Research Article

Statistics in Medicine

Received XXXX

(www.interscience.wiley.com) DOI: 10.1002/sim.0000

Evaluating Surrogate Marker Information using Censored Data: Supplementary Materials

The following notation is used in the Appendix to simplify the presentations: $K_h(x) = K(x/h)/h$, $P(C^{(g)} > u) = W_g^C(u)$, $P(T^{(g)} > u) = W_g^T(u)$, $P(X^{(g)} > u) = W_g^X(u)$, $f_g(s \mid t_0)$ is the density function of S_{gi} given $T_{gi} > t_0$, for $g = A$ or B,

$$r_A(t|s,t_0) = \frac{1}{f_A(s|t_0)\psi_A(t|s,t_0)W_A^C(t)}\ , \quad \text{and} \quad \Psi_{adj}^{(A)}(t,t_0) = \int \psi_A(t \mid s,t_0)dF_B(s|t_0).$$

In addition to the independent censoring assumption and conditions (2.1) and (2.2) stated in the text, we assume the following regularity conditions: (1) the density function of S_{Ai} has a continuous derivatives bounded away from zero, (2) S_{Ai} and S_{Bi} have the same bounded support Ω_S, (3) $\Lambda_A(t|s,t_0)$ has bounded continuous first and second partial derivatives with respect to t and s, (4) $\{n_A/n, n_B/n\} \to (\pi_A, \pi_B) \in (0,1) \times (0,1)$ as $n = n_A + n_B \to \infty$, (5) $h = O(n^{-u}), 1/4 < u < 1/2$, (6) K is a kernel symmetric at zero on the support $(-1,1)$ and $\int K(x)^2dx < \infty$, (7) $P(T^{(g)} > t) > 0$ and $P(T^{(g)} > t_0) < 1$ for $g = A, B$. For simplicity of presentation we suppress the $\gamma(\cdot)$ in $K_h\{\gamma(S_{Ai}) - \gamma(s)\}$ and simply write $K_h(S_{Ai} - s)$.

Appendix A:. Justification for the Proposed Proportion of Treatment Effect Explained

We next show that under conditions (C1)-(C3), $0 \le \Delta_S(t,t_0) \le \Delta(t)$ and $0 \le R_S(t,t_0) \le 1$. Using integration by parts

$$\Delta(t) - \Delta_S(t,t_0) = \int_{s_l}^{s_u} \psi_A(t|s,t_0)\left\{W_A^T(t_0)F_A(ds|t_0) - W_B^T(t_0)F_B(ds|t_0)\right\}$$

$$= \psi_A(t|s_u,t_0)\left\{W_A^T(t_0) - W_B^T(t_0)\right\} - \int_{s_l}^{s_u} \left\{W_A^T(t_0)F_A(s|t_0) - W_B^T(t_0)F_B(s|t_0)\right\}\psi_A(t|ds,t_0)$$

$$= \psi_A(t|s_l,t_0)\left\{W_A^T(t_0) - W_B^T(t_0)\right\} + \int_{s_l}^{s_u} \left\{P(T^{(A)} > t_0, S^{(A)} > s) - P(T^{(B)} > t_0, S^{(B)} > s)\right\}\psi_A(t|ds,t_0)$$

where s_u and s_l are the upper and lower bounds of the support for $S^{(g)}|T^{(g)} > t_0$, respectively. From condition (C2), we have $W_A^T(t_0) - W_B^T(t_0) = P(S^{(A)} > s_l, T^{(A)} > t_0) - P(S^{(B)} > s_l, T^{(B)} > t_0) \ge 0$ and $P(S^{(A)} > s, T^{(A)} > t_0) - P(S^{(B)} > s, T^{(B)} > t_0) \ge 0$. This, together with condition (C1) that $\psi_A(t|s,t_0)$ is an monotone increasing function in s, we have $\Delta(t) - \Delta_S(t,t_0) \ge 0$. Furthermore, from condition (C3) that $\psi_A(t|s,t_0) - \psi_B(t|s,t_0) \ge 0$,

$$\Delta_S(t,t_0) = W_B^T(t_0)\int \psi_A(t|s,t_0)F_B(ds|t_0) - W_B^T(t)$$

$$= \int \left\{\psi_A(t|s,t_0) - \psi_B(t|s,t_0)\right\} W_B^T(t_0)F_B(ds|t_0) \ge 0.$$

Therefore, $0 \le \Delta_S(t,t_0) \le \Delta(t)$ and $0 \le R_S(t,t_0) \le 1$ if we define $0/0 = 0$ to account for the case $\Delta(t) = 0$.

Appendix B:. Asymptotic properties of $\widehat{\Delta}_S(t,t_0)$

Under regularity conditions (1), (3), (5) and (6), $\widehat{\Lambda}_A(t|s,t_0)$ is a consistent estimator for $\Lambda_A(t|s,t_0)$ uniformly in $s \in \Omega_S$ and consequently $\sup_{\Omega_S}|\widehat{\psi}_A(t|s,t_0) - \psi_A(t|s,t_0)| = o_p(1)$, from the uniform consistency of the conditional Nelson

Aalen estimator [1]. This, together with the uniform consistency of the Kaplan-Meier estimator [2], implies that and

$$n_B^{-1}\sum_{X_{Bi}>t_0}\frac{\widehat{\psi}_A(t|S_{Bi},t_0)}{\widehat{W}_B^C(t_0)}-W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)=n_B^{-1}\sum_{X_{Bi}>t_0}\left\{\frac{\widehat{\psi}_A(t|S_{Bi},t_0)}{\widehat{W}_B^C(t_0)}-\frac{\psi_A(t|S_{Bi},t_0)}{W_B^C(t_0)}\right\}$$
$$+n_B^{-1}\sum_{i=1}^{n_B}\left\{\frac{I(X_{Bi}>t_0)\psi_A(t|S_{Bi},t_0)}{W_B^C(t_0)}-W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)\right\}=o_p(1)$$

Similarly, the inverse probability weighted estimator $n_B^{-1}\sum_{i=1}^{n_B}I(X_{Bi}>t)/\widehat{W}_B^C(t)$ is a consistent estimator of the survival probability $W_B^T(t)=P(T^{(B)}>t)$. Therefore, $\widehat{\Delta}_S(t,t_0)-\Delta_S(t,t_0)=o_p(1)$.

We next derive the asymptotic distribution for $\widehat{\mathcal{W}}_S(t,t_0)=n^{\frac{1}{2}}\{\widehat{\Delta}_S(t,t_0)-\Delta_S(t,t_0)\}=\pi_B^{-\frac{1}{2}}\{\widehat{\mathcal{W}}_{S2}(t,t_0)-\widehat{\mathcal{W}}_{S1}(t)\}$, where $\widehat{\mathcal{W}}_{S1}(t)=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\{I(X_{Bi}>t)/\widehat{W}_B^C(t)-W_B^T(t)\}$ and

$$\widehat{\mathcal{W}}_{S2}(t,t_0)=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\left[\frac{\widehat{\psi}_A(t|S_{Bi},t_0)I(X_{Bi}>t_0)}{\widehat{W}_B^C(t_0)}-W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)\right].$$

From the asymptotic properties of the Kaplan-Meier estimator [2],

$$n_B^{\frac{1}{2}}\left\{\frac{1}{\widehat{W}_B^C(t)}-\frac{1}{W_B^C(t)}\right\}=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\frac{1}{W_B^C(t)}\int_0^t\frac{dM_{Bi}^C(z)}{W_B^X(z)},\qquad(1)$$

where $M_{gi}^C(z)=I(X_{gi}\le z)(1-\delta_{gi})-\int_0^z I(X_{gi}\ge u)\lambda_g^C(u)du$ and $\lambda_g^C(u)$ is the hazard function of $C^{(g)}$, for $g=A,B$. It follows that $\widehat{\mathcal{W}}_{S1}(t)=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\xi_{Bi}(t)+o_p(1)$, where

$$\xi_{Bi}(t)=W_B^T(t)\left\{\frac{I(X_{Bi}>t)}{W_B^X(t)}-1+\int_0^t\frac{dM_{Bi}^C(z)}{W_B^X(z)}\right\}.$$

Furthermore

$$\frac{I(X_{Bi}>t)}{W_B^X(t)}-1+\int_0^t\frac{dM_{Bi}^C(z)}{W_B^X(z)}$$
$$=\frac{I(X_{Bi}>t)}{W_B^X(t)}-1+\int_0^t\frac{d\left\{M_{Bi}^C(z)+M_{Bi}^T(z)\right\}}{W_B^X(z)}-\int_0^t\frac{dM_{Bi}^T(z)}{W_B^X(z)}$$
$$=-\int_0^t\frac{dM_{Bi}^T(z)}{W_B^X(z)},$$

where $M_{gi}^T(z)=I(X_{gi}\le z)\delta_{gi}-\int_0^z I(X_{gi}\ge u)\lambda_g^T(u)du$, $\lambda_g^T(u)$ is the hazard function of $T^{(g)}$ and $\lambda_g^X(u)=\lambda_g^T(u)+\lambda_g^C(u)$ is the hazard function for $X^{(g)}=T^{(g)}\wedge C^{(g)}$ for $g=A,B$. Therefore,

$$\xi_{Bi}(t)=-W_B^T(t)\int_0^t\frac{dM_{Bi}^T(z)}{W_B^X(z)}$$

and $n_B^{-1}\sum_{i=1}^{n_B}\{I(X_{Bi}>t)/\widehat{W}_B^C(t)$ is asymptotically equivalent to the KM estimator for the survival function of $T^{(B)}$ at time t.

For $\widehat{\mathcal{W}}_{S2}(t,t_0)$, we may write $\widehat{\mathcal{W}}_{S2}(t,t_0)=I_1+I_2$, where

$$I_1=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\left[\frac{\psi_A(t|S_{Bi},t_0)I(X_{Bi}>t_0)}{\widehat{W}_B^C(t_0)}-W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)\right]$$
$$I_2=n_B^{\frac{1}{2}}\int\left\{\widehat{\psi}_A(t|s,t_0)-\psi_A(t|s,t_0)\right\}\widehat{F}_B(ds,t_0)$$

and $\widehat{F}_B(s,t_0)=n_B^{-1}\sum_{i=1}^{n_B} I(S_{Bi}\le s, X_{Bi}>t_0)/\widehat{W}_B^C(t_0)$. For I_1, from (1) we have $I_1=n_B^{-\frac{1}{2}}\sum_{i=1}^{n_B}\eta_{Bi}^{(1)}(t,t_0)+o_p(1)$, where

$$\begin{aligned}\eta_{Bi}(t,t_0) &= W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)\left[\frac{\psi_A(t|S_{Bi},t_0)}{\Psi_{adj}^{(A)}(t,t_0)}\frac{I(X_{Bi}>t_0)}{W_B^X(t_0)}-1+\int_0^{t_0}\frac{dM_{Bi}^C(u)}{W_B^X(u)}\right]\\ &= W_B^T(t_0)\Psi_{adj}^{(A)}(t,t_0)\left[\left\{\frac{\psi_A(t|S_{Bi},t_0)}{\Psi_{adj}^{(A)}(t,t_0)}-1\right\}\frac{I(X_{Bi}>t_0)}{W_B^X(t_0)}-\int_0^{t_0}\frac{dM_{Bi}^T(u)}{W_B^X(u)}\right]\end{aligned}$$

For I_2, we note that under regularity conditions (1), (3), (5) and (6),

$$\widehat{\Lambda}_A(t|s,t_0)-\Lambda_A(t|s,t_0)=n_A^{-1}\sum_{i=1}^{n_A}\left[K_h(S_{Ai}-s)\int_{t_0}^t r_A(u|s,t_0)\mathcal{H}_{Ai}(du|s,t_0)\right]+O_p(h^2)$$

uniformly in s, where

$$\mathcal{H}_{Ai}(t|s,t_0)=\frac{I(X_{Ai}>t_0)}{W_A^T(t_0)}M_{Ai}^T(t|s,t_0),\ M_{Ai}^T(t|s,t_0)=I(X_{Ai}\le t)\delta_{Ai}-\int_0^t I(X_{Ai}\ge u)\Lambda_A^T(du|s,t_0).$$

It then follows from the functional delta method [3] that

$$\frac{\widehat{\psi}_A(t|s,t_0)-\psi_A(t|s,t_0)}{\psi_A(t|s,t_0)}=-n_A^{-1}\sum_{i=1}^{n_A}\left[K_h(S_{Ai}-s)\int_0^t r_A(z|s,t_0)\mathcal{H}_{Ai}(du|s,t_0)\right]+O_p(h^2).$$

Therefore,

$$\begin{aligned}I_2 &= -n_A^{-1}n_B^{\frac{1}{2}}\sum_{j=1}^{n_A}\int\int_{t_0}^t \psi_A(t|s,t_0)K_h(S_{Aj}-s)r_A(u|s,t_0)\mathcal{H}_{Aj}(du|s,t_0)\hat{F}_B(ds|t_0)+o_p(n_B^{\frac{1}{2}}h^2)\\ &= -n_A^{-1}n_B^{\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t\left\{\int\psi_A(t|s,t_0)K_h(S_{Aj}-s)r_A(u|s,t_0)\hat{F}_B(ds|t_0)\right\}M_{Aj}^T(du|s,t_0)+o_p(1)\end{aligned}$$

Noting that

$$\sup_{u\in[t_0,t],\tilde{s}\in\Omega_S}|H_{\epsilon,n}(u,\tilde{s})|=o_p(1),$$

where

$$H_{\epsilon,n}(u,\tilde{s})=\int\psi_A(t|s,t_0)K_h(\tilde{s}-s)r_A(u|s,t_0)\hat{F}_B(ds|t_0)-\frac{f_B(\tilde{s}|t_0)\psi_A(t|\tilde{s},t_0)}{f_A(u|\tilde{s})\psi_A(u|\tilde{s},t_0)W_A^C(u)},$$

we have

$$\begin{aligned}&n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t\left\{\int\psi_A(t|s,t_0)K_h(S_{Aj}-s)r_A(u|s,t_0)\hat{F}_B(ds|t_0)\right\}M_{Aj}^T(du|s,t_0)\\ =&n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t\frac{f_B(S_{Aj}|t_0)\psi_A(t|S_{Aj},t_0)}{f_A(u|S_{Aj})\psi_A(u|S_{Aj},t_0)W_A^C(u)}M_{Aj}^T(du|s,t_0)+n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t H_{\epsilon,n}(u,S_{Aj})M_{Aj}^T(du|s,t_0)\\ =&n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t\frac{f_B(S_{Aj}|t_0)\psi_A(t|S_{Aj},t_0)}{f_A(u|S_{Aj})\psi_A(u|S_{Aj},t_0)W_A^C(u)}M_{Aj}^T(du|s,t_0)+o_p(1).\end{aligned}$$

Here we used the fact that the variance of $n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\frac{I(X_{Aj}>t_0)}{W_A^T(t_0)}\int_{t_0}^t H_{\epsilon,n}(u,S_{Aj})M_{Aj}^T(du|s,t_0)$ is bounded by

$$E\sup_{(u,\tilde{s})\in[t_0,t]\times\Omega_S}|H_{\epsilon,n}(u,\tilde{s})|^2\left[\int_{t_0}^t\frac{I(X_{Aj}>u)}{W_A^T(t_0)^2}\Lambda_A^T(du|s,t_0)\right]=o(1).$$

Therefore

$$I_2=(\pi_B/\pi_A)^{\frac{1}{2}}n_A^{-\frac{1}{2}}\sum_{j=1}^{n_A}\eta_{Aj}(t,t_0)+o_p(1),$$

where

$$\eta_{Aj}(t,t_0) = \psi_A(t|S_{Aj},t_0)\frac{I(X_{Aj}>t_0)f_B(S_{Aj}|t_0)}{W_A^X(t_0)f_A(S_{Aj}|t_0)}\int_{t_0}^{t}\frac{M_{Aj}^T(du|S_{Aj},t_0)}{\psi_A(u|S_{Aj},t_0)W_A^C(u|t_0)}$$

and $W_A^C(u|t_0) = P(C^{(A)} > u|C^{(A)} > t_0)$. It follows that

$$\widehat{\mathcal{W}}_S(t,t_0) = (\pi_B n_B)^{-\frac{1}{2}}\sum_{i=1}^{n_B}\{\eta_{Bi}(t,t_0) - \xi_{Bi}(t,t_0)\} + (\pi_A n_A)^{-\frac{1}{2}}\sum_{j=1}^{n_A}\eta_{Aj}(t,t_0) + o_p(1),$$

Therefore, under regularity condition (4), $\widehat{\mathcal{W}}_S(t,t_0)$ converges to a normal distribution with mean zero and variance $\sigma_S^2(t,t_0) = \pi_B^{-1}E[\{\eta_{Bi}(t,t_0) - \xi_{Bi}(t,t_0)\}^2] + \pi_A^{-1}E\{\eta_{Aj}(t,t_0)^2\}$ by the central limit theorem.

Appendix C:. Justification of the perturbation-resampling procedure

Let $\widehat{\mathcal{W}}_S^{(b)}(t,t_0) = n^{\frac{1}{2}}\{\widehat{\Delta}_S^{(b)}(t,t_0) - \Delta_S(t,t_0)\}$. Using similar arguments as given in Appendix A, it can be shown that

$$\widehat{\mathcal{W}}_S^{(b)}(t,t_0) = (\pi_B n_B)^{-\frac{1}{2}}\sum_{i=1}^{n_B}\{\eta_{Bi}(t,t_0) - \xi_{Bi}(t,t_0)\}V_{Bi}^{(b)} + (\pi_A n_A)^{-\frac{1}{2}}\sum_{j=1}^{n_A}\eta_{Aj}(t,t_0)V_{Ai}^{(b)} + o_{\tilde{p}}(1)$$

where $\mathbf{V}^{(b)} = \{V_{B1}^{(b)},\cdots,V_{Bn_B}^{(b)},V_{A1}^{(b)},\cdots,V_{An_A}^{(b)}\}$ are i.i.d positive random variables with unit mean and variance as well as a finite third moment and $\tilde{p}$ is the joint probability measure generated by $\mathbf{V}^{(b)}$ and data. Coupled with the expansion for $\widehat{\Delta}_S(t,t_0) - \Delta_S(t,t_0)$, it implies that

$$\begin{aligned} n^{\frac{1}{2}}\{\widehat{\Delta}_S^{(b)}(t,t_0) - \widehat{\Delta}_S(t,t_0)\} &= (\pi_B n_B)^{-\frac{1}{2}}\sum_{i=1}^{n_B}\{\eta_{Bi}(t,t_0) - \xi_{Bi}(t,t_0)\}(V_{Bi}^{(b)} - 1) \\ &\quad + (\pi_A n_A)^{-\frac{1}{2}}\sum_{j=1}^{n_A}\eta_{Aj}(t,t_0)(V_{Ai}^{(b)} - 1) + o_{\tilde{p}}(1). \end{aligned}$$

Conditional on the observed data and thus conditional on $\{\xi_{Bi}(t,t_0),\eta_{Bi}(t,t_0),\eta_{Aj}(t,t_0), i = 1,\cdots,n_A, j = 1,\cdots,n_B\}$, $n^{\frac{1}{2}}\{\widehat{\Delta}_S^{(b)}(t,t_0) - \widehat{\Delta}_S(t,t_0)\}$ converges weakly to a mean zero Gaussian random variable with a variance of $\sigma_S^2(t,t_0)$, as $n\to\infty$. Therefore

$$\sup_x\left|\text{pr}\left(n^{\frac{1}{2}}(\widehat{\Delta}_S^{(b)} - \widehat{\Delta}_S) \le x \mid \text{data}\right) - \text{pr}\left(n^{\frac{1}{2}}(\widehat{\Delta}_S - \Delta_S) \le x\right)\right| = o_p(1),$$

and the conditional distribution of $n^{\frac{1}{2}}\{\widehat{\Delta}_S^{(b)}(t,t_0) - \widehat{\Delta}_S(t,t_0)\}$ can be used to approximate that of $n^{\frac{1}{2}}\{\widehat{\Delta}_S(t,t_0) - \Delta_S(t,t_0)\}$ when the sample size is large.

References

1. Dabrowska DM, *et al.*. Uniform consistency of the kernel conditional kaplan-meier estimate. *The Annals of Statistics* 1989; **17**(3):1157–1167.
2. Kalbfleisch JD, Prentice RL. *The statistical analysis of failure time data*, vol. 360. John Wiley & Sons, 2011.
3. Kosorok MR. *Introduction to empirical processes and semiparametric inference*. Springer Science & Business Media, 2007.