Double inverse-weighted estimation of cumulative treatment effects under non-proportional hazards and dependent censoring

Web Appendix

Douglas E. Schaubel

Department of Biostatistics, University of Michigan, Ann Arbor, MI, 48109-2029, U.S.A.

deschau@umich.edu

and

Guanghui Wei

Amgen Inc., South San Francisco, CA, 94080, U.S.A.

gwei@amgen.com
Supplementary Materials

The first part of this Web Appendix contains proofs of Theorems 1 to 4. The second part consists of the supplementary simulation results and additional analyses pertaining to Section 5.

Proof of Theorem 1

We begin by demonstrating the consistency of

\[
\hat{\Lambda}_j(t) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t \frac{\hat{w}_i^C(s, \hat{\theta})w_{ij}^G(\hat{\beta})}{n^{-1} \sum_{i=1}^{n} Y_{ij}(s)\hat{w}_i^C(s, \hat{\theta})w_{ij}^G(\beta)} dN_{ij}^D(s).
\]

Since \( \hat{w}_i^C(s, \hat{\theta}) \xrightarrow{a.s.} w_i^C(s, \theta_0) \) and \( w_{ij}^G(\hat{\beta}) \xrightarrow{a.s.} w_{ij}^G(\beta_0) \), \( n^{-1} \sum_{i=1}^{n} Y_{ij}(s)\hat{w}_i^C(s, \hat{\theta})w_{ij}^G(\hat{\beta}) \) converges almost surely to \( E \left[ Y_{ij}(s)w_i^C(s, \theta_0)w_{ij}^G(\beta_0) \right] \) by the Strong Law of Large Numbers (SLLN, Pollard, 1990). Let \( \tilde{Z}_i(0, t] = \{Z_i(s); s \in (0, t]\} \), such that \( \tilde{Z}_i(t) = \{Z_i(0), Z_i(0, t]\} \).

We obtain that

\[
E \left[ Y_{ij}(s)w_i^C(s, \theta_0)w_{ij}^G(\beta_0) \right] = E \left[ I(X_i > s, G_i = j) \exp\{\Lambda_i^C(s)\}p_{ij}(\beta_0)^{-1} \right] = E_{\tilde{Z}_i(s)} \left[ P \{X_i > s, G_i = j|\tilde{Z}_i(s)\} \exp\{\Lambda_i^C(s)\}p_{ij}(\beta_0)^{-1} \right] = E_{\tilde{Z}_i(s)} \left[ P \{X_i > s|G_i = j, \tilde{Z}_i(s)\}P \{G_i = j|\tilde{Z}_i(s)\} \exp\{\Lambda_i^C(s)\}p_{ij}(\beta_0)^{-1} \right].
\]

Using successive conditioning, we can write

\[
P \{X_i > s|G_i = j, \tilde{Z}_i(s)\} \exp\{\Lambda_i^C(s)\} = \prod_{r \in (0, s]} P \{X_i > r|X_i > r, G_i = j, \tilde{Z}_i(r)\} \{1 - d\Lambda_i^C(r)\}^{-1}
\]

\[
= \prod_{r \in (0, s]} P \{D_i > r|X_i > r, G_i = j, \tilde{Z}_i(r)\} P \{C_i > r|X_i > r, G_i = j, \tilde{Z}_i(r)\} \{1 - d\Lambda_i^C(r)\}^{-1}
\]

\[
= \prod_{r \in (0, s]} P \{D_i > r|D_i > r, G_i = j, \tilde{Z}_i(r)\}
\]

\[
= P \{D_i > s|G_i = j, \tilde{Z}_i(s)\}.
\]
Combining (1) and (2), we have that

\[
E \left[ Y_{ij}(s) w_i^C(s; \theta_0) w_j^G(\beta_0) \right] \\
= E_{Z_i(s)} \left[ P\{G_i = j| \tilde{Z}_i(s)\} P\{D_i > s| G_i = j, \tilde{Z}_i(s)\} p_{ij}(\beta_0)^{-1} \right] \\
= \int_{Z_i(s)} P\{G_i = j| \tilde{Z}_i(s)\} P\{D_i > s| G_i = j, \tilde{Z}_i(s)\} p_{ij}(\beta_0)^{-1} dF\{\tilde{Z}_i(s)\} \\
= \int_{Z_i(s)} p_{ij}(\beta_0)^{-1} P\{D_i > s, G_i = j, \tilde{Z}_i(s)\} d\tilde{Z}_i(s) \\
= \int_{Z_i(0)} p_{ij}(\beta_0)^{-1} \int_{Z_i(0,s]} P\{D_i > s, G_i = j, Z_i(0), Z_i(0,s]\} dZ_i(0) dZ_i(0,s] \\
= \int_{Z_i(0)} p_{ij}(\beta_0)^{-1} P\{D_i > s, G_i = j, Z_i(0)\} dZ_i(0) \\
= \int_{Z_i(0)} P\{D_i > s| G_i = j, Z_i(0)\} dF\{Z_i(0)\} \\
= E_{Z_i(0)} \left[ P\{D_i > s| G_i = j, Z_i(0)\} \right] \\
\equiv S_j(s).
\]

Similarly, it can be shown that \( n^{-1} \sum_{i=1}^{n} \tilde{w}_i^C(s, \hat{\theta}) w_{ij}^G(\hat{\beta}) dN_i^D(s) \) converges almost surely to \( dF_j(s) \). Combining these results and using continuity, we obtain that \( \hat{\Lambda}_j(t) \) converges almost surely to \( \int_0^t S_j(s)^{-1} dF_j(s) \equiv \Lambda_j(t) \) uniformly for \( t \in [0, \tau] \).

With respect to asymptotic normality, we begin with the decomposition,

\[
n^{1/2}\{\hat{\Lambda}_j(t) - \Lambda_j(t)\} = \hat{\alpha}_{j1}(t) + \hat{\alpha}_{j2}(t) + \hat{\alpha}_{j3}(t) + \hat{\alpha}_{j4}(t),
\]

where

\[
\hat{\alpha}_{j1}(t) = n^{1/2}\{\hat{\Lambda}_j(t, \hat{\beta}, \hat{\theta}, \Lambda_0^C) - \hat{\Lambda}_j(t, \beta_0, \hat{\theta}, \Lambda_0^C)\} \\
\hat{\alpha}_{j2}(t) = n^{1/2}\{\hat{\Lambda}_j(t, \beta_0, \hat{\theta}, \Lambda_0^C) - \hat{\Lambda}_j(t, \beta_0, \theta_0, \Lambda_0^C)\} \\
\hat{\alpha}_{j3}(t) = n^{1/2}\{\hat{\Lambda}_j(t, \beta_0, \theta_0, \Lambda_0^C) - \hat{\Lambda}_j(t, \beta_0, \theta_0, \Lambda_0^C)\} \\
\hat{\alpha}_{j4}(t) = \{\hat{\Lambda}_j(t, \beta_0, \theta_0, \Lambda_0^C) - \Lambda_j(t)\},
\]
for \( j = 0, \cdots, J \), with \( \widehat{\Lambda}_j(t, \widehat{\beta}, \widehat{\theta}, \widehat{\Lambda}_0^C) = \widetilde{\Lambda}_j(t) \) and
\[
\widetilde{\Lambda}_j(t, \beta, \theta, \Lambda_0^C) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t \frac{w_{ij}^G(\beta) \tilde{w}_i^C(s; \theta)}{n-1} \frac{w_{ij}^G(\beta_0) \tilde{w}_i^C(s; \theta)}{n-1} dN_{ij}^D(s)
\]
\[
\tilde{w}^C_i(s; \theta) = Y_{ij}(s) \exp \left\{ \int_0^s \exp \left\{ \beta^T Z_s(u) \right\} d\Lambda_i^C(u; \theta) \right\}
\]
\[
\Lambda_j(t, \beta, \theta, \Lambda_0^C) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t \frac{w_{ij}^G(\beta) w_i^C(s; \theta)}{n-1} \frac{w_{ij}^G(\beta_0) w_i^C(s; \theta)}{n-1} dN_{ij}(s)
\]
\[
w_i^C(s; \theta) = Y_{ij}(s) \exp \{ \Lambda_i^C(s; \theta_0) \}.
\]

We can express \( \widehat{\alpha}_{j1}(t) \) as follows,
\[
\widehat{\alpha}_{j1}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ \frac{w_{ij}^G(\beta)}{R_j^0(s; \beta, \theta)} - \frac{w_{ij}^G(\beta_0)}{R_j^0(s; \beta_0, \theta)} \right\} \tilde{w}_i^C(s; \theta) dN_{ij}^D(s),
\]
such that \( \widehat{\alpha}_{j1}(t) = \hat{\alpha}_{j11}(t) + \hat{\alpha}_{j12}(t) \), where
\[
\hat{\alpha}_{j11}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ \frac{w_{ij}^G(\beta) - w_{ij}^G(\beta_0)}{R_j^0(s; \beta, \theta)} \right\} \tilde{w}_i^C(s; \theta) dN_{ij}^D(s)
\]
(3)
\[
\hat{\alpha}_{j12}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t w_{ij}^G(\beta_0) \left\{ R_j^0(s; \beta, \theta)^{-1} - R_j^0(s; \beta_0, \theta)^{-1} \right\} \tilde{w}_i^C(s; \theta) dN_{ij}^D(s).
\]
(4)

With respect to \( \hat{\alpha}_{j11}(t) \), by a linear Taylor series expansion,
\[
n^{1/2} \{ w_{ij}^G(\beta) - w_{ij}^G(\beta_0) \} = a_{ij}(\beta_0)(\beta - \beta_0) + o_p(1)
\]
(5)
where
\[
a_{ij}(\beta_0) = \left. \frac{\partial w_{ij}^G(\beta)}{\partial \beta} \right|_{\beta_0}
\]
\[
= (1 - G_{i0}) \left[ \sum_{k=1}^{J} \frac{\exp \{ \beta^T X_{ik} \} X_{ik}}{\exp \{ \beta^T X_{ij} \}} X_{ij} \right] - X_{ij} p_{ij}^{-1}(\beta_0) + G_{i0} \sum_{k=1}^{J} \exp \{ \beta^T X_{ik} \} X_{ik}.
\]

From standard maximum likelihood theory,
\[
n^{1/2}(\beta - \beta_0) = \Omega_G^{-1}(\beta_0) n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1)
\]
(6)
\[
\Omega_G(\beta) = E \left[ \sum_{j=1}^{J} p_{ij}(\beta) X_{ij} \left\{ X_{ij}^T - \sum_{k=1}^{J} X_{ik}^T p_{ik}(\beta) \right\} \right]
\]
\[
\psi_i^G(\beta) = \sum_{j=1}^{J} X_{ij} [G_{ij} - p_{ij}(\beta)].
\]
Combining this result with (5),

\[ n^{1/2} \{ w_{ij}^G(\hat{\beta}) - w_{ij}^G(\beta_0) \} = a_{ij}^T(\beta_0)\Omega_G^{-1}(\beta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1). \]  (7)

Using (7), then applying the SLLN and continuity, when \( n \to \infty \), we can re-express \( \hat{\alpha}_{j1}(t) \) as

\[ \hat{\alpha}_{j1}(t) = h_j^T(t)\Omega_G^{-1}(\beta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1) \]  (8)

\[ h_j(t) = E \left[ \int_0^t \frac{w_i^C(s, \theta_0)a_{ij}(\beta_0)}{r_j^{(0)}(s, \beta_0, \theta_0)} dN_{ij}(s) \right]. \] (9)

With respect to \( \hat{\alpha}_{j2}(t) \), combining a Taylor expansion with (6),

\[ n^{1/2} \left\{ R_j^{(0)}(s; \hat{\beta}, \hat{\theta})^{-1} - R_j^{(0)}(s; \beta_0, \theta) \right\} = -\frac{R_j^{(\beta)}(s; \hat{\beta}, \hat{\theta})}{R_j^{(0)}(s; \hat{\beta}, \theta)^2} \Omega_G^{-1}(\beta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1), \]  (9)

where we define

\[ R_j^{(\beta)}(s; \beta, \theta) = \frac{\partial}{\partial \beta} R_j^{(0)}(s; \beta, \theta) = \frac{1}{n} \sum_{i=1}^{n} Y_{ij}(t)w_i^G(\beta)\hat{w}_i^C(s; \theta)a_{ij}(\beta), \]  (10)

which converges almost surely to

\[ r_j^{(\beta)}(s; \beta, \theta) = E[Y_{ij}(t)w_i^G(\beta)w_i^C(s; \theta)a_{ij}(\beta)]. \]

Combining (4) and (6) with (9), then applying the SLLN and continuity,

\[ \hat{\alpha}_{j2}(t) = d_j^T(t)\Omega_G^{-1}(\beta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1) \]

\[ d_j(t) = E \left[ \int_0^t \frac{r_j^{(\beta)}(s; \beta, \theta)}{r_j^{(0)}(s; \beta, \theta)} d\Lambda_j(s) \right]. \]  (11)

Combining (8) and (11),

\[ \hat{\alpha}_{j1}(t) = \{ h_j(t) + d_j(t) \}^T\Omega_G^{-1}(\beta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^G(\beta_0) + o_p(1). \]  (12)
We can write \( \hat{\alpha}_{j2}(t) = \hat{\alpha}_{j21}(t) + \hat{\alpha}_{j21}(t) \), where
\[
\hat{\alpha}_{j21}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t w_i^{C_j}(\beta_0)R_j^0(s; \beta, \theta_0)^{-1}\{\hat{w}_i^C(s; \hat{\theta}) - \hat{w}_i^C(s; \theta_0)\}dN^D_i(s)
\]
\[
\hat{\alpha}_{j22}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t w_i^{C_j}(\beta_0)w_i^C(s; \theta_0)\{R_j^0(s; \beta_0, \hat{\theta})^{-1} - R_j^0(s; \beta_0, \theta_0)^{-1}\}dN^D_i(s).
\]
Through standard partial likelihood theory (Fleming and Harrington, 1991),
\[
n^{1/2}(\hat{\theta} - \theta_0) = \Omega^{-1}_C(\theta_0)n^{-1/2}\sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1) \tag{13}
\]
\[
\Omega_C(\theta) = \int_0^t v_C(t, \theta)r_C^0(t, \theta)d\Lambda^C_0(t),
\]
\[
\psi_i^C(\theta_0) = \int_{0}^{T} \{Z_i^C(t) - z_C(t, \theta_0)\}dM^C_i(t).
\]
Using a Taylor expansion, (13), the SLLN and continuity,
\[
n^{1/2}[\hat{w}_i^C(s, \hat{\theta}) - \hat{w}_i^C(s, \theta_0)] = w_i^C(s, \theta_0)b_i^T(s, \theta_0)\Omega^{-1}_C(\theta_0)n^{-1/2}\sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1)
\]
\[
b_i(s, \theta_0) = \int_{0}^{T} \exp\{\theta_0^TZ_i^C(u)\}\{Z_i^C(u) - z_C(u)\}d\Lambda^C_0(u).
\]
Substituting this expression into \( \hat{\alpha}_{j21}(t) \), then applying the SLLN,
\[
\hat{\alpha}_{j21}(t) = g_j^T(t)\Omega^{-1}_C(\theta_0)n^{-1/2}\sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1) \tag{14}
\]
\[
g_j(t) = E \left[ \int_0^t w_i^{C_j}(\beta_0)w_i^C(s; \theta_0)r_j^0(s; \beta_0, \theta_0)^{-1}b_i(s)dN^D_i(s) \right].
\]
With respect to \( \hat{\alpha}_{j22}(t) \), through another Taylor expansion,
\[
\{R_j^0(s; \beta_0, \theta)\}^{-1} - R_j^0(s; \beta_0, \theta_0)^{-1}
\]
\[
= - \frac{R_j^0(s; \beta_0, \theta_0)^T}{R_j^0(s; \beta_0, \theta_0)^2} \Omega^{-1}_C(\theta_0)n^{-1/2}\sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1), \tag{15}
\]
where we define
\[
R_j^{(\theta)}(s; \beta, \theta) = \frac{\partial}{\partial \theta}R_j^0(s; \beta, \theta) = \frac{1}{n} \sum_{i=1}^{n} Y_{ij}(s)w_i^C(\beta)\hat{w}_i^C(s; \theta)b_i(s), \tag{16}
\]
which converges almost surely to
\[
r_j^{(\theta)}(s; \beta, \theta) = E[Y_{ij}(s)w_i^C(\beta)w_i^C(s; \theta)b_i(s)].
\]
Substituting (15) into \( \hat{\alpha}_{j22}(t) \), then again applying the SLLN and using continuity,

\[
\hat{\alpha}_{j22}(t) = \mathbf{f}_j^T(t) \Omega_C^{-1}(\theta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1) \tag{17}
\]

\[
f_j(t) = \int_0^t \frac{r_j^{(0)}(s; \beta_0, \theta_0)}{R_j^{(0)}(s; \beta_0, \theta_0)} d\Lambda_j(s).
\]

Combining (14) and (17),

\[
\hat{\alpha}_{j2}(t) = \{g_j(t) + f_j(t)\}^T \Omega_C^{-1}(\theta_0)n^{-1/2} \sum_{i=1}^{n} \psi_i^C(\theta_0) + o_p(1). \tag{18}
\]

Regarding \( \hat{\alpha}_{j3}(t) \), we can write

\[
\hat{\alpha}_{j3}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \left\{ \hat{w}_i^C(s, \theta_0) - w_i^C(s, \theta_0) \right\} \frac{w_{ij}^C(\beta_0)}{R_{ij}^{(0)}(s; \beta_0, \theta_0)} dN_{ij}^D(s). \tag{19}
\]

Applying the Functional Delta Method,

\[
n^{1/2} \left\{ \hat{w}_i^C(s, \theta_0) - w_i^C(s, \theta_0) \right\} = w_i^C(s, \theta_0)n^{1/2} \left\{ \hat{\Lambda}_i^C(s, \theta_0) - \Lambda_i^C(s) \right\} = w_i^C(s, \theta_0)n^{-1/2} \sum_{l=1}^{n} \int_0^s \exp\left\{ \theta_0^T Z_i^C(u) \right\} \frac{1}{R_C^l(u, \theta_0)} dM_l(u) + o_p(1).
\]

Subbing this expression into (19), changing the orders of integration and summation, then applying the SLLN,

\[
\hat{\alpha}_{j3}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t q_j(s, t) r_j^{(0)}(s; \beta_0, \theta_0)^{-1} dM_i^C(s) \tag{20}
\]

\[
q_j(s, t) = E \left[ \exp\left\{ \theta_0^T Z_i^C(u) \right\} \int_u^t \frac{w_{ij}^G(\beta_0)w_i^C(s; \theta_0)}{r_j^{(0)}(s; \beta_0, \theta_0)} dN_{ij}^D(s) \right]. \tag{21}
\]

Finally, the quantity \( \hat{\alpha}_{j4}(t) \) can be written asymptotically as

\[
\hat{\alpha}_{j4}(t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t \frac{w_{ij}^G(\beta_0)w_i^C(s, \theta_0)}{r_j(s, \beta_0, \theta_0)} dM_i^D(s). \tag{22}
\]

Combining (12), (18), (21) and (22),

\[
n^{1/2} \{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} = n^{-1/2} \sum_{i=1}^{n} \Phi_{ij}(t) + o_p(1), \tag{23}
\]
where \( \Phi_{ij}(t) \) is as defined in Theorem 1; i.e.,

\[
\Phi_{ij}(t) = \left\{ h_j(t) + d_j(t) \right\}^T \Omega_C^{-1}(\beta_0) n^{-1/2} \Psi_i^C(\beta_0) \\
+ \left\{ g_j(t) + f_j(t) \right\}^T \Omega_C^{-1}(\theta_0) n^{-1/2} \Psi_i^C(\theta_0) \\
+ \int_0^t q_j(s,t) r_j^{(0)}(s; \beta_0, \theta_0)^{-1} dM_i^C(s) \\
+ \int_0^t w_i^C(\beta_0) w_i^C(s, \theta_0) r_j(s; \beta_0, \theta_0) dM_{ij}^D(s).
\]

As \( n \to \infty \), \( n^{1/2}\{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} \) behaves like a sum of independent and identically distributed mean 0 random variates. Therefore, by the multivariate central limit theorem, for any finite set of (say \( k \)) time points, the vector \( n^{1/2} \left[ \{ \hat{\Lambda}_j(t_1) - \Lambda_j(t_1) \}, \ldots, \{ \hat{\Lambda}_j(t_k) - \Lambda_j(t_k) \} \right] \) converges to a mean zero multivariate normal distribution. Further, it can be shown that \( n^{1/2}\{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} \) is tight (Pollard, 1990; van der Vaart and Wellner, 1996) using techniques in Bilias, Gu and Ying (1997); such that \( n^{1/2}\{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} \) converges to a mean 0 Gaussian process with covariance function \( E\{ \Phi_{ij}(s, \beta_0, \theta_0) \Phi_{ij}(t, \beta_0, \theta_0) \} \) for any pair of time points \( (s, t) \in [0, t_U] \times [0, t_U] \).

**Proof of Theorem 2**

As demonstrated in Theorem 1, \( \hat{\Lambda}_j(t) \) converges almost surely to \( \Lambda_j(t) \). Therefore, by the continuity of \( \phi_j(t) \), \( \hat{\phi}_j(t) = \hat{\Lambda}_j(t)/\hat{\Lambda}_0(t) \) converges almost surely to \( \phi_j(t) \), uniformly in \( [t_L, t_U] \).

With respect to asymptotic normality, one can write

\[
n^{1/2}\{ \hat{\phi}_j(t) - \phi_j(t) \} = \hat{\Lambda}_0(t)^{-1} n^{1/2}\{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} + \hat{\Lambda}_j(t) n^{1/2}\{ \hat{\Lambda}_0(t)^{-1} - \Lambda_0(t)^{-1} \}. \tag{24}
\]

Using (23) and the strong consistency of \( \hat{\Lambda}_0(t) \),

\[
\hat{\Lambda}_0(t)^{-1} n^{1/2}\{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} = \Lambda_0(t)^{-1} n^{-1/2} \sum_{i=1}^{n} \Phi_{ij}(t) + o_p(1). \tag{25}
\]

By the strong consistency of \( \hat{\Lambda}_0(t) \), the Functional Delta Method and (23),

\[
\hat{\Lambda}_j(t) n^{1/2}\{ \hat{\Lambda}_0(t)^{-1} - \Lambda_0(t)^{-1} \} = -\Lambda_j(t) \Lambda_0(t)^{-2} n^{-1/2} \sum_{i=1}^{n} \Phi_{ij}(t) + o_p(1). \tag{26}
\]
Combining (25) and (26),
\[
    n^{1/2}\{\hat{\phi}_j(t) - \phi_j(t)\} = n^{-1/2} \sum_{i=1}^{n} \xi^\phi_{ij}(t) + o_p(1)
\]
\[
    \xi^\phi_{ij}(t) = \Lambda_0(t)^{-1} \Phi_{ij}(t) - \Lambda_j(t) \Lambda_0(t)^{-2} \Phi_{i0}(t) + o_p(1),
\]
a sum of independent and identically distributed mean 0 variates, asymptotically. Using arguments similar to those at the end of the Proof of Theorem 1, \(n^{1/2}\{\hat{\phi}_j(t) - \phi_j(t)\}\) converges to a zero-mean Gaussian process with covariance function \(\sigma_j^\phi(s, t) = E\{\xi^\phi_{ij}(s)\xi^\phi_{ij}(t)\}\), for \((s, t) \in [t_L, t_U] \times [t_L, t_U].\)

**Proof of Theorem 3**

Using the almost sure convergence of \(\hat{\Lambda}_j(t)\) to \(\Lambda_j(t)\) and, through continuity, of \(\hat{S}_j(t)\) to \(S_j(t)\), \(\hat{R}\hat{R}_j(t)\) converges almost surely to \(RR_j(t)\) uniformly in \(t \in [t_L, t_U]\). Using the Functional Delta method,
\[
    n^{1/2}\{\hat{R}\hat{R}_j(t) - RR_j(t)\} = \frac{S_j(t)}{F_0(t)} n^{1/2}\{\hat{\Lambda}_j(t) - \Lambda_j(t)\} - \frac{F_j(t) S_0(t)}{F_0(t)^2} n^{1/2}\{\hat{\Lambda}_0(t) - \Lambda_0(t)\} + o_p(1).
\]
Applying this result and (23),
\[
    n^{1/2}\{\hat{R}\hat{R}_j(t) - RR_j(t)\} = \frac{S_j(t)}{F_0(t)} n^{-1/2} \sum_{i=1}^{n} \Phi_{ij}(t) - \frac{F_j(t) S_0(t)}{F_0(t)^2} \sum_{i=1}^{n} \Phi_{i0}(t) + o_p(1). \number{28}
\]
Therefore, we can write
\[
    n^{1/2}\{\hat{R}\hat{R}_j(t) - RR_j(t)\} = n^{-1/2} \sum_{i=1}^{n} \xi^R_{ij}(t) + o_p(1)
\]
\[
    \xi^R_{ij}(t) = \frac{S_j(t)}{F_0(t)} \Phi_{ij}(t) - \frac{F_j(t) S_0(t)}{F_0(t)^2} \Phi_{i0}(t) + o_p(1). \number{29}
\]
The remainder of the proof parallels that of Theorem 2.

**Proof of Theorem 4**

The fact that \(\hat{S}_j(t)\) converges almost surely to \(S_j(t)\) implies that \(\int_0^t \hat{S}_j(u) du \xrightarrow{a.s.} \int_0^t S_j(u) du,\) which in turn implies that \(\hat{\Delta}_j(t)\) converges almost surely to \(\Delta_j(t)\), uniformly in \(t \in [0, \tau]\). By
the Functional Delta Method, \( n^{1/2} \{ \hat{S}_j(t) - S_j(t) \} = -S_j(t)n^{1/2} \{ \hat{\Lambda}_j(t) - \Lambda_j(t) \} \). Integrating, we obtain \( n^{1/2} \{ \hat{\epsilon}_j(t) - \epsilon_j(t) \} = -n^{1/2} \int_0^t S_j(u)\{ \hat{\Lambda}_j(u) - \Lambda_j(u) \} du \). Switching the order of integration yields

\[
n^{1/2} \{ \hat{\epsilon}_j(t) - \epsilon_j(t) \} = -n^{1/2} \int_0^t \{ S_j(t) - S_j(u) \} d\{ \hat{\Lambda}_j(u) - \Lambda_j(u) \}.
\]

Applying result (23),

\[
n^{1/2} \{ \hat{\epsilon}_j(t) - \epsilon_j(t) \} = -n^{-1/2} \sum_{i=1}^n \int_0^t \{ S_j(t) - S_j(u) \} d\Phi_{ij}(u) + o_p(1),
\]

(yielding a result analogous to that in Andersen et al, 1993; pp.279) such that, asymptotically,

\[
n^{1/2} \{ \hat{\Delta}_j(t) - \Delta_j(t) \} = n^{-1/2} \sum_{i=1}^n \xi_{ij}^\Delta(t)
\]

\[
\xi_{ij}^\Delta(t) = -\int_0^t \{ S_j(t) - S_j(u) \} d\Phi_{ij}(u) + \int_0^t \{ S_0(t) - S_0(u) \} d\Phi_{0j}(u).
\]

Convergence to a Normal distribution, for a finite set of time points, and convergence to a Gaussian process both follow from arguments analogous to those used in the proofs of Theorems 1-3.

In Web Table 1, we present simulation results for the set-up described in Section 4, but with the sample size increased from \( n = 200 \) to \( n = 500 \). Upon increasing the sample size, the bias is reduced greatly for all estimators.

Next, we turn to the analysis of the Scientific Registry of Transplant Recipients (SRTR) kidney wait list data, described in Section 5. Estimated regression parameters for the IPCW Cox model are listed in Web Table 2. The kidney transplant hazard decreased by a significant 8% per hospital admission \((p < 0.001)\), reflecting a preference for less sick transplant candidates. Increasing age was also significantly associated with decreasing transplant hazard.

Continuing on with the analysis of the SRTR data, we next examine sensitivity of each of the proposed estimators to confounding by measured covariates. Our implicit assumptions motivating this subanalysis is that the degree of confounding due to unmeasured covariates...
will be no greater than that due to measured covariates; as discussed in Sections 5-6. We examine four different estimators of each of $\phi_1(t)$ (Web Figure 1), $RR_1(t)$ (Web Figure 2) and $\Delta_1(1)$ (Web Figure 3): the proposed IPTW/IPCW estimator (denoted by the solid line); an IPTW estimator (long dashes); an IPTW estimator adjusted for age only (short dashes); and an unadjusted estimator (dotted line).
References


Figure 1. Analysis of wait list mortality by race: Various estimators of the ratio of cumulative hazard functions (Caucasian/African American), $\phi_1(t)$. 
Figure 2. Analysis of wait list mortality by race: Various estimators of the risk ratio (Caucasian/African American), $RR_1(t)$. 
Figure 3. Analysis of wait list mortality by race: Various estimators of the difference in restricted mean lifetime (Caucasian-African American), $\Delta_1(t)$. 
Table 1

Simulation results: Examination of bias at different time points ($n=500$)

<table>
<thead>
<tr>
<th>Setting</th>
<th>C%</th>
<th>$\log \phi_1(t)$</th>
<th>BIAS</th>
<th>$\log \phi_1(t)$</th>
<th>BIAS</th>
<th>$\log \phi_1(t)$</th>
<th>BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>23%</td>
<td>0</td>
<td>0.009</td>
<td>0.283</td>
<td>0.001</td>
<td>0.584</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>0</td>
<td>0.005</td>
<td>0.283</td>
<td>-0.001</td>
<td>0.584</td>
<td>-0.009</td>
</tr>
<tr>
<td>II</td>
<td>13%</td>
<td>0.496</td>
<td>0.007</td>
<td>0.778</td>
<td>0.001</td>
<td>1.076</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>0.496</td>
<td>0.007</td>
<td>0.778</td>
<td>0.010</td>
<td>1.076</td>
<td>-0.011</td>
</tr>
<tr>
<td>III</td>
<td>28%</td>
<td>0.496</td>
<td>0.006</td>
<td>0.778</td>
<td>-0.007</td>
<td>1.076</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>46%</td>
<td>0.496</td>
<td>0.019</td>
<td>0.778</td>
<td>0.018</td>
<td>1.076</td>
<td>0.007</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Setting</th>
<th>C%</th>
<th>$\log \hat{R}_1(t)$</th>
<th>BIAS</th>
<th>$\log \hat{R}_1(t)$</th>
<th>BIAS</th>
<th>$\log \hat{R}_1(t)$</th>
<th>BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>23%</td>
<td>0</td>
<td>0.009</td>
<td>0.238</td>
<td>0.001</td>
<td>0.429</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>0</td>
<td>0.004</td>
<td>0.238</td>
<td>0.001</td>
<td>0.429</td>
<td>-0.007</td>
</tr>
<tr>
<td>II</td>
<td>13%</td>
<td>0.451</td>
<td>0.007</td>
<td>0.619</td>
<td>0.001</td>
<td>0.715</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>0.451</td>
<td>0.007</td>
<td>0.619</td>
<td>0.009</td>
<td>0.715</td>
<td>-0.006</td>
</tr>
<tr>
<td>III</td>
<td>28%</td>
<td>0.451</td>
<td>0.006</td>
<td>0.619</td>
<td>-0.004</td>
<td>0.715</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>46%</td>
<td>0.451</td>
<td>0.018</td>
<td>0.619</td>
<td>0.015</td>
<td>0.715</td>
<td>0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Setting</th>
<th>C%</th>
<th>$\Delta_1(t)$</th>
<th>BIAS</th>
<th>$\Delta_1(t)$</th>
<th>BIAS</th>
<th>$\Delta_1(t)$</th>
<th>BIAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>23%</td>
<td>-0.001</td>
<td>-0.036</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.171</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>40%</td>
<td>-0.001</td>
<td>-0.036</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.171</td>
<td>0.001</td>
</tr>
<tr>
<td>II</td>
<td>13%</td>
<td>-0.042</td>
<td>0.001</td>
<td>-0.197</td>
<td>0.001</td>
<td>-0.502</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>-0.042</td>
<td>0.001</td>
<td>-0.197</td>
<td>-0.001</td>
<td>-0.502</td>
<td>0.003</td>
</tr>
<tr>
<td>III</td>
<td>28%</td>
<td>-0.042</td>
<td>0.001</td>
<td>-0.197</td>
<td>0.003</td>
<td>-0.502</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>46%</td>
<td>-0.042</td>
<td>-0.002</td>
<td>-0.197</td>
<td>-0.005</td>
<td>-0.502</td>
<td>-0.006</td>
</tr>
</tbody>
</table>
Table 2
Analysis of wait list mortality by race: Parameter estimates for censoring model

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Covariate value</th>
<th>Hazard ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospitalizations</td>
<td>0.92</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Gender (reference=Male)</td>
<td>0.95</td>
<td>0.170</td>
<td></td>
</tr>
<tr>
<td>18-29</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>30-40</td>
<td>0.85</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40-50</td>
<td>0.75</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>50-60</td>
<td>0.76</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>≥60</td>
<td>0.66</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Lung disease (reference=No)</td>
<td>0.79</td>
<td>0.168</td>
<td></td>
</tr>
<tr>
<td>0-20</td>
<td>0.90</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>20-25</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-30</td>
<td>0.89</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>30-35</td>
<td>0.94</td>
<td>0.215</td>
<td></td>
</tr>
<tr>
<td>&gt;35</td>
<td>0.76</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Glomerulonephritis</td>
<td>1.11</td>
<td>0.120</td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polycystic kidney disease</td>
<td>1.19</td>
<td>0.013</td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.99</td>
<td>0.920</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1.05</td>
<td>0.250</td>
<td></td>
</tr>
</tbody>
</table>