser. B* **64**, 17-29.

Pollard, D. (1984). *Convergence of Stochastic Processes.* New York: Springer.