Web-based Supplementary Materials for Modeling Restricted Mean Survival Time under General Censoring Mechanisms

Xin Wang (vivianwx@umich.edu)
Douglas E. Schaubel (deschau@umich.edu)
Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, 48109-2029, U.S.A.

Contents

A Asymptotic Properties of The Proposed Estimator
   A.1 Notations .................................................. 2
   A.2 Model Assumptions ........................................... 2
   A.3 Regularity Conditions ..................................... 3
   A.4 Outline of Derivation ..................................... 3
   A.5 Unbiased Estimating Equation ............................. 4
   A.6 Consistency .................................................. 7
   A.7 Asymptotic Distribution .................................. 8

B Model Selection Criteria ....................................... 8

C More Results in Application Data Analysis .................... 9

List of Figures

1 Fitted RMST (L = 1 year) by MELD score for a reference patient: white, male, age=50, Region=5, year=2005, not hospitalized, not on dialysis, blood Type=O, BMI ∈ (20, 25], sodium=130 ................ 10
2 Fitted RMST (L = 5 years) by MELD score for a reference patient: white, male, age=50, Region=5, year=2005, not hospitalized, not on dialysis, blood Type=O, BMI ∈ (20, 25], sodium=130 ................ 10

List of Tables

1 Estimated covariate effects on RMST in the absence of liver transplantation (L = 12 months) .............. 11
2 Estimated covariate effects on RMST in the absence of liver transplantation (L = 60 months) .............. 12
A Asymptotic Properties of The Proposed Estimator

A.1 Notations

To begin with, we review the essential notations needed for further discussion:
- \( i \): subject index, \( i \in \{1, \ldots, n\} \)
- \( D_i \): treatment-free death time
- \( T_i \): dependent censoring time; e.g. treatment
- \( C_i \): independent censoring time; e.g. administrative censoring
- \( \tau \): end of follow up time
- \( L \): pre-specified time point of interest, \( L \leq \tau \)
- \( X_i = D_i \land T_i \land C_i \): observation time
- \( Y_i = X_i \land L \): restricted observation time by \( L \)
- \( \Delta_i = I(D_i \land L \leq T_i \land C_i) \): indicator for restricted survival time \( D_i \land L \)
- \( \Delta_i^D = I(D_i \leq T_i \land C_i) \): death indicator
- \( \Delta_i^T = I(T_i < D_i \land C_i) \): dependent censoring indicator
- \( \Delta_i^C = I(C_i < D_i \land T_i) \): independent censoring indicator
- \( Z_i^D(t) \): time-dependent covariates that predict death \( D_i \)
- \( Z_i^T(t) \): time-dependent covariates that predict dependent censoring \( T_i \)
- \( Z_i^C \): baseline covariates that predict independent censoring \( C_i \)
- \( Z_i(t) \): a covariate set that stacks \( Z_i^D(t), Z_i^T(t), Z_i^C \) together and removes redundancy
- \( Z_i(t) = \{ Z_i(u) : 0 \leq u \leq t \} \): observation history of all the covariates up to time \( t \)
- \( \lambda_i^T(t) \): hazard rate for dependent censoring \( T_i \)
- \( \lambda_i^C(t) \): hazard rate for independent censoring \( C_i \)
- \( \Lambda_i^T(t) = \int_0^t \lambda_i^T(u) du \): cumulative hazard rate for dependent censoring \( T_i \)
- \( \Lambda_i^C(t) = \int_0^t \lambda_i^C(u) du \): cumulative hazard rate for independent censoring \( C_i \)
- \( N_i^P(t) = I(T_i \leq t, \Delta_i^P = 1) \): counting process for death
- \( N_i^T(t) = I(C_i \leq t, \Delta_i^T = 1) \): counting process for dependent censoring
- \( N_i^C(t) = I(C_i \leq t, \Delta_i^C = 1) \): counting process for independent censoring
- \( R_i(t) = I(T_i \geq t) \): at risk process
- \( dM_i^T(t) = dN_i^T(t) - R_i(t)d\Lambda_i^T(t) \): zero mean process for dependent censoring
- \( dM_i^C(t) = dN_i^C(t) - R_i(t)d\Lambda_i^C(t) \): zero mean process for independent censoring

A.2 Model Assumptions

We have made these assumptions in our paper:

(a) Assume restricted mean lifetime conditional on baseline covariates \( \mu_i(L) := E\{D_i \land L | Z_i^D(0)\} \) follows the model structure as below,
\[
g[\mu_i(L)] \equiv g \left[ E \left\{ D_i \land L | Z_i^D(0) \right\} \right] = \beta_D^T Z_i^D(0),
\]
where \( g(*) \) is a given smooth and strictly monotone link function and \( \beta_D \) is of our primary interest.

(b) Assume Cox proportional hazards model for dependent and independent censoring time \( T_i \) and \( C_i \):
\[
\lambda_i^T(t) = \lambda_0^T(t) \exp \left\{ \beta_T^T Z_i^T(t) \right\}, \quad \lambda_i^C(t) = \lambda_0^C(t) \exp \left\{ \beta_C^C Z_i^C \right\}.
\]

(c) Assume no unmeasured confounders for dependent censoring \( T_i \); for any \( t > 0 \),
\[
\lim_{h \to 0} \frac{P \left\{ X_i \in [t, t + h) , \Delta_i^T = 1 | X_i \geq t, Z_i(t) , D_i \right\}}{h} = \lim_{h \to 0} \frac{P \left\{ X_i \in [t, t + h) , \Delta_i^T = 1 | X_i \geq t, Z_i(t) \right\}}{h}.
\]

(d) Assume independent censoring time is independent of either death time or dependent censoring time given baseline covariates; i.e.,
\[
C_i \perp T_i | Z_i(0), \quad C_i \perp D_i | Z_i(0).
\]
A.3 Regularity Conditions

We specify the necessary regularity conditions (i)-(vii) as below.

(i) \{X_i, \Delta_i^D, \Delta_i^T, \Delta_i^C, \bar{Z}_i(X_i)\}, i = 1, \ldots, n are independently and identically distributed.

(ii) \(P(R_i(t) = 1) > 0\) for \(t \in (0, \tau), i = 1, \ldots, n\).

(iii) \(|Z_{ik}(0)| + \int_0^\tau d|Z_{ik}(t)| < M_Z < \infty\) for \(i = 1, \ldots, n\), where \(Z_{ik}(t)\) are the \(k\)th components of \(Z_i(t)\).

(iv) \(\Lambda_i^T(\tau) < \infty, \Lambda_i^C(\tau) < \infty\) and \(\Lambda_i^T(t), \Lambda_i^C(t)\) are absolutely continuous for \(t \in (0, \tau]\).

(v) There exist neighborhoods \(B_T\) of \(\beta_T\) and \(B_C\) of \(\beta_C\) such that for \(k = 0, 1, 2\),

\[
\sup_{t \in (0, \tau], \beta \in B_T} \left\| \frac{1}{n} \sum_{i=1}^{n} \exp \left\{ \beta^T Z_i^T(t) \right\} R_i(t) Z_i^T(t)^{\otimes k} - r_T^{(k)}(t; \beta) \right\| \xrightarrow{p} 0,
\]

\[
\sup_{t \in (0, \tau], \beta \in B_C} \left\| \frac{1}{n} \sum_{i=1}^{n} \exp \left\{ \beta^T Z_i^C(t) \right\} R_i(t) Z_i^C(t)^{\otimes k} - r_C^{(k)}(t; \beta) \right\| \xrightarrow{p} 0,
\]

where \(v^{\otimes 0} = 1, v^{\otimes 1} = v, v^{\otimes 2} = v^2 v\) and

\[
\begin{align*}
 r_T^{(k)}(t; \beta) &= E \left[ \exp \left\{ \beta^T Z_i^T(t) \right\} R_i(t) Z_i^T(t)^{\otimes k} \right], \\
 r_C^{(k)}(t; \beta) &= E \left[ \exp \left\{ \beta^T Z_i^C(t) \right\} R_i(t) Z_i^C(t)^{\otimes k} \right].
\end{align*}
\]  

(vi) Define \(h(x) = \partial g^{-1}(x)/\partial x\), where \(h\) exists and is continuous in an open neighborhood \(B_D\) of \(\beta_D\).

(vii) The matrices \(A(\beta_D), \Omega_T(\beta_T), \Omega_C(\beta_C)\) are each positive definite, where

\[
A(\beta) = E \left[ Z_i^D(0)^{\otimes 2} h \left\{ \beta_D^i Z_i^D(0) \right\} \right],
\]

\[
\Omega_T(\beta) = E \left[ \int_0^\tau \left\{ \frac{r_T^{(2)}(t; \beta)}{r_T^{(0)}(t; \beta)} - \bar{Z}_T(t; \beta)^{\otimes 2} \right\} dN_i^T(t) \right],
\]

\[
\Omega_C(\beta) = E \left[ \int_0^\tau \left\{ \frac{r_C^{(2)}(t; \beta)}{r_C^{(0)}(t; \beta)} - \bar{Z}_C(t; \beta)^{\otimes 2} \right\} dN_i^C(t) \right],
\]

and

\[
\bar{Z}_T(t; \beta) = \frac{r_T^{(1)}(t; \beta)}{r_T^{(0)}(t; \beta)},
\]

\[
\bar{Z}_C(t; \beta) = \frac{r_C^{(1)}(t; \beta)}{r_C^{(0)}(t; \beta)}.
\]

A.4 Outline of Derivation

Two estimating equation mentioned in our paper are

\[
\Phi^*(\beta) := \frac{1}{n} \sum_{i=1}^{n} \Phi_i^*(\beta) := \frac{1}{n} \sum_{i=1}^{n} \Delta_i W_i(Y_i) \left[ Y_i - g^{-1}\left\{ \beta^T Z_i^D(0) \right\} \right] Z_i^D(0) = 0,
\]

where \(W_i(t) = W_i^T(t)W_i^C(t), W_i^T(t) = \exp\{\Lambda_i^T(t)\}\) and \(W_i^C(t) = \exp\{\Lambda_i^C(t)\}\), and

\[
\Phi(\beta) := \frac{1}{n} \sum_{i=1}^{n} \Phi_i(\beta) := \frac{1}{n} \sum_{i=1}^{n} \Delta_i \hat{W}_i(Y_i) \left[ Y_i - g^{-1}\left\{ \beta^T Z_i^D(0) \right\} \right] Z_i^D(0) = 0,
\]

where \(\hat{W}_i(t) = \hat{W}_i^T(t)\hat{W}_i^C(t), \hat{W}_i^T(t) = \exp\{\hat{\Lambda}_i^T(t)\}\) and \(\hat{W}_i^C(t) = \exp\{\hat{\Lambda}_i^C(t)\}\).
We will first show (3) is unbiased, and then (9) satisfies that $\sqrt{n}\Phi(\beta_D)$ converges to a zero-mean Normal with variance $B(\beta_D) = E\{B_i(\beta_D)^{\circ2}\}$, where

$$B(\beta_D) = E\{B_i(\beta_D)^{\circ2}\}, \quad (10)$$

$$B_i(\beta) = \epsilon_i(\beta) + K_T(\beta)\Omega_T(\beta)^{-1}U_i^T(\beta) + \int_0^L H_T(u; \beta)r_T^{(0)}(u; \beta_T)^{-1}dM_i^T(u) + K_C(\beta)\Omega_C(\beta)^{-1}U_i^C(\beta_C) + \int_0^L H_C(u; \beta)r_C^{(0)}(u; \beta_C)^{-1}dM_i^C(u), \quad (11)$$

$$\epsilon_i(\beta) = \Delta_iW_i(Y_i)[Y_i - g^{-1}\{\beta'Z_i^P(0)\}]Z_i^D(0), \quad (12)$$

$$U_i^T(\beta) = \int_0^1 \{Z_i^T(u) - z_T(u; \beta_T)\}dM_i^T(u), \quad (13)$$

$$U_i^C(\beta_C) = \int_0^1 \{Z_i^C - z_C(u; \beta_C)\}dM_i^C(u), \quad (14)$$

$$K_T(\beta) = E\{\epsilon_j(\beta) D_i^T(Y_i')\}, \quad (15)$$

$$K_C(\beta) = E\{\epsilon_j(\beta) D_i^C(Y_i')\}, \quad (16)$$

$$H_T(t; \beta) = E[\epsilon_j(\beta)\exp[\beta_TZ_i^T(t)]R_i(t)], \quad (17)$$

$$H_C(t; \beta) = E[\epsilon_j(\beta)\exp[\beta_CZ_i^C(t)]R_i(t)], \quad (18)$$

$$D_i^T(t) = \int_0^1 \{Z_i^T(u) - z_T(u; \beta_T)\}d\lambda_i^T(u), \quad (19)$$

$$D_i^C(t) = \int_0^1 \{Z_i^C - z_C(u; \beta_C)\}d\lambda_i^C(u), \quad (20)$$

for any subject $i = 1, \ldots, n$, and $\Omega_T(\beta)$, $\Omega_C(\beta)$ are already defined in (4) and (5).

Let $\beta_D$ denote the solution to (9). We will show that

(a) (Consistency) as $n \to \infty$, $\hat{\beta}_D$ converges in probability to $\beta_D$.

(b) (Asymptotic Properties) as $n \to \infty$, $\sqrt{n}(\hat{\beta}_D - \beta_D)$ converges to a zero-mean Normal with variance $A(\beta_D)^{-1}B(\beta_D)A(\beta_D)^{-1}$ with $A(\beta)$ and $B(\beta)$ defined in (3) and (10).

A.5 Unbiased Estimating Equation

Theorem 1 Under regularity conditions [1]–[7], the estimating equation (3) is unbiased at the true value of $\beta_D$; i.e. $E\{\Phi^*(\beta_D)\} = 0$.

Proof 1 As defined in our paper, the $i_{th}$ error term in (3) are independently and identically distributed. It would be enough to show that $E\{\epsilon_i(\beta_D)\} = 0$. This holds because the conditional expectation on $Z_i^D(0)$ is unbiased:

$$E\{\epsilon_i(\beta_D)|Z_i^D(0)\} = Z_i^D(0)E\{W_i(Y_i)\Delta_i|Z_i^D(0)\} - Z_i^D(0)g^{-1}\{\beta_DZ_i^D(0)\}E\{W_i(Y_i)\Delta_iZ_i^D(0)\}$$

$$= Z_i^D(0)E\{W_i(Y_i)\Delta_i|Z_i^D(0)\} - Z_i^D(0)g^{-1}\{\beta_DZ_i^D(0)\}E\{W_i(Y_i)\Delta_i|Z_i^D(0)\}$$

$$= Z_i^D(0)E\left\{\frac{I(T_i \geq D_i \land L; C_i \geq D_i \land L)}{P(T_i \geq D_i \land L; C_i \geq D_i \land L)}|Z_i^D(0)\right\}Z_i^D(0)$$

$$- Z_i^D(0)g^{-1}\{\beta_DZ_i^D(0)\}E\left\{\frac{I(T_i \geq D_i \land L; C_i \geq D_i \land L)}{P(T_i \geq D_i \land L; C_i \geq D_i \land L)}|Z_i^D(0)\right\}$$

$$= Z_i^D(0)E\left\{D_i \land L|Z_i^D(0)\right\} - Z_i^D(0)g^{-1}\{\beta_DZ_i^D(0)\}$$

$$= 0.$$

Then averaging over the baseline covariates, $E\{\epsilon_i(\beta_D)\}$ and therefore $E\{\Phi^*(\beta_D)\}$ will be 0.
Theorem 2 Under regularity conditions [30]-[31], as \( n \to \infty \), \( \sqrt{n} \Phi(\beta_D) \) converges to a zero-mean Normal with variance \( B(\beta_D) \) defined in [30].

Proof 2 As shown in Zhang and Schaumbel (2011), the weight involved with dependent censoring time \( T_i \) can be written as

\[
\sqrt{n} \left\{ \hat{W}_i^T(t) - W_i^T(t) \right\} = \frac{1}{\sqrt{n}} W_i^T(t) \left\{ D_i^T(t)' \Omega_T(\beta_T)^{-1} \sum_{j=1}^{n} U_j^T(\beta_T) + \sum_{j=1}^{n} J_{ij}^T(t) \right\} + o_p(1),
\]

with defined \( D_i^T(t), U_i^T(\beta_T), \Omega_T(\beta_T), r_T^T(t), z_T(t; \beta) \) in [19], [13], [4], [1], [6] and

\[
J_{ij}^T(t) = \int_0^t \exp \left\{ \beta_T'Z_i^T(u) \right\} R_i(u) r_T^T(u; \beta_T)^{-1} dM_j^T(u).
\]

And we can derive the similar formula for independent censoring time \( C_i \),

\[
\sqrt{n} \left\{ \hat{W}_i^C(t) - W_i^C(t) \right\} = \frac{1}{\sqrt{n}} W_i^C(t) \left\{ D_i^C(t)' \Omega_C(\beta_C)^{-1} \sum_{j=1}^{n} U_j^C(\beta_C) + \sum_{j=1}^{n} J_{ij}^C(t) \right\} + o_p(1),
\]

with defined \( D_i^C(t), U_i^C(\beta_C), \Omega_C(\beta), r_T^C(t), z_C(t; \beta) \) in [20], [14], [9], [2], [6] and

\[
J_{ij}^C(t) = \int_0^t \exp \left( \beta_C'Z_i^C(u) \right) R_i(u) r_T^C(u; \beta_C)^{-1} dM_j^C(u).
\]

Rewrite the target vector as

\[
\sqrt{n} \Phi(\beta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i(Y_i)[Y_i - g^{-1}\{\beta'Z_i^D(0)\}]Z_i^D(0) \hat{W}_i^T(Y_i) \hat{W}_i^C(Y_i)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i(Y_i)[Y_i - g^{-1}\{\beta'Z_i^D(0)\}]Z_i^D(0)[W_i^T(Y_i)W_i^C(Y_i)] (21)
\]

\[
+ W_i^C(Y_i) \left\{ \hat{W}_i^T(Y_i) - W_i^T(Y_i) \right\} (22)
\]

\[
+ W_i^T(Y_i) \left\{ \hat{W}_i^C(Y_i) - W_i^C(Y_i) \right\} (23)
\]

\[
+ \left\{ \hat{W}_i^T(Y_i) - W_i^T(Y_i) \right\} \left\{ \hat{W}_i^C(Y_i) - W_i^C(Y_i) \right\}\] (24)

- The first part \([21]\) is just

\[
\frac{21}{} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i(\beta)
\]

where \( \epsilon_i(\beta) \) was defined in [12].

- The second part \([22]\) involves the difference between estimated and true IPCW weights for \( T \):

\[
\frac{22}{} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i(\beta) \left\{ \hat{W}_i^T(Y_i) - W_i^T(Y_i) \right\}
\]

\[
= \frac{1}{n^{1/3}} \sum_{i=1}^{n} \epsilon_i(\beta) \left\{ D_i^T(Y_i)' \Omega_T(\beta_T)^{-1} \sum_{j=1}^{n} U_j^T(\beta_T) + \sum_{j=1}^{n} J_{ij}^T(Y_i) \right\} + o_p(1)
\]

\[
= \frac{1}{n^{1/3}} \sum_{i=1}^{n} \sum_{j=1}^{n} \epsilon_i(\beta) D_i^T(Y_i)' \Omega_T(\beta_T)^{-1} U_j^T(\beta_T) (25)
\]

\[
+ \frac{1}{n^{1/3}} \sum_{i=1}^{n} \sum_{j=1}^{n} \epsilon_i(\beta) J_{ij}^T(Y_i) + o_p(1) (26)
\]
Eq. (25) is simplified as

\[
(25) = \frac{1}{n^{1.5}} \sum_{i=1}^{n} \sum_{j=1}^{n} \epsilon_i(\beta) D_i^T(Y_i)' T_{ij} \Omega_T(\beta_T)^{-1} U_j^T(\beta_T)
\]

\[
= \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left\{ \frac{1}{n} \sum_{i=1}^{n} \epsilon_i(\beta) D_i^T(Y_i)' \right\} \Omega_T(\beta_T)^{-1} U_j^T(\beta_T)
\]

where \( K_T(\beta) = E(\epsilon_i(\beta) D_i^T(Y_i)') \) was defined in (15), then

\[
(25) = \frac{1}{\sqrt{n}} K_T(\beta) \Omega_T(\beta_T)^{-1} \sum_{j=1}^{n} U_j^T(\beta_T)
\]

Since \( J_{ij}^T(Y_i) \) can be written as

\[
J_{ij}^T(Y_i) = \int_0^Y \exp \left\{ \beta_T^T Z_i^T(u) \right\} R_i(u) r_T^0(u; \beta_T)^{-1} dM_j^T(u)
\]

\[
= \int_0^L \exp \left\{ \beta_T^T Z_i^T(u) \right\} I(X_i \geq u) I(X_i \wedge L \geq u) r_T^0(u; \beta_T)^{-1} dM_j^T(u)
\]

\[
= \int_0^L \exp \left\{ \beta_T^T Z_i^T(u) \right\} I(X_i \geq u) r_T^0(u; \beta_T)^{-1} dM_j^T(u)
\]

Eq. (26) is simplified as

\[
(26) = \frac{1}{n^{1.5}} \sum_{i=1}^{n} \sum_{j=1}^{n} \epsilon_i(\beta) \left[ \int_0^L \exp \left\{ \beta_T^T Z_i^T(u) \right\} R_i(u) r_T^0(u; \beta_T)^{-1} dM_j^T(u) \right]
\]

\[
= \frac{1}{\sqrt{n}} \int_0^L \left[ \frac{1}{n} \sum_{i=1}^{n} \epsilon_i(\beta) \exp \left\{ \beta_T^T Z_i^T(u) \right\} R_i(u) \right] r_T^0(u; \beta_T)^{-1} \left\{ \sum_{j=1}^{n} M_j^T(u) \right\}
\]

where \( H_T(u; \beta) = E[\epsilon_i(\beta) \exp(\beta_T^T Z_i^T(u)) R_i(u)] \) was defined in (17), then

\[
(26) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^L H_T(u; \beta) r_T^0(u; \beta_T)^{-1} dM_i^T(u)
\]

To sum up, (22) can be rewritten as:

\[
(22) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[ K_T(\beta) \Omega_T(\beta_T)^{-1} U_i^T(\beta_T) + \int_0^L H_T(u; \beta) r_T^0(u; \beta_T)^{-1} dM_i^T(u) \right] + \alpha_T(1)
\]

- Similarly, (23) can be rewritten as:

\[
(23) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[ K_C(\beta) \Omega_C(\beta_C)^{-1} U_i^C(\beta_C) + \int_0^L H_C(u; \beta) r_C^0(u; \beta_C)^{-1} dM_i^C(u) \right] + \alpha_T(1)
\]

where \( K_C(\beta) = E[\epsilon_i(\beta) D_i^C(Y_i)'] \) and \( H_C(u; \beta) = E[\epsilon_i(\beta) \exp(\beta_C^T Z_i^C) R_i(u)] \) were defined in (18).
• Eq. (24) can be rewritten as:

\[
24 \quad \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) |Y_i - \xi^{-1} (x_i^T z_i (0)) z_i^T (0) \{ \hat{W}_i^{C} (Y_i) - w_i^{C} (Y_i) \} \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) |Y_i - \xi^{-1} (x_i^T z_i (0)) z_i^T (0) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

\[
= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i (Y_i) \{ \hat{W}_i^{T} (Y_i) - w_i^{T} (Y_i) \}
\]

Eq. (27)-(30) can be shown to be negligible.

To sum up, we can rewrite \( \sqrt{n} \Phi (\beta) \) as

\[
\sqrt{n} \Phi (\beta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} B_i (\beta) + o_p (1),
\]

where as defined in (11). Since we have defined \( B_i (\beta) = E \{ B_i (\beta) \} \) in (10), then we have proven that

\[
\sqrt{n} \Phi (\beta_D) \xrightarrow{D} \text{Normal}(0, B(\beta_D)),
\]

following that the mean of each term in the summation above is 0 at \( \beta_D \).

### A.6 Consistency

**Theorem 3** Under regularity conditions [1]-(31)], as \( n \to \infty \), \( \hat{\beta}_D \xrightarrow{p} \beta_D \).

**Proof 3** We use the Inverse Function Theorem (Foutz, 1977) by verifying the following conditions:

- \( \partial \Phi (\beta) / \partial \beta' \) exists and is continuous in an open neighborhood \( B_D \) of \( \beta_D \).
- \( -n^{-1} \partial \Phi (\beta) / \partial \beta' \) is positive definite with probability 1 as \( n \to \infty \).
- \( -n^{-1} \partial \Phi (\beta) / \partial \beta' \) converges in probability to a fixed function uniformly in an open neighborhood \( B_D \) of \( \beta_D \).
- Asymptotic unbiasedness of the estimating function: \( -\Phi (\beta_D) / n \xrightarrow{p} 0 \).

We know that

\[
\frac{\partial \Phi (\beta)}{\partial \beta'} = - \sum_{i=1}^{n} \Delta_i \hat{W}_i^{T} (Y_i) \hat{W}_i^{C} (Y_i) h \left\{ \beta' Z_i^D (0) \right\} Z_i^D (0)^{\otimes 2}
\]

where \( h(x) = \partial g^{-1}(x) / \partial x \). We will show that this derivative vector satisfies all the necessary conditions above.

- The first condition here holds because of the regularity condition [11], which states that \( h \) exists and is continuous in an open neighborhood \( B_D \) of \( \beta_D \).
• As to the second condition here, we know
\[
- \frac{1}{n} \frac{\partial \Phi(\beta)}{\partial \beta} \bigg|_{\beta = \beta_0} \\
= E \left[ \Delta_i W_i^T (Y_i) W_i^C (Y_i) h \left\{ \beta_i^D Z_i^D (0) \right\} Z_i^D (0)\right] + o_p(1)
\]
\[
= E \left[ E \left\{ \frac{I (T_i \geq D_i \wedge L)}{P(T_i \geq D_i \wedge L)} \right\} |D_i, Z_i^D (0)\right] h \left\{ \beta_i^D Z_i^D (0) \right\} Z_i^D (0)\right] + o_p(1)
\]
\[
= E \left[ E \left\{ \frac{I (T_i \geq D_i \wedge L)}{P(T_i \geq D_i \wedge L)} \right\} |D_i, Z_i^D (0)\right] h \left\{ \beta_i^D Z_i^D (0) \right\} Z_i^D (0)\right] + o_p(1)
\]
\[
= E \left[ h \left\{ \beta_i^D Z_i^D (0) \right\} Z_i^D (0)\right] + o_p(1)
\]
\[
= A(\beta).
\]

where \(A(\beta)\) is defined as \(\square\). Since we have assumed \(A(\beta_D)\) is positive definite, the second condition holds here too.

• The third condition holds by the law of large numbers.

• Finally, since we have proven that
\[
\sqrt{n} \Phi(\beta_D) \xrightarrow{D} \text{Normal}(0, B(\beta_D)).
\]

The last condition holds by Chebyshev’s inequality.

Having verified all the four conditions, we can argue that \(\hat{\beta}_D \xrightarrow{p} \beta_D\) follows from Inverse Function Theorem.

### A.7 Asymptotic Distribution

#### Theorem 4
Under regularity conditions \(\square\), as \(n \to \infty\),
\[
\sqrt{n} \left( \hat{\beta}_D - \beta_D \right) \xrightarrow{D} \text{Normal}(0, A(\beta_D)^{-1} B(\beta_D) A(\beta_D)^{-1}).
\]

#### Proof 4
Taylor expansion of \(\Phi(\hat{\beta}_D)\) around \(\beta_D\) is:
\[
0 = \Phi(\hat{\beta}_D) = \Phi(\beta_D) + \frac{\partial \Phi(\beta)}{\partial \beta} \bigg|_{\beta = \beta} \left( \hat{\beta}_D - \beta_D \right),
\]
where \(\beta\) lies between \(\hat{\beta}_D\) and \(\beta_D\). So
\[
\sqrt{n} \left( \hat{\beta}_D - \beta_D \right) = - \left\{ \frac{\partial \Phi(\beta)}{\partial \beta} \bigg|_{\beta = \beta} \right\}^{-1} \sqrt{n} \Phi(\beta_D)
\]
\[
= \left\{ - \frac{1}{n} \sum_{i=1}^{n} \Delta_i W_i^T (Y_i) Z_i^D (0)\right\}^{-1} \sqrt{n} \Phi(\beta_D)
\]
\[
= A(\beta_D)^{-1} \sqrt{n} \Phi(\beta_D) + o_p(1).
\]

Following Theorem \(\square\) it holds that
\[
\sqrt{n} \left( \hat{\beta}_D - \beta_D \right) \xrightarrow{D} \text{Normal}\left(0, A(\beta_D)^{-1} B(\beta_D) A(\beta_D)^{-1}\right).
\]

### B Model Selection Criteria

We suggest using Concordance Statistics (IOC), Mean Absolute Deviation (MAD) and Mean Squared Deviation (MSD) to select the proper link function. To simplify the notation, we denote \(D_i^L = D_i \wedge L\) and its predicted value as \(\tilde{D}_i^L = g^{-1} \{ \beta_i^D Z_i^D (0)\}\). Due to the occurrence of censoring, we observe \(X_i = D_i^L \wedge T_i \wedge C_i\) for subject \(i\).
Our version of IOC is adapted from Frank Harrell’s formula of concordance (Harrell, 1996; Heagerty, 2005, Uno et al., 2011):

\[
IOC = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_i \hat{W}_i(Y_i) \hat{W}_j(Y_j) I \left(Y_i < Y_j, \hat{D}^L_i < \hat{D}^L_j\right)}{\sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_i \hat{W}_i(Y_i) \hat{W}_j(Y_j) I \left(Y_i < Y_j\right)}.
\]

It converges to a censoring distribution free quantity \(P(\hat{D}^L_i < \hat{D}^L_j | D^L_i < D^L_j)\) because

(i) as to the numerator,

\[
\frac{1}{n^2} \Delta_i \hat{W}_i(Y_i) \hat{W}_j(Y_j) I \left(Y_i < Y_j, \hat{D}^L_i < \hat{D}^L_j\right) \xrightarrow{p} E \left\{ I (D^L_i \leq T_i \land C_i) W_i(Y_i) W_j(Y_j) I (Y_i < Y_j) I \left(\hat{D}^L_i < \hat{D}^L_j\right) \right\}
\]

\[
= E \left\{ I (D^L_i \leq T_i \land C_i) W_i(D^L_i) W_j(D^L_j) I (T_j \land C_j > D^L_i) I (D^L_j > D^L_i) I \left(\hat{D}^L_i < \hat{D}^L_j\right) \right\}
\]

\[
= E \left[ E \left\{ I (T_i \land C_i \geq D_i \land L) I (T_j \land C_j > D_i \land L) P(T_j \land C_j > D_i \land L) I \left(D^L_i < D^L_j, \hat{D}^L_i < \hat{D}^L_j\right) | Z_i^D(0), D_i \right\} \right]
\]

\[
= P(D^L_i < D^L_j, \hat{D}^L_i < \hat{D}^L_j).
\]

(ii) Similarly, the denominator follows that

\[
\frac{1}{n^2} \Delta_i \hat{W}_i(Y_i) \hat{W}_j(Y_j) I(Y_i < Y_j) \xrightarrow{p} P(D^L_i < D^L_j).
\]

(iii) So

\[
IOC \xrightarrow{p} P(\hat{D}^L_i < \hat{D}^L_j | D^L_i < D^L_j).
\]

We can also use the similar trick to prove that

\[
MAD := \frac{1}{n} \sum_{i=1}^{n} \Delta_i \hat{W}_i(Y_i) \left| Y_i - g^{-1}\left\{ \hat{\beta}_D Z_i^D(0) \right\} \right| \xrightarrow{p} E \left| D^L_i - \hat{D}^L_i \right|.
\]

\[
MSD := \frac{1}{n} \sum_{i=1}^{n} \Delta_i \hat{W}_i(Y_i) \left[ Y_i - g^{-1}\left\{ \hat{\beta}_D Z_i^D(0) \right\} \right]^2 \xrightarrow{p} E \left[ D^L_i - \hat{D}^L_i \right]^2.
\]

### C More Results in Application Data Analysis

Below are the plots of RMST within 1 year and 5 years post wait-list for chronic ESLD patients with different MELD scores.
Below are the estimated effects of prognostic factors on pre-transplant survival time within 1 year and 5 year post wait-list for chronic ESLD patients.
Web Table 1: Estimated covariate effects on RMST in the absence of liver transplantation ($L = 12$ months)

<table>
<thead>
<tr>
<th>UNOS Region</th>
<th>Linear</th>
<th>Log</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{Z}^D(0)$</td>
<td>$\hat{\beta}_D$</td>
<td>ASE$_1$</td>
</tr>
<tr>
<td>Intercept</td>
<td>$12.02$</td>
<td>$0.14$</td>
<td>$&lt; 0.01$</td>
</tr>
<tr>
<td>Year-2005</td>
<td>$0.1$</td>
<td>$0.01$</td>
<td>$&lt; 0.01$</td>
</tr>
<tr>
<td>Age-50 (Years)</td>
<td>$-0.05$</td>
<td>$&lt; 0.01$</td>
<td>$&lt; 0.01$</td>
</tr>
<tr>
<td>Sodium-130 (mmol/l)</td>
<td>$0.13$</td>
<td>$0.01$</td>
<td>$&lt; 0.01$</td>
</tr>
<tr>
<td>MELD Score-6</td>
<td>$-0.32$</td>
<td>$&lt; 0.01$</td>
<td>$&lt; 0.01$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender</th>
<th>Reference Group: Male</th>
<th>Female</th>
<th>Reference Group: White</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
<td>0.05</td>
<td>0.88</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th>Reference Group: O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>0.19</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.02</td>
</tr>
<tr>
<td>Asian</td>
<td>0.2</td>
</tr>
<tr>
<td>Others</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Blood Type</th>
<th>Reference Group: (20, 25]</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-0.05</td>
</tr>
<tr>
<td>B</td>
<td>-0.05</td>
</tr>
<tr>
<td>AB</td>
<td>-0.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Reference Group: No or Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hepatitis C</td>
<td>-0.09</td>
</tr>
<tr>
<td>Noncholestatic</td>
<td>0.29</td>
</tr>
<tr>
<td>Cholestatic</td>
<td>-0.05</td>
</tr>
<tr>
<td>Acute Hepatic Necrosis</td>
<td>0.9</td>
</tr>
<tr>
<td>Metastatic Disease</td>
<td>-0.45</td>
</tr>
<tr>
<td>Malignant Neoplasm</td>
<td>-1.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BMI</th>
<th>Reference Group: (20, 25]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 20]</td>
<td>-0.46</td>
</tr>
<tr>
<td>(25, 30]</td>
<td>0.08</td>
</tr>
<tr>
<td>&gt; 30</td>
<td>0.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hospitalized</th>
<th>Reference Group: Not Hospitalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICU</td>
<td>-1.63</td>
</tr>
<tr>
<td>not ICU</td>
<td>-1.43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dialysis</th>
<th>Reference Group: No or Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.81</td>
</tr>
</tbody>
</table>

An offset of $L = 12$ months is applied for log link.
Web Table 2: Estimated covariate effects on RMST in the absence of liver transplantation ($L = 60$ months)

<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>Log</th>
<th>Logistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Z^r_v(0)$</td>
<td>$\hat{\beta}_D$</td>
<td>ASE$_1$</td>
</tr>
<tr>
<td>Intercept</td>
<td>46.91</td>
<td>0.88</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Year-2005</td>
<td>-3.2</td>
<td>0.07</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Age-50 (Years)</td>
<td>0.54</td>
<td>0.02</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Sodium-130 (mmol/l)</td>
<td>0.66</td>
<td>0.04</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>MELD Score-6</td>
<td>-1.19</td>
<td>0.02</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>UNOS Region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.37</td>
<td>0.78</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>2</td>
<td>-2.31</td>
<td>0.64</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>3</td>
<td>-2.31</td>
<td>0.76</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
<td>0.63</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.66</td>
<td>0.94</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>-0.55</td>
<td>0.68</td>
<td>0.41</td>
</tr>
<tr>
<td>7</td>
<td>-1.59</td>
<td>0.83</td>
<td>0.05</td>
</tr>
<tr>
<td>8</td>
<td>-0.105</td>
<td>0.65</td>
<td>0.1</td>
</tr>
<tr>
<td>9</td>
<td>-3.94</td>
<td>0.83</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>10</td>
<td>-3.46</td>
<td>0.77</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.54</td>
<td>0.36</td>
<td>0.14</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-0.99</td>
<td>0.61</td>
<td>0.1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.5</td>
<td>0.49</td>
<td>0.3</td>
</tr>
<tr>
<td>Asian</td>
<td>1.63</td>
<td>0.87</td>
<td>0.06</td>
</tr>
<tr>
<td>Others</td>
<td>-0.86</td>
<td>1.56</td>
<td>0.58</td>
</tr>
<tr>
<td>Blood Type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.06</td>
<td>0.36</td>
<td>0.86</td>
</tr>
<tr>
<td>B</td>
<td>-0.51</td>
<td>0.59</td>
<td>0.38</td>
</tr>
<tr>
<td>AB</td>
<td>0.55</td>
<td>1.07</td>
<td>0.61</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hepatitis C</td>
<td>-1.54</td>
<td>0.54</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Noncholestatic</td>
<td>2.34</td>
<td>0.55</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>Cholestatic</td>
<td>-0.98</td>
<td>0.77</td>
<td>0.2</td>
</tr>
<tr>
<td>Acute Hepatic Necrosis</td>
<td>2.4</td>
<td>1.5</td>
<td>0.11</td>
</tr>
<tr>
<td>Metastatic Disease</td>
<td>-2.58</td>
<td>1.41</td>
<td>0.07</td>
</tr>
<tr>
<td>Malignant Neoplasm</td>
<td>-8.59</td>
<td>0.69</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0, 20]</td>
<td>-2.64</td>
<td>0.86</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>(25, 30]</td>
<td>-0.15</td>
<td>0.46</td>
<td>0.75</td>
</tr>
<tr>
<td>&gt; 30</td>
<td>-0.12</td>
<td>0.47</td>
<td>0.8</td>
</tr>
<tr>
<td>Hospitalized</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not ICU</td>
<td>0.176</td>
<td>0.59</td>
<td>$&lt;0.01$</td>
</tr>
<tr>
<td>ICU</td>
<td>-0.18</td>
<td>0.75</td>
<td>0.01</td>
</tr>
<tr>
<td>Dialysis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>2.01</td>
<td>0.77</td>
<td>0.01</td>
</tr>
</tbody>
</table>

An offset of $L = 60$ months is applied for log link.