Supporting information for "SMIM: a unified framework of survival sensitivity analysis using multiple imputation and martingale" by Yang, Zhang, Liu, and Guan.

Web Appendix S1 provides the preliminary for the proofs. Web Appendix S2 establishes the asymptotic linearization of $\hat{S}_{0,\mathrm{mi}}(t)$. Web Appendix S3 describes the $\sigma-$fields. Web Appendices S4 and S5 provide the proofs of Theorem 1 and Theorem 2. Web Appendix S6 presents a comprehensive simulation study.

Web Appendix S1 Preliminary

We adopt the counting process theory of Andersen and Gill (1982) in our theoretical framework. We state the existing results which will be used in our proof throughout.

To simplify the exposition, we introduce additional notation. We use $\xrightarrow{p}$ and $\xrightarrow{d}$ to represent "converge in probability as $n \to \infty$" and "converge in distribution as $n \to \infty$", respectively. Also, let $n_1/n \to p_1 \in (0,1)$ and $n_0/n \to p_0 \in (0,1)$, as $n \to \infty$. We do not state this condition formally as an assumption because it holds trivially for most of clinical trials where the two treatment groups are relatively balanced in their sample sizes.

Let $X_i^{\otimes l}$ denote 1 for $l = 0$, X_i for $l = 1$, and $X_i X_i^{\mathrm{T}}$ for $l = 2$. Define

$$U_a^{(l)}(\beta_a, t) = \frac{1}{n_a}\sum_{i=1}^{n} \mathbf{1}(A_i = a) X_i^{\otimes l} e^{\beta_a^{\mathrm{T}} X_i} Y_i(t),\ u_a^{(l)}(\beta_a, t) = \mathbb{E}\left\{X^{\otimes l} e^{\beta_a^{\mathrm{T}} X} Y(t)\right\},$$

where $u_a^{(l)}(\beta_a, t)$ is the expectation of $U_a^{(l)}(\beta_a, t)$, for $l = 0, 1, 2$. Moreover, define

$$E_a(\beta_a, t) = \frac{U_a^{(1)}(\beta_a, t)}{U_a^{(0)}(\beta_a, t)},\ e_a(\beta_a, t) = \frac{u_a^{(1)}(\beta_a, t)}{u_a^{(0)}(\beta_a, t)}.$$

The maximum partial likelihood estimator $\hat{\beta}_a$ solves

$$\mathcal{S}_{a,n}(\beta_a) = \frac{1}{n_a}\sum_{i=1}^{n} \mathbf{1}(A_i = a) \int_0^{\tau} \left\{X_i - \frac{U_1^{(1)}(\beta_a, u)}{U_1^{(0)}(\beta_a, u)}\right\} \mathrm{d}N_i(u) = 0.$$

We state the standard asymptotic results for $\hat{\beta}_a$ and $\hat{\lambda}_a(\cdot)$ requiring certain regularity conditions. To avoid too many technical distractions, we omit the exact conditions in Assumption S1 for the consistency and uniform convergency of the estimators of Cox models.

Assumption S1 *i) (Positivity) There exists a constant c such that with probability one, $S_a(t \mid X_i) \geq c > 0$ for t in $[0, \tau]$ and $a = 0, 1$. ii) Conditions A–D in Andersen and Gill (1982) hold for treatment group $a = 0, 1$.*

Following Andersen and Gill (1982), we have

$$n_a^{1/2}(\hat{\beta}_a - \beta_a) \quad = \quad \Gamma_a^{-1} \frac{1}{n_a^{1/2}} \sum_{i=1}^{n} \mathbf{1}(A_i = a) H_{a,i} + o_p(1), \tag{S1}$$

where $\Gamma_a = \mathbb{E}\{-\partial \mathcal{S}_{a,n}(\beta_a)/\partial \beta_a^{\mathrm{T}}\}$ is the Fisher information matrix of β_a, $H_{a,i} = \int_0^L \{X_i - e_a(\beta_a, u)\} \mathbf{1}(A_i = a) \mathrm{d}M_{a,i}(u)$, and

$$\mathrm{d}M_{a,i}(t) = \mathrm{d}N_i(u) - e^{\beta_a^{\mathrm{T}} X_i} Y_i(u) \lambda_{a,0}(u) \mathrm{d}u. \tag{S2}$$

Moreover, $n^{1/2}\{S_a(t \mid X_i; \hat{\theta}) - S_a(t \mid X_i)\}$ converges uniformly to a Gaussian process in $[0, L]$ for all X_i.

Web Appendix S2 Asymptotic linearization of $\hat{S}_{0,\mathrm{mi}}(t)$

To obtain the asymptotic linearization of $\hat{S}_{0,\mathrm{mi}}(t)$, we have

$$n^{1/2}\left\{\hat{S}_{0,\mathrm{mi}}(t) - S_0^{\mathrm{sen}}(t)\right\}$$

$$= \frac{n^{1/2}}{mn_0} \sum_{j=1}^{m} \sum_{i=1}^{n} (1 - A_i)\{1 - Y_i(t)\} \left[\mathbf{1}(T_i^{*(j)} \geq t) - S_0\{t \mid H_i(t); \hat{\theta}\}\right] \tag{S3}$$

$$+ \frac{n^{1/2}}{n_0} \sum_{i=1}^{n} (1 - A_i) \left[Y_i(t) + \{1 - Y_i(t)\}(1 - I_i) S_0\{t \mid H_i(t); \theta\} - S_0^{\mathrm{sen}}(t)\right] \tag{S4}$$

$$+ \frac{n^{1/2}}{n_0} \sum_{i=1}^{n} (1 - A_i) \phi_{0,i}(t) + o_p(1), \tag{S5}$$

where the exact expression of $\phi_{0,i}(t)$ is given in Section Web Appendix S4, reflecting the estimation of $\{\lambda_0(\cdot), \beta_0\}$. In our context, the imputation for the control group uses the information only from the control group. By the imputation and estimation procedures, (S3)–(S5) have (conditional) mean zero.

Web Appendix S3 σ-fields for the martingales

We consider the σ-fields as follows

$$\mathcal{F}_{n,k} = \begin{cases} \sigma\left(O_1, \ldots, O_k\right), & \text{for } k = i \quad (1 \le i \le n_1), \\ \sigma\left(O_1, \ldots, O_{n_1}, T_1^{*(1)}, \ldots, T_i^{*(j)}\right), & \text{for } k = n_1 + (i-1)m + j \\ & \quad (1 \le i \le n_1, 1 \le j \le m), \\ \sigma\Big(O_1, \ldots, O_{n_1}, T_1^{*(1)}, \ldots, T_{n_1}^{*(m)}, & \text{for } k = (1+m)n_1 + i \\ \quad O_{n_1+1}, \ldots, O_k\Big), & \quad (n_1 + 1 \le i \le n), \\ \sigma\Big(O_1, \ldots, O_{n_1}, T_1^{*(1)}, \ldots, T_{n_1}^{*(m)}, & \text{for } k = (1+m)n_1 + n_0 + (i-1)m + j \\ \quad O_{n_1+1}, \ldots, O_n, T_{n_1+1}^{*(1)}, \ldots, T_i^{*(j)}\Big), & \quad (n_1 + 1 \le i \le n, 1 \le j \le m). \end{cases}$$

Web Appendix S4 Proof of Theorem 1

We first derive the martingale representation of the MI estimator under δ-adjusted Cox models and control-based Cox models, separately. Then, we apply the martingale CLT to derive the asymptotic distribution of the MI estimator.

S4.1 Delta-adjusted Cox models

A key step is to separate the imputation step and the estimation step. We start with treatment group $a = 1$. For the imputations, it is important to recognize that $T_i^{*(j)}$ follows a time-dependent Cox model with the conditional survival function $S_1\{t \mid H_i(t); \hat{\theta}\}$ for $t > U_i$, where

$$S_1\{t \mid H_i(t); \theta\} = \begin{cases} \exp\left\{-\int_{U_i}^t \lambda_1(u) e^{\beta_1^{\mathrm{T}} X_i} \mathrm{d}u\right\}, & \text{if } A_i = 1, R_i = 1, \\ \exp\left\{-\delta \int_{U_i}^t \lambda_0(u) e^{\beta_0^{\mathrm{T}} X_i} \mathrm{d}u\right\}, & \text{if } A_i = 1, R_i = 2. \end{cases}$$

We express the MI estimator of $S_1^{\delta\text{-adj}}(t)$ as

$$n^{1/2}\left\{\hat{S}_{1,\mathrm{mi}}(t) - S_1^{\delta\text{-adj}}(t)\right\}$$

$$
\begin{aligned}
&= \frac{n^{1/2}}{mn_1}\sum_{j=1}^{m}\sum_{i=1}^{n} A_i\{\mathbf{1}(T_i^{*(j)} \geq t) - S_1^{\delta\text{-adj}}(t)\} \\
&= \frac{n^{1/2}}{mn_1}\sum_{j=1}^{m}\sum_{i=1}^{n} A_i[\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t);\hat{\theta}\}] + \frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i[S_1\{t \mid H_i(t);\hat{\theta}\} - S_1^{\delta\text{-adj}}(t)] \\
&= \frac{n^{1/2}}{mn_1}\sum_{j=1}^{m}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}[\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t);\hat{\theta}\}] \qquad \text{(S6)} \\
&\quad + \frac{n^{1/2}}{n_1}\sum_{i=1}^{n}\Big[A_iY_i(t) + A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\} - S_1^{\delta\text{-adj}}(t)\Big] + o_p(1), \qquad \text{(S7)}
\end{aligned}
$$

where (S6) follows because $\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t);\hat{\theta}\} = 0$ for subject i with $\{A_i = 1, Y_i(t) = 1\}$, and (S7) follows because $A_i\{1-Y_i(t)\}I_iS_1\{t \mid H_i(t);\hat{\theta}\} = 0$.

By the counting process theory, we can express the term $n_1^{-1/2}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\}$ in (S7) further as

$$
\begin{aligned}
&\frac{1}{n_1^{1/2}}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\} \\
&= \frac{1}{n_1^{1/2}}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\exp\left\{-\int_{U_i}^{t}\hat{\lambda}_1(u)\delta^{\mathbf{1}(R_i=2)}e^{\hat{\beta}_1^{\mathrm{T}}X_i}\mathrm{d}u\right\} \\
&= \frac{1}{n_1^{1/2}}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\theta\} \\
&\quad + \frac{1}{n_1^{1/2}}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\theta\}\left[-\int_{U_i}^{t}\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}\left\{\hat{\lambda}_1(u)-\lambda_1(u)\right\}\mathrm{d}u\right] \qquad \text{(S8)} \\
&\quad + \left[\frac{1}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\theta\}\left\{-\int_{U_i}^{t}\lambda_1(u)\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}X_i\mathrm{d}u\right\}\right] \qquad \text{(S9)} \\
&\quad \times n_1^{1/2}\left(\hat{\beta}_1-\beta_1\right). \qquad \text{(S10)}
\end{aligned}
$$

For (S8), we further express the key term as

$$
\begin{aligned}
&\int_{U_i}^{t}\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}\left\{\hat{\lambda}_1(u)-\lambda_1(u)\right\}\mathrm{d}u \\
&= \int_{U_i}^{t}\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}\left\{\frac{n_1^{-1}\sum_{j=1}^{n}A_j\mathrm{d}N_j(u)}{U_1^{(0)}(\hat{\beta}_1,u)} - \frac{n_1^{-1}\sum_{j=1}^{n}A_j\mathrm{d}N_j(u)}{U_1^{(0)}(\beta_1,u)}\right\} \\
&\quad + \int_{U_i}^{t}\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}\left\{\frac{n_1^{-1}\sum_{j=1}^{n}A_j\mathrm{d}N_j(u)}{U_1^{(0)}(\beta_1,u)} - \lambda_1(u)\mathrm{d}u\right\} \\
&= -\left[\int_{C_i}^{t}\delta^{\mathbf{1}(R_i=2)}e^{\beta_1^{\mathrm{T}}X_i}\frac{U_1^{(1)}(\beta_1,u)}{\left\{U_1^{(0)}(\beta_1,u)\right\}^2}\left\{n_1^{-1}\sum_{j=1}^{n}\mathrm{d}N_j(u)\right\}\right]^{\mathrm{T}}\left(\hat{\beta}_1-\beta_1\right)
\end{aligned}
$$

$$
\begin{aligned}
&+\int_{U_i}^{t} \delta^{\mathbf{1}(R_i=2)} e^{\beta_1^{\mathrm{T}} X_i} \frac{n_1^{-1} \sum_{j=1}^{n} A_j \mathrm{d}M_{1,j}(u)}{U_1^{(0)}(\beta_1, u)} + o_p(1) \\
= \quad & -\left\{\int_{U_i}^{t} \delta^{\mathbf{1}(R_i=2)} e^{\beta_1^{\mathrm{T}} X_i} e_1(\beta_1, u)\lambda_1(u)\mathrm{d}u\right\}^{\mathrm{T}} \left(\hat{\beta}_1 - \beta_1\right) \\
&+\int_{U_i}^{t} \delta^{\mathbf{1}(R_i=2)} e^{\beta_1^{\mathrm{T}} X_i} \frac{n_1^{-1} \sum_{j=1}^{n} A_j \mathrm{d}M_{1,j}(u)}{U_1^{(0)}(\beta_1, u)} + o_p(1),
\end{aligned} \tag{S11}
$$

where $\mathrm{d}M_{1,j}(u)$ is defined in (S2). Denote

$$
\begin{aligned}
g_{a,0}(t) &= \mathbb{E}\left[\mathbf{1}(A_i = a)\{1 - Y_i(t)\}(1 - I_i) S_a\{t \mid H_i(t); \theta\} \delta^{\mathbf{1}(R_i=2)} e^{\beta_a^{\mathrm{T}} X_i}\right], \\
g_{a,1}(t) &= \mathbb{E}\left[\mathbf{1}(A_i = a)\{1 - Y_i(t)\}(1 - I_i) S_a\{t \mid H_i(t); \theta\} \left\{\int_{U_i}^{t} \delta^{\mathbf{1}(R_i=2)} e^{\beta_a^{\mathrm{T}} X_i} X_i \lambda_a(u)\mathrm{d}u\right\}\right], \\
g_{a,2}(t) &= \mathbb{E}\left[\mathbf{1}(A_i = a)\{1 - Y_i(t)\}(1 - I_i) S_a\{t \mid H_i(t); \theta\} \left\{\int_{U_i}^{t} \delta^{\mathbf{1}(R_i=2)} e^{\beta_a^{\mathrm{T}} X_i} e_a(\beta_a, u)\lambda_a(u)\mathrm{d}u\right\}\right],
\end{aligned}
$$

for $a = 0, 1$.

Plugging (S11) in (S8) becomes

$$
\begin{aligned}
& \frac{1}{n_1^{1/2}} \sum_{i=1}^{n_1} A_i\{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \hat{\theta}\} \\
= \quad & \frac{1}{n_1^{1/2}} \sum_{i=1}^{n_1} A_i\{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \theta\} \\
& + \{g_{1,2}(t) - g_{1,1}(t)\}^{\mathrm{T}} n_1^{1/2} \left(\hat{\beta}_1 - \beta_1\right) - n_1^{-1/2} \sum_{j=1}^{n} \int_{U_j}^{t} \frac{g_{1,0}(u)}{s_0(\beta_1, u)} A_j \mathrm{d}M_{1,j}(u) + o_p(1) \\
= \quad & \frac{1}{n_1^{1/2}} \sum_{i=1}^{n_1} A_i\{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \theta\} \\
& + \frac{1}{n_1^{1/2}} \sum_{i=1}^{n_1} \left[\{g_{1,2}(t) - g_{1,1}(t)\}^{\mathrm{T}} \Gamma_1^{-1} A_i H_{1,i} - \int_{U_i}^{t} \frac{g_{1,0}(u)}{s_0(\beta_1, u)} A_i \mathrm{d}M_{1,i}(u)\right] + o_p(1),
\end{aligned} \tag{S12}
$$

where the second equality follows by (S1).

Combining (S6) and (S12) leads to

$$
\begin{aligned}
& n^{1/2}\left\{\hat{S}_{1,\mathrm{mi}}(t) - S_1^{\delta\text{-adj}}(t)\right\} \\
= \quad & \frac{n^{1/2}}{m n_1} \sum_{j=1}^{m} \sum_{i=1}^{n_1} \left[A_i\{1 - Y_i(t)\}\{\mathbf{1}(T_i^{*(j)} \geq t) - S_1(t \mid O_i; \hat{\theta}_1)\}\right] \\
& + \frac{n^{1/2}}{n_1} \sum_{i=1}^{n_1} A_i \left[\phi_{11,i}(t) + Y_i(t) + \{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \theta\} - S_1^{\delta\text{-adj}}(t)\right]
\end{aligned} \tag{S13}
$$

$$+o_p(1).$$

where

$$\phi_{11,i}(t) = \{g_{1,2}(t) - g_{1,1}(t)\}^{\mathrm{T}} \Gamma_1^{-1} H_{1,i} - \int_{U_i}^{t} \frac{g_{1,0}(u)}{u_0(\beta_1, u)} \mathrm{d}M_{1,i}(u). \tag{S14}$$

Similarly, for treatment group $a = 0$, define

$$\phi_{0,i}(t) = \{g_{0,2}(t) - g_{0,1}(t)\}^{\mathrm{T}} \Gamma_0^{-1} H_{0,i} - \int_{U_i}^{t} \frac{g_{0,0}(u)}{u_0(\beta_0, u)} \mathrm{d}M_{0,i}(u), \tag{S15}$$

We have

$$\begin{aligned} & n^{1/2}\left\{\hat{S}_{0,\mathrm{mi}}(t) - S_0^{\delta\text{-adj}}(t)\right\} \\ = \; & \frac{n^{1/2}}{mn_0}\sum_{j=1}^{m}\sum_{i=1}^{n}(1-A_i)\{1-Y_i(t)\}[\mathbf{1}(T_i^{*(j)} \geq t) - S_0\{t \mid H_i(t); \hat{\theta}\}] \\ & + \frac{n^{1/2}}{n_0}\sum_{i=1}^{n}(1-A_i)\left\{\phi_{0,i}(t) + Y_i(t) + \{1-Y_i(t)\}(1-I_i)S_0\{t \mid H_i(t); \theta\} - S_0^{\delta\text{-adj}}(t)\right\} \\ & + o_p(1). \end{aligned} \tag{S16}$$

The martingale series approximation of $\hat{\Delta}_{\tau,\mathrm{mi}}$ follows by plugging (S13) and (S16) into

$$\begin{aligned} n^{1/2}\left(\hat{\Delta}_{\tau,\mathrm{mi}} - \Delta_\tau^{\delta\text{-adj}}\right) &= n^{1/2}\left[\Psi_\tau\{\hat{S}_{1,\mathrm{mi}}(t), \hat{S}_{0,\mathrm{mi}}(t)\} - \Delta_\tau^{\delta\text{-adj}}\right] \\ &= \sum_{a=0}^{1}\int_0^\tau \psi_a(t)\left\{\hat{S}_{a,\mathrm{mi}}(t) - S_a(t)\right\}\mathrm{d}t + o_p(1) = \sum_{k=1}^{(1+m)n}\xi_{n,k} + o_p(1), \end{aligned}$$

where the $\xi_{n,k}$ terms are given in (13) with $\phi_{10,i}(t) = 0$ and $\phi_{11,i}(t)$ and $\phi_{0,i}(t)$ given in (S14) and (S15), respectively.

S4.2 Control-based Cox models

We focus on the treatment group $a = 1$. Under the control-based imputation model, the MI estimator $\hat{S}_{1,\mathrm{mi}}(t)$ depends on not only the parameter estimator in the treatment group but also the parameter estimator in the control group. Following the same steps for (S7), we express the MI estimator as

$$n^{1/2}\left\{\hat{S}_{1,\mathrm{mi}}(t) - S_1^{\delta\text{-cb}}(t)\right\}$$

$$= \frac{n^{1/2}}{mn_1}\sum_{j=1}^{m}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}[\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t);\hat{\theta}\}] \tag{S17}$$

$$+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\left[Y_i(t) + \{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\} - S_1^{\delta\text{-cb}}(t)\right], \tag{S18}$$

where under the imputation based on the control-based Cox model,

$$S_1\{t \mid H_i(t);\theta\} = \begin{cases} \exp\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}\mathrm{d}u\right\}, & \text{if } A_i=1, R_i=1, \\ \exp\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}\mathrm{d}u\right\}, & \text{if } A_i=1, R_i=2, \end{cases}$$

for $t \geq U_i$.

By the counting process theory, we can further express $n^{1/2}n_1^{-1}\sum_{i=1}^{n}A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\}$ in (S18) as

$$\begin{aligned}
& \frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)S_1\{t \mid H_i(t);\hat{\theta}\} \\
= & \frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)\exp\left\{-\int_{U_i}^{t}\hat{\lambda}_1(u)e^{\hat{\beta}_1^{\mathrm{T}}X_i}\mathrm{d}u\right\} \\
& +\frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)\exp\left\{-\delta\int_{U_i}^{t}\hat{\lambda}_0(u)e^{\hat{\beta}_0^{\mathrm{T}}X_i}\mathrm{d}u\right\} \\
= & \frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)\exp\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}\mathrm{d}u\right\} \\
& +\frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)\exp\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}\mathrm{d}u\right\} \\
& +\frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)S_1\{t \mid H_i(t);\theta\}\left[-\int_{U_i}^{t}e^{\beta_1^{\mathrm{T}}X_i}\left\{\hat{\lambda}_1(u)-\lambda_1(u)\right\}\mathrm{d}u\right] \\
& +\left[\frac{1}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)S_1\{t \mid H_i(t);\theta\}\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}X_i\mathrm{d}u\right\}\right] \\
& \times n^{1/2}\left(\hat{\beta}_1-\beta_1\right) \\
& +\frac{n^{1/2}}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)S_1\{t \mid H_i(t);\theta\}\left[-\delta\int_{U_i}^{t}e^{\beta_0^{\mathrm{T}}X_i}\left\{\hat{\lambda}_0(u)-\lambda_0(u)\right\}\mathrm{d}u\right] \\
& +\left[\frac{1}{n_1}\sum_{i=1}^{n} A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)S_1\{t \mid H_i(t);\theta\}\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}X_i\mathrm{d}u\right\}\right] \\
& \times n^{1/2}\left(\hat{\beta}_0-\beta_0\right). \qquad \text{(S19)}
\end{aligned}$$

Denote

$$
\begin{aligned}
\tilde{g}_{1,0}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)S_1\{t\mid H_i(t);\theta\}e^{\beta_1^{\mathrm{T}}X_i}\right],\\
\tilde{g}_{1,1}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)S_1\{t\mid H_i(t);\theta\}\left\{\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}X_i^{\mathrm{T}}\mathrm{d}u\right\}\right],\\
\tilde{g}_{1,2}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)S_1\{t\mid H_i(t);\theta\}\left\{\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}e_1(\beta_1,u)^{\mathrm{T}}\mathrm{d}u\right\}\right],\\
\tilde{g}_{0,0}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)S_1\{t\mid H_i(t);\theta\}\delta e^{\beta_0^{\mathrm{T}}X_i}\right],\\
\tilde{g}_{0,1}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)S_1\{t\mid H_i(t);\theta\}\left\{\delta\int_{U_i}^{t}\lambda_0(u)e^{\delta\beta_0^{\mathrm{T}}X_i}X_i^{\mathrm{T}}\mathrm{d}u\right\}\right],\\
\tilde{g}_{0,2}(t) &= \mathbb{E}\left[A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)S_1\{t\mid H_i(t);\theta\}\left\{\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}e_0(\beta_0,u)^{\mathrm{T}}\mathrm{d}u\right\}\right].
\end{aligned}
$$

Then, we can express (S19) further as

$$
\begin{aligned}
&= \frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)\exp\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}\mathrm{d}u\right\}\\
&\quad+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)\exp\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}\mathrm{d}u\right\}\\
&\quad+\{\tilde{g}_{1,2}(t)-\tilde{g}_{1,1}(t)\}^{\mathrm{T}}n^{1/2}\left(\hat{\beta}_1-\beta_1\right)-\frac{n^{1/2}}{n_1}\sum_{j=1}^{n}\int_{U_i}^{t}\frac{\tilde{g}_{1,0}(u)}{s_0(\beta_1,u)}A_j\mathrm{d}M_{1,j}(u)\\
&\quad+\{\tilde{g}_{0,2}(t)-\tilde{g}_{0,1}(t)\}^{\mathrm{T}}n^{1/2}\left(\hat{\beta}_0-\beta_0\right)-\frac{n^{1/2}}{n_1}\sum_{j=1}^{n}\int_{U_i}^{t}\frac{\tilde{g}_{0,0}(u)}{s_0(\beta_0,u)}(1-A_j)\mathrm{d}M_{0,j}(u)+o_p(1)\\
&= \frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)\exp\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}\mathrm{d}u\right\}\\
&\quad+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)\exp\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}\mathrm{d}u\right\}\\
&\quad+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n}\left[\{\tilde{g}_{1,2}(t)-\tilde{g}_{1,1}(t)\}^{\mathrm{T}}\Gamma_1^{-1}A_iH_{1,i}-\int_{U_i}^{t}\frac{\tilde{g}_{1,0}(u)}{s_0(\beta_1,u)}A_i\mathrm{d}M_{1,i}(u)\right]\\
&\quad+\frac{n^{1/2}}{n_0}\sum_{i=1}^{n}\left[\{\tilde{g}_{0,2}(t)-\tilde{g}_{0,1}(t)\}^{\mathrm{T}}\Gamma_0^{-1}(1-A_i)H_{0,i}-\int_{U_i}^{t}\frac{\tilde{g}_{0,0}(u)}{s_0(\beta_0,u)}(1-A_i)\mathrm{d}M_{0,i}(u)\right]+o_p(1)\\
&= \frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=1)\exp\left\{-\int_{U_i}^{t}\lambda_1(u)e^{\beta_1^{\mathrm{T}}X_i}\mathrm{d}u\right\}\\
&\quad+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n_1}A_i\{1-Y_i(t)\}(1-I_i)\mathbf{1}(R_i=2)\exp\left\{-\delta\int_{U_i}^{t}\lambda_0(u)e^{\beta_0^{\mathrm{T}}X_i}\mathrm{d}u\right\}
\end{aligned}
$$

$$+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n}A_i\left[\{\tilde{g}_{1,2}(t)-\tilde{g}_{1,1}(t)\}^{\mathrm{T}}\Gamma_1^{-1}H_{1,i}-\int_{U_i}^{t}\frac{\tilde{g}_{1,0}(u)}{s_0(\beta_1,u)}\mathrm{d}M_{1,i}(u)\right]$$
$$+\frac{n^{1/2}}{n_0}\sum_{i=1}^{n}(1-A_i)\left[\{\tilde{g}_{0,2}(t)-\tilde{g}_{0,1}(t)\}^{\mathrm{T}}\Gamma_0^{-1}H_{0,i}-\int_{U_i}^{t}\frac{\tilde{g}_{0,0}(u)}{s_0(\beta_0,u)}\mathrm{d}M_{0,i}(u)\right]+o_p(1). \quad \text{(S20)}$$

Combining (S17) and (S20) leads to

$$\begin{aligned}
&n^{1/2}\left\{\hat{S}_{1,\mathrm{mi}}(t)-S_1^{\delta\text{-cb}}(t)\right\}\\
=\ &\frac{n^{1/2}}{mn_1}\sum_{j=1}^{m}\sum_{i=1}^{n}A_i\{1-Y_i(t)\}\left[\mathbf{1}(T_i^{*(j)}\geq t)-S_1\{t\mid H_i(t);\hat{\theta}\}\right]+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n}(1-A_i)\phi_{10,i}(t)\\
&+\frac{n^{1/2}}{n_1}\sum_{i=1}^{n}A_i\left[\phi_{11,i}(t)+Y_i(t)+\{1-Y_i(t)\}(1-I_i)S_1\{t\mid H_i(t);\theta\}-S_1^{\delta\text{-cb}}(t)\right]+o_p(1),
\end{aligned}$$

where

$$\phi_{11,i}(t) = \left[\{\tilde{g}_{1,2}(t)-\tilde{g}_{1,1}(t)\}^{\mathrm{T}}\Gamma_1^{-1}H_{1,i}-\int_{U_i}^{t}\frac{\tilde{g}_{1,0}(u)}{s_0(\beta_1,u)}\mathrm{d}M_{1,i}(u)\right] \quad \text{(S21)}$$
$$\phi_{10,i}(t) = \left[\{\tilde{g}_{0,2}(t)-\tilde{g}_{0,1}(t)\}^{\mathrm{T}}\Gamma_0^{-1}H_{0,i}-\int_{U_i}^{t}\frac{\tilde{g}_{0,0}(u)}{s_0(\beta_0,u)}\mathrm{d}M_{0,i}(u)\right]. \quad \text{(S22)}$$

Because the imputation mechanism for the censored control subjects is the same, the martingale representation for $\hat{S}_{0,\mathrm{mi}}(t)$ remains the same as in (S16). Finally, we can decompose $\hat{\Delta}_{\tau,\mathrm{mi}}$ by the martingale representation

$$n^{1/2}(\hat{\Delta}_{\tau,\mathrm{mi}}-\Delta_\tau^{\delta\text{-bc}})=\sum_{k=1}^{(1+m)n}\xi_{n,k}+o_p(1),$$

where the $\xi_{n,k}$ terms are given in (13) with $\phi_{11,i}(t)$, $\phi_{10,i}(t)$, and $\phi_{0,i}(t)$ given in (S21), (S22) and (S15), respectively.

For both the δ-adjusted and control-based Cox models, it follows by the martingale CLT, $n^{1/2}(\hat{\Delta}_{\tau,\mathrm{mi}}-\Delta_\tau^{\mathrm{sen}})$ converges to a Normal distribution with mean zero and a finite variance

$$V_{\tau,\mathrm{mi}}^{\mathrm{sen}}=\sum_{k=1}^{(1+m)n}\mathbb{E}(\xi_{n,k}^2\mid\mathcal{F}_{n,k-1})=\sum_{a=0}^{1}\left(\sigma_{a,1}^2+\sigma_{a,2}^2\right), \quad \text{(S23)}$$

where sen denotes either δ-adj or δ-cb, and

$$
\begin{aligned}
\sigma_{0,1}^2 &= \frac{1}{p_0}\mathbb{E}\left(\left[\int_0^\tau \psi_0(t)\{(1-A_i)[\phi_{10,i}(t)+\phi_{0,i}(t)+Y_i(t)\right.\right. \\
&\quad \left.\left. +\{1-Y_i(t)\}(1-I_i)S_a\{t\mid H_i(t);\theta\}]-S_a^{\text{sen}}(t)\}\mathrm{d}t\right]^2\right) \\
\sigma_{1,1}^2 &= \frac{1}{p_1}\mathbb{E}\left\{\left(\int_0^\tau \psi_1(t)A_i[\phi_{11,i}(t)+Y_i(t)\right.\right. \\
&\quad \left.\left. +\{1-Y_i(t)\}(1-I_i)S_a\{t\mid H_i(t);\theta\}-S_a^{\text{sen}}(t)]\mathrm{d}t\right)\right\}, \\
\sigma_{a,2}^2 &= \frac{1}{p_a m}\mathbb{V}\left[\int_0^\tau \psi_a(t)\mathbf{1}(A_i=a)\{1-Y_i(t)\}\{\mathbf{1}(T_i^{*(j)}\geq t)-S_a\{t\mid H_i(t);\theta\}\}\mathrm{d}t\right],
\end{aligned}
$$

for $a=0,1$.

Web Appendix S5 Remarks

Remark 1 *There are many choices for generating μ_k, such as the the standard normal distribution, Mammen's two point distribution, a simpler distribution with probability 0.5 of being 1 and probability 0.5 of being -1, or the nonparametric bootstrap weights. The wild bootstrap procedure is not sensitive to the choice of the sampling distribution of μ_k. We adopt the standard normal distribution in the simulation study.*

Remark 2 *It is worth discussing the connection between the martingale representation (10) and existing results in the survival literature. Under CCAR, Zhao et al. (2016) derived an asymptotic linearization for the RMST estimator and proposed the perturbation-resampling variance estimation by adding independent noises to the linearized terms. In this simpler case, by setting the sensitivity parameter δ to be 1 and omitting the imputation step, our martingale representation with the first n_1 terms reduces to their linearization. The slight difference lies in the distribution for generating the resampling weights. In the wild bootstrap, the resampling weight distribution has mean 1; while in the perturbation, the resampling weight distribution has mean 0. The difference would only affect the center of the bootstrap replicates of $\hat{\Delta}_{\tau,\text{mi}}$ but not the variability and thus variance estimation. Our framework allows for CAR and sensitivity analysis using δ-adjustment/control-based models, taking into account variability from both parameter estimation and imputation.*

Web Appendix S6 Proof of Theorem 2

We provide the proof of Theorem 2, which draws on the martingale central limit theory (Hall and Heyde, 1980) and the asymptotic property of weighted sampling of martingale difference arrays (Pauly, 2011).

First, by the law of large numbers, we have

$$
\begin{aligned}
&\sum_{k=1}^{n_1} \xi_{n,k}^2 \\
= \;& \frac{n}{n_1^2} \sum_{i=1}^{n_1} \left(\int_0^\tau \psi_1(t) A_i \left[\phi_{11,i}(t) + Y_i(t) + \{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \theta\} - S_1^{\text{sen}}(t) \right] \mathrm{d}t \right)^2 \\
\xrightarrow{p} \;& \frac{1}{p_1} \mathbb{E} \left\{ \left(\int_0^\tau \psi_1(t) A_i \left[\phi_{11,i}(t) + Y_i(t) + \{1 - Y_i(t)\}(1 - I_i) S_1\{t \mid H_i(t); \theta\} - S_1^{\text{sen}}(t) \right] \mathrm{d}t \right)^2 \right\} \\
= \;& \sigma_{1,1}^2,
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{k=n_1+1}^{(1+m)n_1} \xi_{n,k}^2 \\
= \;& \frac{n}{n_1^2} \sum_{i=1}^{n_1} \frac{1}{m^2} \sum_{j=1}^{m} \left[\int_0^\tau \psi_1(t) A_i \{1 - Y_i(t)\} [\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t); \hat{\theta}\}] \mathrm{d}t \right]^2 \\
\xrightarrow{p} \;& \frac{1}{p_1 m} \mathbb{E} \left(\text{var} \left[\int_0^\tau \psi_1(t) A_i \{1 - Y_i(t)\} [\mathbf{1}(T_i^{*(j)} \geq t) - S_1\{t \mid H_i(t); \hat{\theta}\}] \mathrm{d}t \mid O_{1:n} \right] \right) \\
= \;& \sigma_{1,2}^2,
\end{aligned}
$$

as $n \to \infty$. Similarly, by the law of large numbers, we have $\sum_{k=(1+m)n_1+1}^{(1+m)n_1+n_0} \xi_{n,k}^2 \xrightarrow{p} \sigma_{0,1}^2$, and $\sum_{k=(1+m)n_1+n_0+1}^{(1+m)n} \xi_{n,k}^2 \xrightarrow{p} \sigma_{0,2}^2$. Therefore, we have

$$
\sum_{k=1}^{(1+m)n} \xi_{n,k}^2 \xrightarrow{p} V_{\tau,\text{mi}}^{\text{sen}}, \tag{S24}
$$

as $n \to \infty$.

Second, we show

$$
\max_{1 \leq k \leq (1+m)n} |\xi_{n,k}| \xrightarrow{p} 0, \tag{S25}
$$

as $n \to \infty$. Toward this end, for any $\epsilon > 0$,

$$
\mathbb{P} \left(\max_{1 \leq k \leq n_1} |\xi_{n,k}| > \epsilon \right) \;\; \leq \;\; n_1 \mathbb{P} \left(|\xi_{n,k}| > \epsilon \right) = n_1 \mathbb{P} \left(\xi_{n,k}^4 > \epsilon^4 \right)
$$

$$\leq \quad \frac{n^2}{n_1^3\epsilon^4}\mathbb{E}\left(\int_0^\tau \psi_1(t)A_i\left[S_1\{t \mid H_i(t);\hat{\theta}\} - S_1^{\text{sen}}(t)\right]\mathrm{d}t\right)^4 \to 0,$$

where the second inequality follows from the Markov inequality, and the convergence follows because the expectation term is bounded due to the natural range of the survival functions. Similarly, we have

$$\mathbb{P}\left(\max_{n_1+1\leq k\leq(1+m)n_1}|\xi_{n,k}| > \epsilon\right) \leq \frac{n^2}{n_1^3m^3\epsilon^4}\mathbb{E}\left\{\int_0^\tau \psi_1(t)A_i[\mathbf{1}(T_i^{*(j)}\geq t) - S_1\{t \mid H_i(t);\hat{\theta}\}]\mathrm{d}t\right\}^4 \to 0,$$

as $n\to\infty$. Therefore, $\mathbb{P}(\max\limits_{1\leq k\leq(1+m)n_1}|\xi_{n,k}| > \epsilon) \to 0$, as $n\to\infty$. Similarly, $\mathbb{P}(\max\limits_{(1+m)n_1+1\leq k\leq(1+m)n}|\xi_{n,k}| > \epsilon) \to 0$, as $n\to\infty$. Then (S25) holds.

Third, we show

$$\sup_n \mathbb{E}\left(\max_{1\leq k\leq(1+m)n}\xi_{n,k}^2\right) < \infty. \tag{S26}$$

For any n, by Assumption S1,

$$\begin{aligned}\mathbb{E}\left(\max_{1\leq k\leq n_1}\xi_{n,k}^2\right) &\leq \mathbb{E}\left(n_1\xi_{n,k}^2\right)\\ &= \frac{n}{n_1}\mathbb{E}\left(\int_0^\tau \left[\psi_1(t)A_iS_1\{t \mid H_i(t);\hat{\theta}\} - S_1^{\text{sen}}(t)\right]\mathrm{d}t\right)^2 < \infty,\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left(\max_{n_1+1\leq k\leq(1+m)n_1}\xi_{n,k}^2\right) &\leq \mathbb{E}\left(nm\xi_{n,k}^2\right)\\ &= \frac{n}{mn_1}\mathbb{E}\left(\int_0^\tau \psi_1(t)A_i[\mathbf{1}(T_i^{*(j)}\geq t) - S_1\{t \mid H_i(t);\hat{\theta}\}]\mathrm{d}t\right)^2 < \infty.\end{aligned}$$

Therefore, $\mathbb{E}(\max_{1\leq k\leq(1+m)n_1}\xi_{n,k}^2) \leq \mathbb{E}(\max_{1\leq k\leq n_1}\xi_{n,k}^2) + \mathbb{E}(\max_{n_1+1\leq k\leq(1+m)n_1}\xi_{n,k}^2) < \infty$. Similarly, $\mathbb{E}(\max_{n_1(1+m)+1\leq k\leq n(1+m)}\xi_{n,k}^2) < \infty$. Then (S26) follows.

Given the results in (S24) and (S25), the martingale CLT implies that

$$\sum_{k=1}^{(1+m)n}\xi_{n,k} \xrightarrow{d} \mathcal{N}(0, V_{\tau,\text{mi}}^{\text{sen}}),$$

as $n \to \infty$. Given the results in (S24), (S25), and (S26), Theorem 2.1 in Pauly (2011) yields

$$\sup_r \left| \mathbb{P}\left\{ \{(1+m)n\}^{1/2} \sum_{k=1}^{(1+m)n} \frac{u_k}{\{n(1+m)\}^{1/2}} \xi_{n,k} \le r \,\middle|\, O_{1:n} \right\} - \Phi\left(\frac{r}{\sigma}\right) \right| \xrightarrow{p} 0, \tag{S27}$$

as $n \to \infty$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

Let $W_L = n^{-1/2} \sum_{k=1}^{(1+m)n} \xi_{n,k} u_k$. By Theorem 1 and (S27), we have

$$\sup_r \left| \mathbb{P}\left(n^{1/2} W_L \le r \mid O_{1:n}\right) - \mathbb{P}\left\{ n^{1/2}\left(\hat{\Delta}_{\tau,\mathrm{mi}} - \Delta_\tau^{\mathrm{sen}}\right) \le r \right\} \right| \xrightarrow{p} 0, \tag{S28}$$

as $n \to \infty$.

Lastly, to prove Theorem 2, it remains to show that

$$\mathbb{P}\left\{ n^{1/2}(W_L - W_L^*) \mid O_{1:n} \right\} \xrightarrow{p} 0, \tag{S29}$$

as $n \to \infty$. To unify the notation for both treatment group, define $\Phi_{1,i}(t) = \phi_{11,i}(t)$, $\Phi_{0,i}(t) = \phi_{10,i}(t) + \phi_{0,i}(t)$, $\hat{\Phi}_{1,i}(t) = \hat{\phi}_{11,i}(t)$, and $\hat{\Phi}_{0,i}(t) = \hat{\phi}_{10,i}(t) + \hat{\phi}_{0,i}(t)$. The difference between W_L and W_L^* can be decomposed to six parts,

$$n^{1/2}(W_L - W_L^*) = \sum_{k=1}^{n(1+m)} n^{-1/2} u_k (n^{1/2}\hat{\xi}_{n,k} - n^{1/2}\xi_{n,k}) = \sum_{a=0}^{1}\sum_{l=1}^{3} R_{al,n},$$

where

$$\begin{aligned}
R_{a1,n} &= \sum_{i=1}^{n} \frac{n^{1/2}}{n_a} u_i \mathbf{1}(A_i = a) \int_0^\tau \psi_a(t) \left\{ \hat{S}_{a,\mathrm{mi}}(t) - S_a^{\mathrm{sen}}(t) \right\} \mathrm{d}t, \\
R_{a2,n} &= \sum_{i=1}^{n} \frac{n^{1/2}}{n_a} u_i \mathbf{1}(A_i = a) \int_0^\tau \psi_a(t) \left\{ \hat{\Phi}_{a,i}(t) - \Phi_{a,i}(t) \right\} \mathrm{d}t, \\
R_{a3,n} &= \sum_{i=1}^{n} \frac{n^{1/2}}{n_a} u_i \mathbf{1}(A_i = a) \\
&\quad \times \int_0^\tau \psi_a(t) \{1 - Y_i(t)\}(1 - I_i) \left[S_a\{t \mid H_i(t); \hat{\theta}\} - S_a\{t \mid H_i(t); \theta\} \right] \mathrm{d}t,
\end{aligned}$$

for $a = 0, 1$.

Given that the bootstrap weights satisfy $\mathbb{E}(u_k^2 \mid O_{1:n}) = 1$, we have

$$\begin{aligned}\mathbb{E}\left(R_{a1,n}^2 | O_{1:n}\right) &= \frac{n}{n_a^2} n_a \mathbb{E}(u_i^2) \left[\int_0^\tau \psi_a(t)\left\{\hat{S}_{a,\mathrm{mi}}(t) - S_a^{\mathrm{sen}}(t)\right\} \mathrm{d}t\right]^2 \\ &= \frac{n}{n_a}\left[\int_0^\tau \psi_a(t)\left\{\hat{S}_{a,\mathrm{mi}}(t) - S_a^{\mathrm{sen}}(t)\right\} \mathrm{d}t\right]^2 \xrightarrow{p} 0,\end{aligned}$$

as $n \to \infty$, for $a = 0, 1$. Also, we have

$$\mathbb{E}\left(R_{a2,n}^2 | O_{1:n}\right) = \frac{n}{n_a^2}\sum_{i=1}^{n} \mathbf{1}(A_i = a)\left[\int_0^\tau \psi_a(t)\left\{\hat{\Phi}_{a,i}(t) - \Phi_{a,i}(t)\right\} \mathrm{d}t\right]^2 \xrightarrow{p} 0,$$

as $n \to \infty$, for $a = 0, 1$, where the convergence follows by Assumption S1 and the results in Section Web Appendix S1. Similarly, we have

$$\begin{aligned}&\mathbb{E}\left(R_{a3,n}^2 | O_{1:n}\right) \\ &= \frac{n}{n_a^2}\sum_{i=1}^{n_a} \mathbf{1}(A_i = a)\left[\int_0^\tau \psi_a(t)\{1 - Y_i(t)\}(1 - I_i)\ \left\{S_a(t \mid O_i; \hat{\theta}_a) - S_a\{t \mid H_i(t); \theta\}\right\} \mathrm{d}t\right]^2 \xrightarrow{p} 0,\end{aligned}$$

as $n \to \infty$, for $a = 0, 1$. Therefore, for any $\epsilon > 0$,

$$\mathbb{P}\{|R_{a1,n}| > \epsilon \mid O_{1:n}\} \xrightarrow{p} 0, \quad \mathbb{P}\{|R_{a2,n}| > \epsilon \mid O_{1:n}\} \xrightarrow{p} 0, \quad \mathbb{P}\{|R_{a3,n}| > \epsilon \mid O_{1:n}\} \xrightarrow{p} 0,$$

as $n \to \infty$, for $a = 0, 1$. Then we obtain (S29). The conclusion of Theorem 2 follows.

Web Appendix S7 Simulation study

We conduct simulation studies to evaluate the finite sample performance of the proposed SMIM framework. For illustration, we focus on the δ-adjusted and control-based models for sensitivity analysis and the RMST as the treatment effect estimand. We start with a simple setup with one covariate in Section S7.1 and then consider a setting motivated by the ACTG175 trial data in Section S7.2.

S7.1 Simulation one: a simple setup

For both the treatment and control groups, each with sample size $n \in \{500, 1000\}$, the confounder is generated by $X_i \sim \mathcal{N}(0, 1)$. In the treatment group, T follows the Cox model with the hazard rate $\lambda_1(t \mid X_i) = \lambda_1(t)\exp(\beta_1 X_i)$, where $\lambda_1(t) = 0.35$ and $\beta_1 = 0.75$. We consider censoring due to the end of the study and premature dropout. We generate the censoring time to dropout, C_i, according to a Cox model with the hazard rate $\lambda_C(t \mid X_i) = \lambda_C(t)\exp(\beta_C X_i)$, where $\lambda_C(t) = 0.15$ and $\beta_C = 0.75$. The maximum follow up time is $L = 3.25$. The observed time is $U_i = T_i \wedge C_i \wedge L$. If $U_i = T_i$, the event indicator is $I_i = 1$; if $U_i = L$, then $I_i = 0$ and the censoring type is $R_i = 1$; if $U_i = C_i$, then $I_i = 0$ and the censoring type is $R_i = 2$. Under the data generating mechanism, the average percentages of $I_i = 1$, $R_i = 1$, and $R_i = 2$ are 53%, 25%, and 22%, respectively. In the control group, T_i follows the hazard rate $\lambda_0(t \mid X) = \lambda_0(t)\exp(\beta_0 X_i)$, where $\lambda_0(t) = 0.40$ and $\beta_0 = 0.75$. The censoring time C_i follows the same model as in the treatment group. For the dropout subjects with $R_i = 2$ in treatment group, the hazard rate for events after censoring are $\delta\lambda_1(t)\exp(\beta_1 X_i)$ for delta-adjusted model and $\lambda_0(t)\exp(\beta_0 X_i)$ for control-based models. For the dropout subjects with $R_i = 2$ in control group, the hazard rate for event after censoring remains the same, which correspondsto the case when the control treatment is a placebo or the standard of care. The true RMST estimand under the δ-adjusted model is $\Delta_\tau^{\delta\text{-adj}} = \mu_{1,\tau}^{\delta\text{-adj}} - \mu_{0,\tau}$ with $\tau = 3$. We assess the proposed method to implement the sensitivity analysis for the treatment group when the true parameter δ is 1.5, while the analysis parameter δ varies in a pre-specified set $\{0.5, 1, 1.5, 2, 2.5\}$. The true RMST estimand under the control-based model are $\Delta_\tau^{\text{control-adj}}$ with $\tau = 3$.

We use MI for imputing the censored event times following Steps MI-1-1, MI-1-2 and MI-1-3 in Section 3 with imputation size $m \in \{10, 20, 50\}$. We compare the standard MI inference and the proposed wild bootstrap inference. For the standard MI inference, the $100(1-\alpha)\%$ confidence intervals are calculated as $(\hat{\Delta}_{\tau,\text{mi}} - z_{1-\alpha/2}\hat{V}_{\text{mi}}^{1/2}, \hat{\Delta}_{\tau,\text{mi}} + z_{1-\alpha/2}\hat{V}_{\text{mi}}^{1/2})$, where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$th quantile of the standard normal distribution. For the proposed wild bootstrap procedure, we sample the weights μ_k from the standard normal distribution, and calculate the variance estimate $\hat{V}_{\text{WB}}$ based on 100 replications. The corresponding $100(1-\alpha)\%$ confidence intervals are calculated as $(\hat{\Delta}_{\tau,\text{mi}} - z_{1-\alpha/2}\hat{V}_{\text{WB}}^{1/2}, \hat{\Delta}_{\tau,\text{mi}} + z_{1-\alpha/2}\hat{V}_{\text{WB}}^{1/2})$. We assess the performance in terms of the relative bias of the variance estimator and the coverage rate of confidence intervals. The relative bias of the variance estimators are calculated as $\{\mathbb{E}(\hat{V}_{\text{mi}}^{1/2}) - \mathbb{V}(\hat{\Delta}_{\tau,\text{mi}}^{1/2})\}/\mathbb{V}(\hat{\Delta}_{\tau,\text{mi}}^{1/2}) \times 100\%$ and $\{\mathbb{E}(\hat{V}_{\text{WB}}^{1/2}) - \mathbb{V}(\hat{\Delta}_{\tau,\text{mi}}^{1/2})\}/\mathbb{V}(\hat{\Delta}_{\tau,\text{mi}}^{1/2}) \times 100\%$. The coverage rate of the $100(1-\alpha)\%$ confidence intervals is estimated by the percentage of the Monte Carlo samples for which the confidence intervals contain the true value.

Table S1 presents the simulation results for the sensitivity analysis of δ-adjusted estimand $\Delta_\tau^{\delta\text{-adj}}$ based on 1000 Monte Carlo samples. When the imputation model is correctly specified with $\delta = 1.5$, the MI point estimator $\hat{\Delta}_{\tau,\mathrm{mi}}$ is unbiased of the true estimand $\Delta_\tau^{\delta\text{-adj}}$. When the analysis sensitivity parameter is lower (higher) than the true parameter $\delta = 1.5$, the MI point estimator produces higher (lower) RMST for the treatment group, and therefore $\hat{\Delta}_{\tau,\mathrm{mi}}$ is biased upward (downward). When the true sensitivity parameter is correctly specified, Rubin's combining rule overestimates the true standard deviation with the relative bias ranging from 7.0% to 12.2%; consequently, the coverage rates are larger than the nominal level 95%. In contrast, our proposed wild bootstrap procedure is unbiased; as a result, the coverage rates of the confidence intervals are close to the nominal level. Moreover, the proposed method is not sensitive to the number of imputations m. We observed similar behavior for the sensitivity analysis of control-based models for sensitivity analysis and summarized in Table S2.

S7.2 Simulation two: ACTG175

We consider a simulation setup that is similar to ACTG175 data. The confounder is generated by $X_{1i} \sim \mathcal{N}(0, 1)$ and $X_{2i} \sim \text{Bernoulli}(0.15)$. In the treatment group, T follows the Cox model with the hazard rate $\lambda_1(t \mid X_{1i}X_{2i}) = \lambda_1(t)\exp(\beta_1 X_{1i} + \beta_2 X_{2i})$, where $\lambda_1 = 0.03$, $\beta_1 = 0.24$ and $\beta_2 = 0.04$. We consider censoring due to the end of the study and premature dropout. We generate the censoring time to dropout, C_i, according to a Cox model with the hazard rate $\lambda_C(t \mid X_{1i}X_{2i}) = \lambda_C(t)\exp(\beta_{C1}X_{1i} + \beta_{C2}X_{2i})$, where $\lambda_C(t) = 0.01$, $\beta_{C1} = 0.24$, $\beta_{C2} = 0.20$. The maximum follow up time is $L = 40$. The observed time is $U_i = T_i \wedge C_i \wedge L$. If $U_i = T_i$, the event indicator is $I_i = 1$; if $U_i = L$, then $I_i = 0$ and the censoring type is $R_i = 1$; if $U_i = C_i$,then $I_i = 0$ and the censoring type is $R_i = 2$. Under the data generating mechanism, the average percentages of $I_i = 1$, $R_i = 1$, and $R_i = 2$ are 60%, 20% and 20%, respectively.In the control group, T_i follows the hazard rate $\lambda_0(t \mid X_{1i}, X_{2i}) = \lambda_0(t)\exp(\beta_{01}X_{1i} + \beta_{02}X_{2i})$, where $\lambda_0(t) = 0.03$, $\beta_{01} = -0.55$ and $\beta_{02} = 0.65$. The censoring time C_i follows the same model as in the treatment group. For the dropout subjects with $R_i = 2$ in treatment group, the hazard rate for events after censoring are

$\delta\lambda_1(t)\exp(\beta_1 X_{1i} + \beta_2 X_{2i})$ for delta-adjusted model and $\lambda_0(t)\exp(\beta_{01} X_{1i} + \beta_{02} X_{2i})$ for control-based models. For the dropout subjects with $R_i = 2$ in control group, the hazard rate for event after censoring remains the same, which correspondsto the case when the control treatment is a placebo or the standard of care. The true RMST estimand under the δ-adjusted model is $\Delta_\tau^{\delta\text{-adj}} = \mu_{1,\tau}^{\delta\text{-adj}} - \mu_{0,\tau}$ with $\tau = 24$. We assess the proposed method to implement the sensitivity analysis for the treatment group when the true parameter δ is 2, while the analysis parameter δ varies in a pre-specified set $\{1, 2, 3, 4, 5\}$. The true RMST estimand under the control-based model are $\Delta_\tau^{\text{control-adj}}$ with $\tau = 24$. The estimation procedure are the same as the first simulation study. The simulation results is summarized in Table S3 and Table S4 with similar observation in the first simulation study.

Table S1: Simulation results for the true estimand $\Delta_\tau^{\delta\text{-adj}} = 0.054$ with the true sensitivity parameter $\delta = 1.5$: point estimate, true standard deviation, relative bias of the standard error estimator, coverage of interval estimate using Rubin's method and the proposed wild bootstrap method

n	m	Model	Point est ($\times 10^2$)	True sd ($\times 10^2$)	Standard error ($\times 10^2$) Rubin[a]	WB	Relative Bias (%) Rubin[a]	WB	Coverage (%) for 95% CI Rubin[a]	WB
500	10	δ =0.50	15.8	6.93	7.43	6.78	7.24	-2.18	71.0	66.2
		δ =1.00	9.3	6.91	7.41	6.74	7.31	-2.43	94.3	90.8
		δ =1.50	5.0	6.89	7.38	6.74	7.11	-2.15	**97.0**	**95.1**
		δ =2.00	2.0	6.87	7.35	6.75	6.94	-1.74	94.5	92.0
		δ =2.50	-0.3	6.85	7.32	6.77	6.84	-1.30	89.3	85.7
	20	δ =0.50	15.8	6.92	7.41	6.76	7.12	-2.28	71.3	65.5
		δ =1.00	9.3	6.90	7.39	6.73	7.14	-2.53	93.9	90.3
		δ =1.50	5.1	6.88	7.36	6.73	6.99	-2.22	**96.6**	**94.9**
		δ =2.00	2.0	6.86	7.33	6.74	6.89	-1.76	94.4	91.9
		δ =2.50	-0.3	6.84	7.31	6.75	6.84	-1.28	89.4	86.0
	50	δ =0.50	15.8	6.90	7.41	6.75	7.37	-2.07	71.3	65.6
		δ =1.00	9.3	6.88	7.38	6.72	7.38	-2.32	94.1	91.0
		δ =1.50	5.0	6.86	7.35	6.72	7.22	-2.01	**96.6**	**95.0**
		δ =2.00	2.0	6.84	7.32	6.73	7.09	-1.56	94.7	91.7
		δ =2.50	-0.3	6.82	7.30	6.75	7.01	-1.10	89.3	85.9
	N/A	Tian et.al. 2014	9.4	7.10	-	7.56	-	6.40	-	92.9
1000	10	δ =0.50	16.3	4.72	5.25	4.80	11.19	1.58	45.4	37.5
		δ =1.00	9.8	4.68	5.24	4.77	11.87	1.98	87.9	84.2
		δ =1.50	5.6	4.66	5.21	4.78	11.98	2.57	**97.7**	**95.2**
		δ =2.00	2.5	4.64	5.19	4.79	12.04	3.22	94.4	91.3
		δ =2.50	0.2	4.62	5.18	4.80	12.14	3.85	85.2	80.9
	20	δ =0.50	16.3	4.71	5.25	4.79	11.39	1.76	45.0	37.8
		δ =1.00	9.8	4.67	5.23	4.77	12.02	2.08	87.9	84.6
		δ =1.50	5.6	4.64	5.21	4.77	12.14	2.68	**97.7**	**95.0**
		δ =2.00	2.5	4.62	5.19	4.78	12.20	3.35	94.1	91.5
		δ =2.50	0.2	4.61	5.17	4.79	12.28	3.97	85.7	81.8
	50	δ =0.50	16.3	4.70	5.24	4.79	11.39	1.78	45.3	37.5
		δ =1.00	9.8	4.66	5.22	4.76	12.06	2.13	88.0	84.6
		δ =1.50	5.5	4.64	5.20	4.76	12.19	2.74	**97.5**	**95.2**
		δ =2.00	2.5	4.61	5.18	4.77	12.27	3.41	94.1	91.4
		δ =2.50	0.2	4.60	5.17	4.78	12.34	4.03	85.4	81.7
	N/A	Tian et.al. 2014	9.9	4.90	-	5.35	-	9.28	-	88.2

Table S2: Simulation results for the true estimand $\Delta_{\tau}^{\text{control-adj}} = 1.783$ based on control-based method: point estimate, true standard deviation, relative bias of the standard error estimator, coverage of interval estimate using Rubin's method and the proposed wild bootstrap method

					Standard error ($\times 10^2$)		Relative Bias (%)		Coverage (%) for 95% CI	
n	Model	m	Point est ($\times 10^2$)	True sd ($\times 10^2$)	Rubin[a]	WB	Rubin[a]	WB	Rubin[a]	WB
500	Control-based	10	179.0	4.58	5.24	4.76	14.34	3.87	**97.2**	**95.1**
		20	179.0	4.58	5.22	4.75	13.95	3.62	**97.4**	**95.2**
		50	179.0	4.57	5.22	4.74	14.16	3.76	**97.3**	**95.3**
	Tian et.al. 2014	N/A	184.6	4.81	-	5.34	-	10.93	-	80.2
1000	Control-based	10	179.1	3.30	3.70	3.37	11.94	1.97	**96.6**	**94.4**
		20	179.1	3.30	3.69	3.36	12.10	2.03	**96.5**	**94.2**
		50	179.1	3.29	3.69	3.36	12.21	2.14	**96.4**	**94.5**
	Tian et.al. 2014	N/A	184.8	3.53	-	3.78	-	7.08	-	61.1

Table S3: Simulation results for the true estimand $\Delta_\tau^{\text{control-adj}} = 0.513$ based on control-based method: point estimate, true standard deviation, relative bias of the standard error estimator, coverage of interval estimate using Rubin's method and the proposed wild bootstrap method

					Standard error ($\times 10^2$)		Relative Bias (%)		Coverage (%) for 95% CI	
n	m	Model	Point est ($\times 10^2$)	True sd ($\times 10^2$)	Rubin[a]	WB	Rubin[a]	WB	Rubin[a]	WB
500	10	δ =1	84.9	55.9	60.8	58.3	8.74	4.25	93.8	92.1
		δ =2	50.3	56.5	60.4	58.0	6.91	2.69	96.2	**95.1**
		δ =3	26.1	56.8	59.6	57.9	4.93	2.03	94.2	93.2
		δ =4	8.2	56.8	59.0	58.3	3.79	2.53	90.0	88.8
		δ =5	-5.0	57.0	58.6	58.5	2.82	2.65	84.6	84.4
	20	δ =1	86.7	54.2	60.6	55.3	11.83	2.09	93.4	90.7
		δ =2	52.2	54.5	60.0	55.2	10.15	1.37	97.5	**95.4**
		δ =3	27.9	54.8	59.3	55.4	8.17	1.07	95.1	93.2
		δ =4	10.1	54.9	58.9	55.5	7.15	1.03	89.6	86.5
		δ =5	-2.9	55.0	58.5	55.6	6.28	1.16	85.0	82.3
	50	δ =1	85.4	54.4	60.4	53.3	11.02	-1.96	94.3	89.4
		δ =2	51.0	55.0	60.0	53.4	9.08	-2.79	97.2	**94.9**
		δ =3	26.8	55.2	59.3	53.5	7.40	-3.11	94.3	91.7
		δ =4	9.2	55.3	58.7	53.6	6.16	-3.04	90.9	86.6
		δ =5	-4.2	55.3	58.3	53.8	5.55	-2.59	86.2	80.5
	N/A	Tian et.al. 2014	92.8	55.4	-	55.5	-	0.29	-	88.5
1000	10	δ =1	87.0	38.6	43.0	41.1	11.25	6.46	90.7	88.1
		δ =2	52.4	38.8	42.7	41.0	10.10	5.73	97.2	**96.6**
		δ =3	28.1	38.7	42.0	41.2	8.47	6.48	93.3	92.5
		δ =4	10.6	38.7	41.7	41.3	7.59	6.56	84.9	84.2
		δ =5	-3.0	38.8	41.4	41.3	6.80	6.58	76.4	76.0
	20	δ =1	85.4	39.5	42.8	39.1	8.46	-0.93	90.2	86.5
		δ =2	50.8	39.8	42.6	38.9	7.09	-2.20	96.5	**95.2**
		δ =3	26.2	40.1	42.0	39.1	4.82	-2.34	91.4	88.7
		δ =4	8.5	40.1	41.6	39.3	3.71	-2.03	82.4	78.5
		δ =5	-4.9	40.1	41.3	39.3	2.96	-2.05	72.4	69.4
	50	δ =1	86.8	39.1	42.7	37.8	9.19	-3.22	88.1	83.2
		δ =2	52.3	39.5	42.4	37.6	7.30	-4.67	96.3	**93.9**
		δ =3	28.0	39.7	41.8	37.9	5.38	-4.60	92.7	89.3
		δ =4	10.4	39.8	41.4	37.9	4.22	-4.71	83.6	78.3
		δ =5	-2.9	39.8	41.2	38.1	3.37	-4.44	74.8	69.8
	N/A	Tian et.al. 2014	93.0	39.6	-	39.3	-	-0.87	-	81.3

Table S4: Simulation results for the true estimand $\Delta_{\tau}^{\text{control-adj}} = 0.843$ based on control-based method: point estimate, true standard deviation, relative bias of the standard error estimator, coverage of interval estimate using Rubin's method and the proposed wild bootstrap method

					Standard error ($\times 10^2$)		Relative Bias (%)		Coverage (%) for 95% CI	
n	Model	m	Point est ($\times 10^2$)	True sd ($\times 10^2$)	Rubin^a	WB	Rubin^a	WB	Rubin^a	WB
500	Control-based	10	85.2	55.1	60.6	58.1	9.90	5.36	96.8	**95.7**
		20	84.4	53.8	60.3	55.2	12.03	2.53	97.3	**95.3**
		50	87.1	53.2	60.2	53.5	13.08	0.55	97.0	**94.7**
	Tian et.al. 2014	N/A	93.9	54.4	-	55.5	-	1.95	-	95.2
1000	Control-based	10	86.1	38.9	42.8	41.1	10.16	5.81	96.8	**96.3**
		20	83.5	38.7	42.7	39.1	10.53	1.13	96.3	**95.3**
		50	86.6	38.1	42.5	37.9	11.56	-0.70	96.2	**94.6**
	Tian et.al. 2014	N/A	93.1	39.3	-	39.3	-	-0.10	-	94.2