Online Appendix of "Competing Risks Regression for Clustered Data via the Marginal Additive Subdistribution Hazard Model"

Xinyuan Chen[1] Denise Esserman[2] Fan Li[3]

S1 Notations and Assumptions

In this ection, we re-state the notations and assumptions for the right-censored data scenario for facilitate the reading of the subsequent Sections. Recall the following notations defined in the paper:

$$
\begin{aligned}
\hat{\boldsymbol{S}}^{(r)}(t) &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\hat{\omega}_{ij}(t)Y_{ij}(t)\boldsymbol{X}_{ij}(t)^{\otimes r},\\
\tilde{\boldsymbol{s}}^{(r)}(t) &= \lim_{n\to\infty}\hat{\boldsymbol{S}}^{(r)}(t), \;\; r=0,1,\\
\hat{\boldsymbol{X}}(t) &= \hat{\boldsymbol{S}}^{(1)}(t)/\hat{S}^{(0)}(t), \;\; \tilde{\boldsymbol{x}}(t) = \tilde{\boldsymbol{s}}^{(1)}(t)/\tilde{s}^{(0)}(t),\\
\hat{\boldsymbol{A}}(\tau) &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\hat{\omega}_{ij}(t)Y_{ij}(t)\left\{\boldsymbol{X}_{ij}(t)-\hat{\boldsymbol{X}}(t)\right\}^{\otimes 2}\mathrm{d}t.
\end{aligned}
$$

We assume the following regularity conditions to be held throughout the Online Appendix. Without loss of generality, we define $\boldsymbol{\beta}_{k,0}$ as the true value of $\boldsymbol{\beta}_k$ and $\boldsymbol{\beta}_{k,0}\in\mathcal{B}$, where $\mathcal{B}$ is a compact set of $\mathbb{R}^p$ with non-empty interior. And denote $\Lambda_{0k,0}(t)$ as the true baseline subdistribution hazard function.

(1) Conditional on each cluster, $\{T_{ij},\Delta_{ij},\epsilon_{ij},Y_{ij},\boldsymbol{X}_{ij}(t)\}$ are i.i.d. replicates of $\{T,\Delta,\epsilon,Y,\boldsymbol{X}(t)\}$ for $j=1,\ldots,m_i$.

(2) Within each cluster, T and C are independent conditional on $\boldsymbol{X}(\cdot)$.

[1]Xinyuan Chen is Assistant Professor, Department of Mathematics and Statistics, Mississippi State University, Mississippi State, MS, 39762 (email: xchen@math.msstate.edu).

[2]Denise Esserman is Professor, Department of Biostatistics, Yale University School of Public Health, New Haven, CT, 06511 (email: denise.esserman@yale.edu).

[3]Fan Li is Assistant Professor, Department of Biostatistics, Yale University School of Public Health, New Haven, CT, 06511 (email: fan.f.li@yale.edu).

(3) There exists a maximum follow-up time $\tau < \infty$ such that $\mathbb{P}(T \leq \tau) > 0$.

(4) The baseline subdistribution hazard function satisfies $\int_0^\tau \mathrm{d}\Lambda_{0k,0}(t) < \infty$.

(5) $\boldsymbol{X}(t)$ has bounded support on $\mathbb{R}^p$ almost surely and the matrix

$$\boldsymbol{A}(\tau) = \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^\tau \hat{\omega}_{ij}(t) Y_{ij}(t) \left\{ \boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t) \right\}^{\otimes 2} \mathrm{d}t$$

is positive definite.

We define $G(t) = \mathbb{P}(C \geq t)$ as the marginal survival function of the censoring variable C. When the set of censoring variables are correleated within each cluster, the consistency and asymptotic normality of the Kaplan-Meier estimator of $G(t)$, $\hat{G}(t)$, can be established following details in Appendix B.1 of Zhou et al. (2012). We will use this conclusion in deriving the asymptotic properties of the GEE estimator for the regression coefficients.

We maintain the notation in the main paper such that $N_{ij}^c(t) = \mathbb{I}(T_{ij} \leq t, \Delta_{ij} = 0)$ is the counting process of censoring and $Y_{ij}^c(t) = 1 - N_{ij}^c(t-)$ is the at-risk indicator process. It follows that

$$M_{ij}^c(t) = N_{ij}^c(t) - \int_0^t Y_{ij}^c(u) \mathrm{d}\Lambda_0^c(u)$$

is a martingale for the marginal data filtration $\mathcal{F}_{ij}^c(t) = \{N_{ij}^c(u), Y_{ij}^c(u), \boldsymbol{X}_{ij}(u); u \leq t\}$ generated from each participant j of cluster i, and and $\Lambda_0^c(t)$ is the common cumulative hazard of the censoring survival function.

S2 Proof of Theorem 1

We define the martingale evaluated at the true parameters as

$$M_{ij}(\boldsymbol{\beta}_{k,0}, t) = N_{ij}(t) - \int_0^t Y_{ij}(u) \left\{ \mathrm{d}\Lambda_{0k,0}(u) + \boldsymbol{X}_{ij}(u)^T \boldsymbol{\beta}_{k,0} \mathrm{d}u \right\}. \qquad \text{(S2.1)}$$

Then the IPCW score function for $\boldsymbol{\beta}_k$ evaluated at $\boldsymbol{\beta}_{k,0}$ is

$$\boldsymbol{U}(\boldsymbol{\beta}_{k,0}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\hat{\omega}_{ij}(t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t),$$

which can be rewritten into

$$\begin{aligned}\boldsymbol{U}(\boldsymbol{\beta}_{k,0}) =& \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\omega_{ij}(t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t)\\ &+ \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\left\{\hat{\omega}_{ij}(t) - \omega_{ij}(t)\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t)\\ &+ o_p(n^{1/2}),\end{aligned} \tag{S2.2}$$

where $\omega_{ij}(t) = r_{ij}(t)G(t)/G(Z_{ij} \wedge t)$ is the true weight at time t. We can show that the first term on the right-hand side of (S2.2) is equal to

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\right\}\omega_{ij}(t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) + o_p(n^{1/2})$$

because

$$n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\tilde{\boldsymbol{x}}(t) - \hat{\boldsymbol{X}}(t)\right\}\omega_{ij}(t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) \xrightarrow{p} 0$$

by invoking Lemma A.1 of Spiekerman and Lin (1998). For the second term, we follow the argument given in Fine and Gray (1999) and Zhou et al. (2012) by rewriting the $\hat{\omega}_{ij}(t) - \omega_{ij}(t)$ in terms of the Kaplan-Meier estimator of the censoring time survival function as

$$\begin{aligned}&\frac{\hat{G}(t)}{\hat{G}(Z_{ij} \wedge t)} - \frac{G(t)}{G(Z_{ij} \wedge t)}\\ &= \mathbb{I}(Z_{ij} < t)\left\{\frac{\hat{G}(t)}{\hat{G}(Z_{ij})} - \frac{G(t)}{G(Z_{ij})}\right\}\\ &= \mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left\{\frac{\hat{G}(t)}{\hat{G}(Z_{ij})}\frac{G(Z_{ij})}{G(t)} - 1\right\}\\ &= \mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left[\exp\left\{-\hat{\Lambda}_0^c(t) + \hat{\Lambda}_0^c(Z_{ij}) + \Lambda_0^c(t) - \Lambda_0^c(Z_{ij}) - 1\right\}\right]\end{aligned}$$

$$= \mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left\{-\hat{\Lambda}_0^c(t) + \hat{\Lambda}_0^c(Z_{ij}) + \Lambda_0^c(t) - \Lambda_0^c(Z_{ij})\right\} + o_p(1).$$

Consistency of of $\hat{G}(t)$ implies $\|\hat{\Lambda}_0^c(t) - \Lambda_0^c(t)\| \xrightarrow{p} 0$, which results in $-\hat{\Lambda}_0^c(t) + \hat{\Lambda}_0^c(Z_{ij}) + \Lambda_0^c(t) - \Lambda_0^c(Z_{ij}) \xrightarrow{p} 0$. Define $N_{\cdot\cdot}^c(t) = \sum_{i=1}^n \sum_{j=1}^{m_i} N_{ij}^c(t)$, $Y_{\cdot\cdot}^c(t) = \sum_{i=1}^n \sum_{j=1}^{m_i} Y_{ij}^c(t)$, and $M_{\cdot\cdot}^c(t) = \sum_{i=1}^n \sum_{j=1}^{m_i} M_{ij}^c(t)$. Since

$$\hat{\Lambda}_0^c(t) = \int_0^t \frac{\mathrm{d}N_{\cdot\cdot}^c(u)}{Y_{\cdot\cdot}^c(u)},$$

we can further write

$$\begin{aligned}
&\frac{\hat{G}(t)}{\hat{G}(Z_{ij} \wedge t)} - \frac{G(t)}{G(Z_{ij} \wedge t)} \\
&= -\mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left[\left\{\hat{\Lambda}_0^c(t) - \Lambda_0^c(t)\right\} - \left\{\hat{\Lambda}_0^c(Z_{ij}) - \Lambda_0^c(Z_{ij})\right\}\right] + o_p(1) \\
&= -\mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left\{\int_0^t \frac{\mathrm{d}M_{\cdot\cdot}^c(u)}{Y_{\cdot\cdot}^c(u)} - \int_0^{Z_{ij}} \frac{\mathrm{d}M_{\cdot\cdot}^c(u)}{Y_{\cdot\cdot}^c(u)}\right\} + o_p(1) \\
&= -\mathbb{I}(Z_{ij} < t)\frac{G(t)}{G(Z_{ij})}\left\{\int_{Z_{ij}}^t \frac{\mathrm{d}M_{\cdot\cdot}^c(u)}{n\pi(u)}\right\} + o_p(1),
\end{aligned}$$

where $\pi(u) = \lim_{n\to\infty} n^{-1} Y_{\cdot\cdot}^c(u)$. Based on the above relationship, the second term on the right-hand side of (S2.2) becomes

$$\begin{aligned}
&\sum_{i=1}^n \sum_{j=1}^{m_i} \int_0^\tau \{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\} r_{ij}(t) \left\{\frac{\hat{G}(t)}{\hat{G}(Z_{ij} \wedge t)} - \frac{G(t)}{G(Z_{ij} \wedge t)}\right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) + o_p(n^{1/2}) \\
&= -\sum_{i=1}^n \sum_{j=1}^{m_i} \int_0^\tau \{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\} \mathbb{I}(Z_{ij} < t) r_{ij}(t) \frac{G(t)}{G(Z_{ij})}\left\{\int_{Z_{ij}}^t \frac{\mathrm{d}M_{\cdot\cdot}^c(u)}{n\pi(u)}\right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) + o_p(n^{1/2}) \\
&= -\int_0^\tau \sum_{i=1}^n \sum_{j=1}^{m_i} \int_0^\tau \{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\} \omega_{ij}(t) \mathbb{I}(Z_{ij} < u \le t) \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) \frac{\mathrm{d}M_{\cdot\cdot}^c(u)}{n\pi(u)} + o_p(n^{1/2}) \\
&= \int_0^\tau \frac{\boldsymbol{q}(u)}{\pi(u)} \mathrm{d}M_{\cdot\cdot}^c(u) + o_p(n^{1/2}) \\
&= \sum_{i=1}^n \sum_{j=1}^{m_i} \int_0^\tau \frac{\boldsymbol{q}(u)}{\pi(u)} \mathrm{d}M_{ij}^c(u) + o_p(n^{1/2}),
\end{aligned}$$

where

$$\boldsymbol{q}(u) = -\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n}\sum_{j=1}^{m_i} \int_0^\tau \left\{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\right\} \omega_{ij}(t)\mathbb{I}(Z_{ij} < u \le t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t).$$

Hence

$$n^{-1/2}\boldsymbol{U}(\boldsymbol{\beta}_{k,0}) = n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}(\boldsymbol{\eta}_{ij} + \boldsymbol{\psi}_{ij}) + o_p(1),$$

where

$$\boldsymbol{\eta}_{ij} = \int_0^\tau \left\{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\right\} \omega_{ij}(t)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t),$$

and

$$\boldsymbol{\psi}_{ij} = \int_0^\tau \frac{\boldsymbol{q}(u)}{\pi(u)}\mathrm{d}M_{ij}^c(u).$$

Therefore, $n^{-1/2}\boldsymbol{U}(\boldsymbol{\beta}_{k,0})$ is asymptotically equivalent to a sum of n i.i.d. random variables. Applying the multivariate central limit theorem, $n^{-1/2}\boldsymbol{U}(\boldsymbol{\beta}_{k,0})$ is asymptotically normal with mean zero and covariance matrix $\boldsymbol{\Omega} = \mathbb{E}\left\{(\boldsymbol{\eta}_{i\cdot} + \boldsymbol{\psi}_{i\cdot})^{\otimes 2}\right\}$, where $\boldsymbol{\eta}_{i\cdot} = \sum_{j=1}^{m_i}\boldsymbol{\eta}_{ij}$ and $\boldsymbol{\psi}_{i\cdot} = \sum_{j=1}^{m_i}\boldsymbol{\psi}_{ij}$. Denote the (minus) hessian matrix as

$$\boldsymbol{I}(\boldsymbol{\beta}) = -\frac{\partial}{\partial\boldsymbol{\beta}}\boldsymbol{U}(\boldsymbol{\beta}) = \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \hat{\omega}_{ij}(t)Y_{ij}(t)\left\{\boldsymbol{X}_{ij}(t) - \bar{\boldsymbol{X}}(t)\right\}^{\otimes 2}\mathrm{d}t.$$

Assumption (5) indicates that $n^{-1}\boldsymbol{I}(\boldsymbol{\beta})$ converges in probability uniformly in $\boldsymbol{\beta}$ to a continuous limit with $\lim_{n\to\infty}\boldsymbol{I}(\boldsymbol{\beta}) = \boldsymbol{A}(\tau)$, which is a positive definite matrix. By (S2.2), we can also see that $n^{-1}\boldsymbol{U}(\boldsymbol{\beta}_{k,0}) = o_p(1)$. Therefore, we have $\hat{\boldsymbol{\beta}}_k \xrightarrow{p} \boldsymbol{\beta}_{k,0}$.

For asymptotic normality, we consider a first-order Taylor expansion of $\boldsymbol{U}(\boldsymbol{\beta}_k)$ at $\boldsymbol{\beta}_{k,0}$:

$$\begin{aligned} n^{1/2}\left(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0}\right) &= \boldsymbol{A}^{-1}(\tau)\left\{n^{-1/2}\boldsymbol{U}(\boldsymbol{\beta}_{k,0})\right\} + o_p(1) \\ &= \boldsymbol{A}^{-1}(\tau)\left\{n^{-1/2}\sum_{i=1}^{n}(\boldsymbol{\eta}_{i\cdot} + \boldsymbol{\psi}_{i\cdot})\right\} + o_p(1), \end{aligned}$$

which is a sum of bounded i.i.d. random variables. Utilizing the multivariate central limit theorem, $n^{1/2}(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0})$ is asymptotically normal with mean zero and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}_k} = \boldsymbol{A}^{-1}(\tau)\boldsymbol{\Omega}\boldsymbol{A}^{-1}(\tau)$, which can be consistently estimated by $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}_k} = \hat{\boldsymbol{A}}^{-1}(\tau)\hat{\boldsymbol{\Omega}}\hat{\boldsymbol{A}}^{-1}(\tau)$, where

$$
\begin{aligned}
\hat{\boldsymbol{A}}(\tau) &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\hat{\omega}_{ij}(t)Y_{ij}(t)\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}^{\otimes 2}\mathrm{d}t,\\
\hat{\boldsymbol{\Omega}} &= n^{-1}\sum_{i=1}^{n}\left(\hat{\boldsymbol{\eta}}_{i\cdot} + \hat{\boldsymbol{\psi}}_{i\cdot}\right)^{\otimes 2},\\
\hat{\boldsymbol{\eta}}_{i\cdot} &= \sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\hat{\omega}_{ij}(t)\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, t),\\
\hat{\boldsymbol{\psi}}_{i\cdot} &= \sum_{j=1}^{m_i}\int_0^{\tau}\frac{\hat{\boldsymbol{q}}(t)}{\hat{\pi}(t)}\mathrm{d}\hat{M}^c_{ij}(t),\\
\hat{\boldsymbol{q}}(u) &= -n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{\tau}\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\hat{\omega}_{ij}(t)\mathbb{I}(Z_{ij} < u \le t)\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, t),\\
\hat{\pi}(u) &= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}Y^c_{ij}(u),\\
\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, t) &= \mathrm{d}N_{ij}(t) - Y_{ij}(t)\left\{\mathrm{d}\hat{\Lambda}_{0k}(t) + \boldsymbol{X}_{ij}(t)^T\hat{\boldsymbol{\beta}}_k\mathrm{d}t\right\},\\
\mathrm{d}\hat{M}^c_{ij}(t) &= \mathrm{d}N^c_{ij}(t) - Y^c_{ij}(t)\mathrm{d}\hat{\Lambda}^c_0(t),\\
\hat{\Lambda}^c_0(t) &= \int_0^{\tau}\frac{\sum_{i=1}^{n}\sum_{j=1}^{m_i}\mathrm{d}N^c_{ij}(t)}{\sum_{i=1}^{n}\sum_{j=1}^{m_i}Y^c_{ij}(t)}.
\end{aligned}
$$

For the baseline cumulative subdistribution hazard function, we have

$$
\begin{aligned}
&\hat{\Lambda}_{0k}(t) - \Lambda_{0k,0}(t) \\
&\quad = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^{t}\frac{\hat{\omega}_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, u)}{\hat{S}^{(0)}(u)} - \left(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0}\right)^T\int_0^{t}\hat{\boldsymbol{X}}(u)\mathrm{d}u
\end{aligned}
\tag{S2.3}
$$

Since $\hat{\boldsymbol{\beta}}_k$ is consistent and $\hat{S}^{(0)}$ is assumed to be bounded away from zero, the right-hand side of (S2.3) converges in probability to zero uniformly in $t \in [0, \tau]$.

Notice that the first term on the right-hand side of (S2.3) can be further rearranged as

$$n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{\omega_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\hat{S}^{(0)}(u)} + n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{\{\hat{\omega}_{ij}(u)-\omega_{ij}(u)\}\,\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\hat{S}^{(0)}(u)},$$

and

$$\begin{aligned}
&n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{\{\hat{\omega}_{ij}(u)-\omega_{ij}(u)\}\,\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\hat{S}^{(0)}(u)}\\
&= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{r_{ij}(u)}{\hat{S}^{(0)}(u)}\left\{\frac{\hat{G}(u)}{\hat{G}(Z_{ij}\wedge u)}-\frac{G(u)}{G(Z_{ij}\wedge u)}\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)\\
&= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{r_{ij}(u)}{\hat{S}^{(0)}(u)}\left\{-\frac{G(u)}{G(Z_{ij})}\mathbb{I}(Z_{ij}<u)\int_{Z_{ij}}^{u}\frac{\mathrm{d}M_{\cdot\cdot}^{c}(v)}{n\pi(v)}\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)+o_p(n^{-1/2})\\
&= n^{-1}\int_0^\tau\left\{-n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \mathbb{I}(Z_{ij}<v\le u\le t)\frac{\omega_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\hat{S}^{(0)}(u)}\right\}\frac{\mathrm{d}M_{\cdot\cdot}^{c}(v)}{\pi(v)}+o_p(n^{-1/2})\\
&= n^{-1}\int_0^\tau \frac{p(t,v)}{\pi(v)}\mathrm{d}M_{\cdot\cdot}^{c}(v)+o_p(n^{-1/2})\\
&= n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \frac{p(t,v)}{\pi(v)}\mathrm{d}M_{ij}^{c}(v)+o_p(n^{-1/2}),
\end{aligned}$$

where

$$p(t,v)=\lim_{n\to\infty}-n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \mathbb{I}(Z_{ij}<v\le u\le t)\frac{\omega_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\hat{S}^{(0)}(u)}.$$

Combining results from the consistency of $\hat{\boldsymbol{\beta}}_k$, we have

$$\begin{aligned}
\hat{\Lambda}_{0k}(t)-\Lambda_{0k,0}(t) =& n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \frac{\omega_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\tilde{s}^{(0)}(u)} + n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \frac{p(t,v)}{\pi(v)}\mathrm{d}M_{ij}^{c}(v)\\
&+\left\{n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}(\boldsymbol{\eta}_{ij}+\boldsymbol{\psi}_{ij})\right\}^{T}\boldsymbol{A}^{-1}(\tau)\left\{-\int_0^t \tilde{\boldsymbol{x}}(u)\mathrm{d}u\right\}+o_p(n^{-1/2}).
\end{aligned}$$

It follows that,

$$n^{1/2}\left\{\hat{\Lambda}_{0k}(t)-\Lambda_{0k,0}(t)\right\}=n^{-1/2}\sum_{i=1}^{n}H_{\Lambda,i}(t)+o_p(1),$$

where

$$H_{\Lambda,i}(t) = \sum_{j=1}^{m_i} \int_0^t \frac{\omega_{ij}(u)\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u)}{\tilde{s}^{(0)}(u)} + \sum_{j=1}^{m_i} \int_0^\tau \frac{p(t,v)}{\pi(v)} \mathrm{d}M_{ij}^c(v)$$
$$+ \left\{ \sum_{j=1}^{m_i} (\boldsymbol{\eta}_{ij} + \boldsymbol{\psi}_{ij}) \right\}^T \boldsymbol{A}^{-1}(\tau) \left\{ -\int_0^t \tilde{\boldsymbol{x}}(u)\mathrm{d}u \right\}.$$

By the multivariate central limit theorem, $n^{1/2}\left\{\hat{\Lambda}_{0k}(t) - \Lambda_{0k,0}(t)\right\}$ converges in finite-dimensional distribution to a mean-zero Gaussian process. Using similar arguments to those in Lin et al. (2000), it can be shown that $n^{-1/2}\sum_{i=1}^n H_{\Lambda,i}(t)$ is tight. Therefore, it follows that $n^{1/2}\left\{\hat{\Lambda}_{0k}(t) - \Lambda_{0k,0}(t)\right\}$ converges weakly to a mean-zero Gaussian process with its variance $\boldsymbol{\Sigma}_{\Lambda_k}(t) = \mathbb{E}\left\{H_{\Lambda,1}(t)\right\}^2$.

S3 Lemma 1 and Proof

Lemma 1 *Under regularity conditions (R1), (R2), and (R4),* $n^{-1/2}\boldsymbol{W}(t,\boldsymbol{x})$ *converges weakly to* $n^{-1/2}\tilde{\boldsymbol{W}}(t,\boldsymbol{x}) = \sum_{i=1}^n \boldsymbol{Q}_i(t,\boldsymbol{x})$*, a mean-zero Gaussian process in* $l^\infty[0,\tau]$ *with the covariance function between* $(t,\boldsymbol{x})$ *and* $(t^*,\boldsymbol{x}^*)$ *given by* $\mathbb{E}\left\{\boldsymbol{Q}_i(t,\boldsymbol{x})\boldsymbol{Q}_i(t^*,\boldsymbol{x}^*)^T\right\}$*, where*

$$\boldsymbol{Q}_i(t,\boldsymbol{x}) = \sum_{j=1}^{m_i} \int_0^t \omega_{ij}(u) \left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\} \mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\} - \boldsymbol{g}(u,\boldsymbol{x})\right] \mathrm{d}M_{ij}(\boldsymbol{\beta}_k,u)$$
$$- \boldsymbol{h}(t,\boldsymbol{x})\boldsymbol{A}^{-1}(\tau) \sum_{j=1}^{m_i} \int_0^\tau \left\{\boldsymbol{X}_{ij}(t) - \tilde{\boldsymbol{x}}(t)\right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_k,t).$$

Proof of Lemma 1

The cumulative sum of residuals with IPCW is given by

$$\boldsymbol{W}(t,\boldsymbol{x})$$
$$= \sum_{i=1}^n \sum_{j=1}^{m_i} \int_0^t \hat{\omega}_{ij}(u) \boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\} \mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\} \mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k,u)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u) \tag{S3.1}$$

$$- \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\left\{\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u) - \hat{M}_{ij}(\boldsymbol{\beta}_{k,0},u)\right\} \tag{S3.2}$$

$$- \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\left\{\mathrm{d}\hat{M}_{ij}(\boldsymbol{\beta}_{k,0},u) - \hat{M}_{ij}(\hat{\boldsymbol{\beta}}_{k},u)\right\} \tag{S3.3}$$

For (S3.2), we have

$$\mathrm{d}\hat{M}_{ij}(\boldsymbol{\beta}_{k,0},u) = \mathrm{d}N_{ij}(u) - Y_{ij}(u)\left\{\mathrm{d}\hat{\Lambda}^*_{0k}(u) + \boldsymbol{X}_{ij}(u)^T\boldsymbol{\beta}_{k,0}\mathrm{d}u\right\}, \tag{S3.4}$$

where

$$\mathrm{d}\hat{\Lambda}^*_{0k}(u) = \frac{\sum_{l=1}^{n}\sum_{h=1}^{n_l}\hat{\omega}_{lh}(u)\left\{\mathrm{d}N_{lh}(u) - Y_{lh}(u)\boldsymbol{X}_{lh}(u)^T\boldsymbol{\beta}_{k,0}\mathrm{d}u\right\}}{\sum_{l=1}^{n}\sum_{h=1}^{n_l}\hat{\omega}_{lh}(u)Y_{lh}(u)}. \tag{S3.5}$$

Plug (S3.4) and (S3.5) into (S3.2) and exchange summation, we have

$$(S3.2) = -\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\hat{\boldsymbol{g}}(u,\boldsymbol{x})\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u). \tag{S3.6}$$

For (S3.3), Taylor expansion leads to

$$\begin{aligned}
(S3.3) &= -\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\left\{\nabla_{\boldsymbol{\beta}}\mathrm{d}\hat{M}_{ij}(\boldsymbol{\beta},u)(\boldsymbol{\beta}_{k,0}-\hat{\boldsymbol{\beta}}_k) + o_p(1)\right\} \\
&= -\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\left[Y_{ij}(u)\left\{\boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u)\right\}\mathrm{d}u(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0}) + o_p(1)\right] \\
&= -\hat{\boldsymbol{h}}(t,\boldsymbol{x})n(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0}) + o_p(1) \\
&= -\hat{\boldsymbol{h}}(t,\boldsymbol{x})\hat{\boldsymbol{A}}^{-1}(\tau)\boldsymbol{U}(\boldsymbol{\beta}_{k,0}) + o_p(1) \\
&= -\hat{\boldsymbol{h}}(t,\boldsymbol{x})\hat{\boldsymbol{A}}^{-1}(\tau)\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \hat{\omega}_{ij}(t)\left\{\boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t)\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},t) + o_p(1)
\end{aligned} \tag{S3.7}$$

Using similar arguments as in Li et al. (2015), we have

$$(S3.1) \to \sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \omega_{ij}(u)\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u)\le \boldsymbol{x}\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u), \tag{S3.8}$$

and

$$(S3.6) \to -\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \omega_{ij}(u)\boldsymbol{g}(u,\boldsymbol{x})\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},u), \tag{S3.9}$$

The tightness of (S3.7) follows the uniform convergence of $\hat{\boldsymbol{h}}(t,\boldsymbol{x})$, $\hat{\boldsymbol{A}}(t)$, and $\hat{\omega}_{ij}(t)$, which gives

$$(S3.7) \to -\boldsymbol{h}(t,\boldsymbol{x})\boldsymbol{A}^{-1}(\tau)\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^\tau \omega_{ij}(t)\left\{\boldsymbol{X}_{ij}(t)-\tilde{\boldsymbol{x}}(t)\right\}\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0},t) \tag{S3.10}$$

Combining the above intermediate results, we have $n^{-1/2}\boldsymbol{W}(t,\boldsymbol{x})$ converges weakly to $n^{-1/2}\tilde{\boldsymbol{W}}(t,\boldsymbol{x})$.

S4 Proof of Theorem 2

The proof of Theorem 2 largely follows the arguments provided in Yin (2007) with certain modifications. First, the perturbed version of the stochastic process is defined as $\hat{\boldsymbol{W}}(t,\boldsymbol{x}) = \sum_{i=1}^{n}\hat{\boldsymbol{Q}}_i(t,\boldsymbol{x})\xi_i$, where the ξ_i's are generated independently from $\mathcal{N}(0,1)$. Define $\boldsymbol{W}^*(t,\boldsymbol{x}) = \sum_{i=1}^{n}\boldsymbol{Q}_i(t,\boldsymbol{x})\xi_i$.

Based on the conditional multiplier CLT in van der Vaart and Wellner (1996) [30, Theorem 2.9.6], $n^{-1/2}\boldsymbol{W}^*(t,\boldsymbol{x})$ converges weakly to the same Gaussian process, $n^{-1/2}\tilde{\boldsymbol{W}}(t,\boldsymbol{x})$. Therefore, it suffices to prove that $n^{-1/2}\|\boldsymbol{W}^*(t,\boldsymbol{x})-\hat{\boldsymbol{W}}(t,\boldsymbol{x})\| \to 0$ in probability, where

$$\boldsymbol{f}(t,\boldsymbol{x}) = \max_j \sup_{t\in[0,\tau],\boldsymbol{x}\in[-1,1]^p}\{|f_j(t,\boldsymbol{x})|, j=1,\ldots,p\}$$

for a (multivariate) function $\boldsymbol{f}(t,\boldsymbol{x}) = \{f_1(t,\boldsymbol{x}),\ldots,f_p(t,\boldsymbol{x})\}$, and $f_j : ([0,\tau]\times[-1,1]^p) \to \mathbb{R}$.

Define

$$\hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k,t) = \sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\left\{\boldsymbol{X}_{ij}(u)-\hat{\boldsymbol{X}}(u)\right\}\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k,u),$$

and

$$\boldsymbol{\Phi}_i(\boldsymbol{\beta}_{k,0}, t) = \sum_{j=1}^{m_i} \int_0^t \omega_{ij}(u) \left\{ \boldsymbol{X}_{ij}(u) - \tilde{\boldsymbol{x}}(u) \right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, u).$$

We then have

$$\begin{aligned}
&\left\| n^{-1/2} \sum_{i=1}^{n} \xi_i \left\{ \boldsymbol{\Phi}_i(\boldsymbol{\beta}_{k,0}, t) - \hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k, t) \right\} \right\| \\
\leq &\left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \xi_i \omega_{ij}(Z_{ij}) \left\{ \hat{\boldsymbol{X}}(Z_{ij}) - \tilde{\boldsymbol{x}}(Z_{ij}) \right\} \Delta_{ij} \mathbb{I}(Z_{ij} \leq t) \right\| \\
&+ \left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \omega_{ij}(u) Y_{ij}(u) \left[\left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} \mathrm{d}\hat{\Lambda}_{0k}(u) - \left\{ \boldsymbol{X}_{ij}(u) - \tilde{\boldsymbol{x}}(u) \right\} \mathrm{d}\Lambda_{0k,0}(u) \right] \right\| \\
&+ \left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \omega_{ij}(u) Y_{ij}(u) \left[\left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} \hat{\boldsymbol{\beta}}_k' - \left\{ \boldsymbol{X}_{ij}(u) - \tilde{\boldsymbol{x}}(u) \right\} \boldsymbol{\beta}_{k,0}' \right] \boldsymbol{X}_{ij}(u) \mathrm{d}u \right\| \\
&+ \left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \left\{ \omega_{ij}(u) - \hat{\omega}_{ij}(u) \right\} \left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} \mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, u) \right\|. \qquad \text{(S4.1)}
\end{aligned}$$

The first three terms on the right-hand side of (S4.1) converge to zero in probability following the proof of Theorem 2 in Yin (2007). The last term can be decomposed using the Taylor expansion on $\hat{M}_{ij}^k(\hat{\boldsymbol{\beta}}_k, t)$ following the proof of Lemma 1, where

$$\begin{aligned}
&\left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \left\{ \omega_{ij}(u) - \hat{\omega}_{ij}(u) \right\} \left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} \mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, u) \right\| \\
\leq &\left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \left\{ \omega_{ij}(u) - \hat{\omega}_{ij}(u) \right\} \left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, u) \right\| \\
&+ \left\| n^{-1/2} \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^t \xi_i \hat{\omega}_{ij}(u) \hat{\boldsymbol{l}}(u) \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, u) \right\| \qquad \text{(S4.2)} \\
&+ \left\| n^{-1/2} \hat{\boldsymbol{m}}(t) \hat{\boldsymbol{A}}^{-1}(\tau) \sum_{i=1}^{n} \sum_{j=1}^{m_i} \int_0^\tau \hat{\omega}_{ij}(t) \left\{ \boldsymbol{X}_{ij}(t) - \hat{\boldsymbol{X}}(t) \right\} \mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, t) \right\|
\end{aligned}$$

where

$$\hat{\boldsymbol{l}}(u) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m_i} \left\{ \omega_{ij}(u) - \hat{\omega}_{ij}(u) \right\} \left\{ \boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u) \right\} Y_{ij}(u)}{\sum_{i=1}^{n} \sum_{j=1}^{m_i} \hat{\omega}_{ij}(u) Y_{ij}(u)},$$

and

$$\hat{\boldsymbol{m}}(t) = n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \xi_i \left\{\omega_{ij}(u) - \hat{\omega}_{ij}(u)\right\} Y_{ij}(u) \left\{\boldsymbol{X}_{ij}(u) - \hat{\boldsymbol{X}}(u)\right\}^{\otimes 2} \mathrm{d}u.$$

The terms on the right-hand side of (S4.2) converge to zero because of the uniform convergence of $\hat{\omega}_{ij}(u)$. Thus as $n \to \infty$,

$$\left\| n^{-1/2}\sum_{i=1}^{n} \xi_i \left\{\boldsymbol{\Phi}_i(\boldsymbol{\beta}_{k,0}, t) - \hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k, t)\right\}\right\| \to 0. \tag{S4.3}$$

Similarly, we can define

$$\hat{\boldsymbol{\Psi}}_i(\hat{\boldsymbol{\beta}}_k, t, \boldsymbol{x}) = \sum_{j=1}^{m_i}\int_0^t \hat{\omega}_{ij}(u)\left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\} - \hat{\boldsymbol{g}}(u, \boldsymbol{x})\right]\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k, u),$$

and

$$\boldsymbol{\Psi}_i(\boldsymbol{\beta}_{k,0}, t, \boldsymbol{x}) = \sum_{j=1}^{m_i}\int_0^t \omega_{ij}(u)\left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\} - \boldsymbol{g}(u, \boldsymbol{x})\right]\mathrm{d}M_{ij}(\boldsymbol{\beta}_{k,0}, u).$$

In a similar fashion, we have

$$\begin{aligned}
&\left\| n^{-1/2}\sum_{i=1}^{n}\xi_i\left\{\boldsymbol{\Psi}_i(\boldsymbol{\beta}_{k,0}, t, \boldsymbol{x}) - \hat{\boldsymbol{\Psi}}_i(\hat{\boldsymbol{\beta}}_k, t, \boldsymbol{x})\right\}\right\| \\
\le &\left\| n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\xi_i\omega_{ij}(Z_{ij})\left[\hat{\boldsymbol{g}}\{Z_{ij}, \boldsymbol{X}_{ij}(u)\} - \boldsymbol{g}\{Z_{ij}, \boldsymbol{X}_{ij}(u)\}\right]\Delta_{ij}\mathbb{I}(Z_{ij} \le t)\right\| \\
&+ \left\| n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \xi_i\omega_{ij}(u)Y_{ij}(u)\left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\}\left\{\mathrm{d}\hat{\Lambda}_{0k}(u) - \mathrm{d}\Lambda_{0k,0}(u)\right\}\right.\right. \\
&\quad \left.\left. + \boldsymbol{g}(u, \boldsymbol{x})\mathrm{d}\Lambda_{0k,0}(u) - \hat{\boldsymbol{g}}(u, \boldsymbol{x})\mathrm{d}\hat{\Lambda}_{0k}(u)\right]\right\| \qquad \text{(S4.4)} \\
&+ \left\| n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \xi_i\omega_{ij}(u)Y_{ij}(u)\left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u) \le \boldsymbol{x}\right\}(\hat{\boldsymbol{\beta}}_k - \boldsymbol{\beta}_{k,0})^T\right.\right. \\
&\quad \left.\left. + \hat{\boldsymbol{g}}(u, \boldsymbol{x})\boldsymbol{\beta}_{k,0}^T - \boldsymbol{g}(u, \boldsymbol{x})\hat{\boldsymbol{\beta}}_k^T\right]\boldsymbol{X}_{ij}(u)\mathrm{d}u\right\|
\end{aligned}$$

$$+\left\|n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{m_i}\int_0^t \xi_i\left\{\omega_{ij}(u)-\hat{\omega}_{ij}(u)\right\}\left[\boldsymbol{f}\left\{\boldsymbol{X}_{ij}(u)\right\}\mathbb{I}\left\{\boldsymbol{X}_{ij}(u)\leq \boldsymbol{x}\right\}-\hat{\boldsymbol{g}}(u,\boldsymbol{x})\right]\mathrm{d}\hat{M}_{ij}(\hat{\boldsymbol{\beta}}_k,u)\right\|.$$

Following arguments in the proof of showing $\left\|n^{-1/2}\sum_{i=1}^{n}\xi_i\left\{\boldsymbol{\Phi}_i(\boldsymbol{\beta}_{k,0},t)-\hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k,t)\right\}\right\|\to 0$, all four terms on the right-hand side of (S4.4) converge to zero. Thus we have as $n\to\infty$,

$$\left\|n^{-1/2}\sum_{i=1}^{n}\xi_i\left\{\boldsymbol{\Psi}_i(\boldsymbol{\beta}_{k,0},t,\boldsymbol{x})-\hat{\boldsymbol{\Psi}}_i(\hat{\boldsymbol{\beta}}_k,t,\boldsymbol{x})\right\}\right\|\to 0. \tag{S4.5}$$

Combining the intermediate results from above, we have

$$\begin{aligned}
&\left\|n^{-1/2}\boldsymbol{W}^*(t,\boldsymbol{x})-n^{-1/2}\hat{\boldsymbol{W}}(t,\boldsymbol{x})\right\|\\
\leq&\left\|n^{-1/2}\sum_{i=1}^{n}\xi_i\left\{\boldsymbol{\Psi}_i(\boldsymbol{\beta}_{k,0},t,\boldsymbol{x})-\hat{\boldsymbol{\Psi}}_i(\hat{\boldsymbol{\beta}}_k,t,\boldsymbol{x})\right\}\right\|\\
&+\left\|n^{-1/2}\sum_{i=1}^{n}\xi_i\boldsymbol{h}(t,\boldsymbol{x})\boldsymbol{A}^{-1}(\tau)\left\{\hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k,t)-\boldsymbol{\Phi}_i(\boldsymbol{\beta}_{k,0},t)\right\}\right\|\\
&+\left\|n^{-1/2}\sum_{i=1}^{n}\xi_i\left\{\hat{\boldsymbol{h}}(t,\boldsymbol{x})\hat{\boldsymbol{A}}^{-1}(\tau)-\boldsymbol{h}(t,\boldsymbol{x})\boldsymbol{A}^{-1}(\tau)\right\}\hat{\boldsymbol{\Phi}}_i(\hat{\boldsymbol{\beta}}_k,t)\right\|
\end{aligned} \tag{S4.6}$$

The first term on the right-hand side of (S4.6) converges to zero by (S4.5), and by (S4.3) and the boundedness of $\boldsymbol{h}(t,\boldsymbol{x})$ and $\boldsymbol{A}(t)$, the second term converges to zero. The uniform convergence of $\hat{\boldsymbol{A}}(t)$ to $\boldsymbol{A}(t)$ and $\hat{\boldsymbol{h}}(t,\boldsymbol{x})$ to $\boldsymbol{h}(t,\boldsymbol{x})$ for $t\in[0,\tau]$ and $\boldsymbol{x}\in[-1,1]^p$, together with (S4.3), entail that the third term in (S4.6) converges to zero.

S5 Compatibility/Collapsibility of the conditional and marginal additive subdistribution hazard model in Simulation Study 1 and Additional Simulation Results

The conditional CIF for cause 1 is

$$F_1(t; \boldsymbol{X}_{ij}, \nu_i) = 1 - \left\{1 - (\rho + \nu_i)(1 - e^{-t})\right\} \exp\left\{-\boldsymbol{X}_{ij}^T \boldsymbol{\beta}_1 (1 - e^{-t})\right\}.$$

Let f_ν denote the density function of the frailty ν_i. Then the marginal CIF for cause 1 is

$$\begin{aligned}
F_1(t; \boldsymbol{X}_{ij}) &= \int_{v_i} F_1(t; X_{ij}, \nu_i) f_\nu(\nu_i) \mathrm{d}\nu_i \\
&= 1 - \left\{1 - (\rho + \mu_\nu)(1 - e^{-t})\right\} \exp\left\{-\boldsymbol{X}_{ij}^T \boldsymbol{\beta}_1 (1 - e^{-t})\right\},
\end{aligned}$$

where $\mu_\nu = \mathbb{E}(\nu_i) = \int_{\nu_i} f_\nu(\nu_i) \mathrm{d}\nu_i$. The marginal subdistribution hazard is

$$\begin{aligned}
\lambda_{ij1}(t) &= -\frac{\mathrm{d} \log\left\{1 - F_1(t; \boldsymbol{X}_{ij})\right\}}{\mathrm{d}t} \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} \log\left\{1 - (\rho + \mu_\nu)(1 - e^{-t})\right\} + \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{X}_{ij}^T \boldsymbol{\beta}_1 (1 - e^{-t}) \\
&= \lambda_{01}(t) + \boldsymbol{X}_{ij}^T \boldsymbol{\beta}_1 e^{-t},
\end{aligned}$$

which is again of the additive structure.

Web Table 1: Simulation results when the marginal censoring rate is around 40%. The true value of β_1 is 1 and θ refers to the rate parameter of the frailty distribution in the data generating process. CRC: clustered and right-censored; CCC: clustered and complete-censoring; UCRC: unclustered and right-censored; UCCC: unclustered and complete-censoring.

Number of Clusters = 100							Number of Clusters = 250						
Cluster Size	θ	Approaches	$\mathbb{E}(\hat{\beta}_1)$	$s(\hat{\beta}_1)$	$\mathbb{E}(\hat{s})$	Coverage (%)	Cluster Size	θ	Approaches	$\mathbb{E}(\hat{\beta}_1)$	$s(\hat{\beta}_1)$	$\mathbb{E}(\hat{s})$	Coverage (%)
10	0.7	CRC	0.986	0.247	0.244	94.60	10	0.7	CRC	1.004	0.156	0.151	95.50
		CCC	0.989	0.247	0.244	94.60			CCC	1.004	0.157	0.151	95.30
		UCRC	0.986	0.247	0.213	90.60			UCRC	1.004	0.156	0.137	90.40
		UCCC	0.989	0.247	0.214	90.10			UCCC	1.004	0.157	0.137	90.50
	1.0	CRC	0.999	0.243	0.232	95.80		1.0	CRC	1.008	0.154	0.152	96.40
		CCC	1.001	0.244	0.232	95.70			CCC	1.008	0.154	0.152	96.40
		UCRC	0.999	0.243	0.208	89.70			UCRC	1.008	0.154	0.130	89.90
		UCCC	1.001	0.244	0.209	90.10			UCCC	1.008	0.154	0.130	90.20
20	0.7	CRC	0.994	0.209	0.194	94.10	20	0.7	CRC	1.001	0.133	0.131	94.80
		CCC	0.994	0.209	0.194	94.20			CCC	1.001	0.132	0.131	94.60
		UCRC	0.994	0.209	0.155	87.80			UCRC	1.001	0.133	0.096	88.70
		UCCC	0.994	0.209	0.155	87.20			UCCC	1.001	0.132	0.096	88.80
	1.0	CRC	1.005	0.202	0.198	95.80		1.0	CRC	1.006	0.125	0.123	95.50
		CCC	1.005	0.202	0.199	95.80			CCC	1.006	0.125	0.123	95.60
		UCRC	1.005	0.202	0.146	88.60			UCRC	1.006	0.125	0.096	89.40
		UCCC	1.005	0.202	0.147	88.80			UCCC	1.006	0.125	0.097	88.90

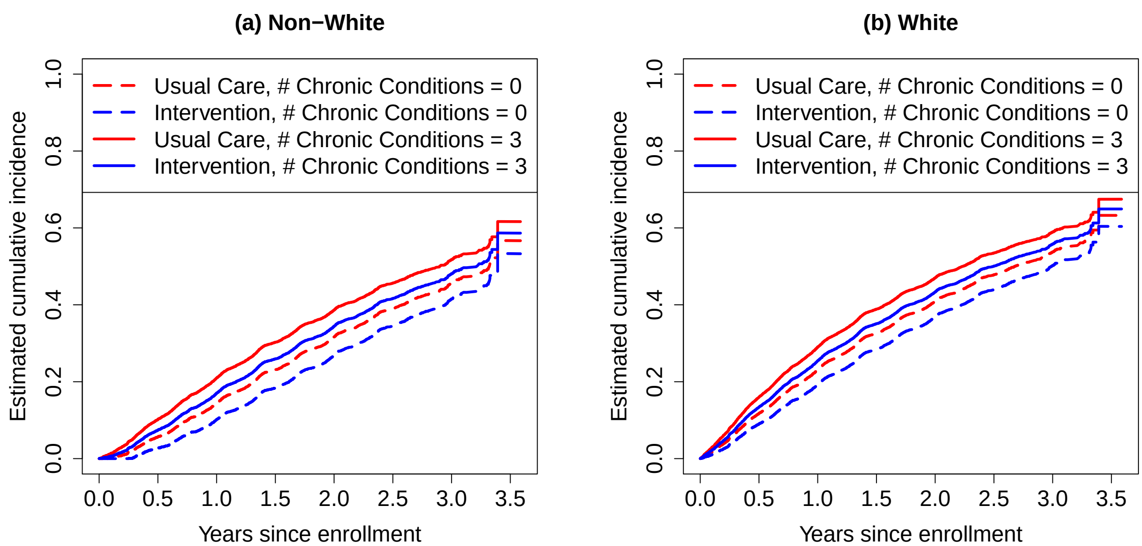
S6 Additional Analysis Results of the STRIDE Study

As an additional illustration, we consider the marginal additive subdistribution hazard model in Section 6 of the main text, but multiply each baseline covariate by e^{-t} to create a set of time-dependent covariate. The estimated coefficients and their standard errors are reported in Web Table 2. Compared to Table 3 in the main paper, the estimated coefficients become larger in magnitude. Further, the estimated cumulative incidence functions for the eight hypothetical patients are presented in Web Figure 1; these predicted cumulative incidence functions are almost identical to those in Figure 1 of the main paper.

Web Table 2: Estimation and model checking results for the analysis of STRIDE study, when all covariates are multipled by e^{-t}.

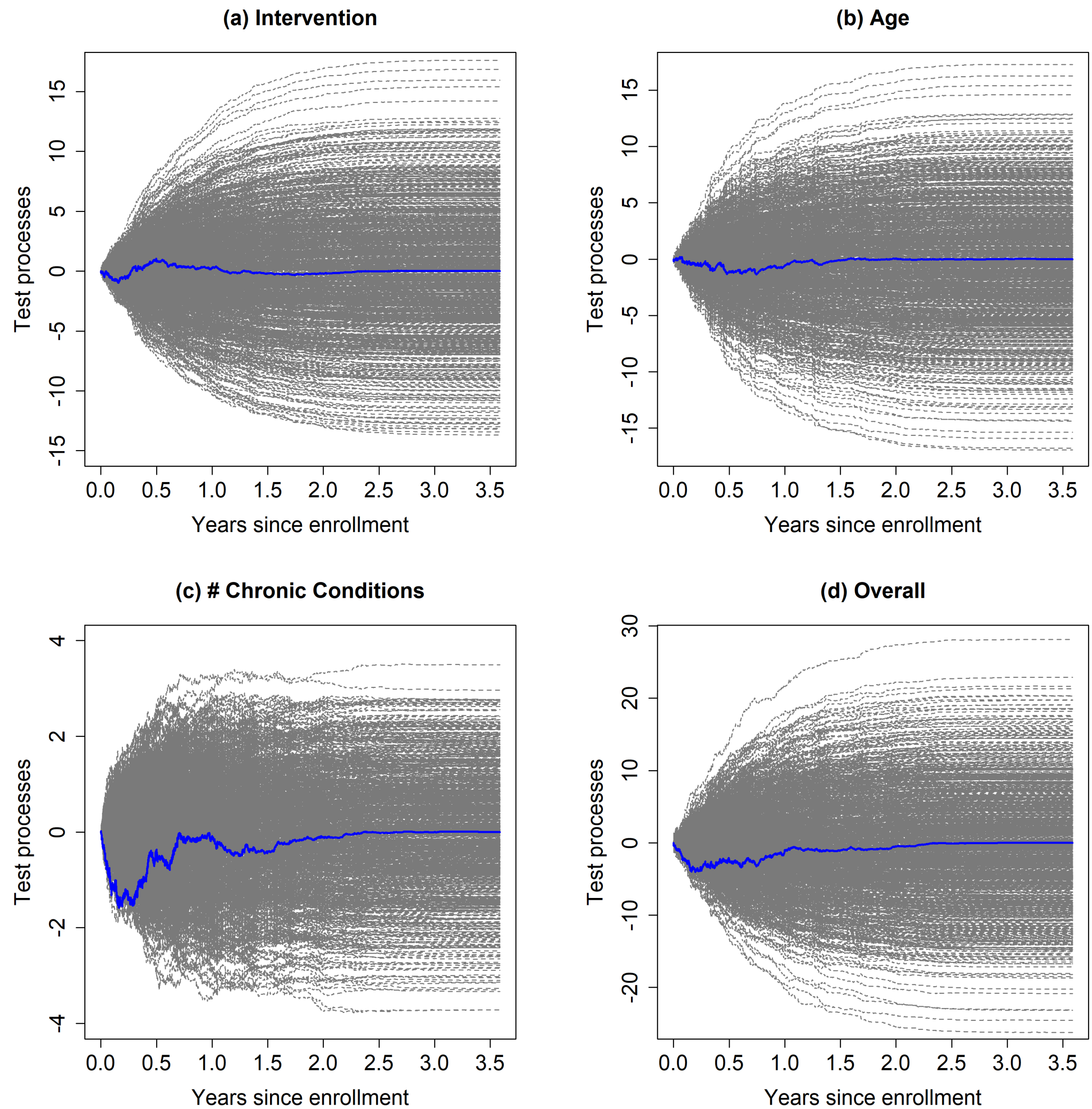
	Model Fitting		Model Checking	
	Estimate	Robust SE	Test Statistic	p-Value
`Intervention` ($\hat{\beta}_1$)	-0.0775	1.6462	1.0354	0.921
`Urban` ($\hat{\beta}_2$)	0.0326	2.6760	1.3580	0.865
`Age` ($\hat{\beta}_3$)	-0.0024	0.1610	0.6435	0.950
`Female` ($\hat{\beta}_4$)	0.0206	1.9728	0.6208	0.920
`White` ($\hat{\beta}_5$)	0.1698	2.6811	1.9655	0.347
`# Chronic Conditions` ($\hat{\beta}_6$)	0.0418	0.7032	1.5813	0.297
Overall	–	–	5.0487	0.984

We further illustrate the proposed goodness-of-fit tests with results for each covariate, as well as the overall model fit, when the data are analyzed by the marginal additive subdistribution hazard model with covariates multiplied by e^{-t}. Each test is based on 1000 simulated test processes with the test statistics and p-values given in Web Table 2. The p-value for each covariate is at least 0.297. The test for overall model fit yields a p-value of 0.984, indicating no evidence from the data against the model assumption, but with a slightly smaller p-value



Web Figure 1: Estimated cumulative incidence functions for self-reported fall injury among four typical White and non-White patients based on the model with the "created" time-dependent covariates. Each patient is assumed to come from an urban practice, with age 76 years old and female.

compared to the analysis in the main paper. For four different tests, Web Figure 2 also graphically illustrates that the observed test process can be completely covered by the 1000 simulated processes.



Web Figure 2: Plot of the observed test process (blue) and simulated curves (gray) under the null in the STRIDE study for model checking based on three covariates (intervention, age and number of chronic coexisting conditions), as well as the overall fit. This illustration is based on the model with the "created" time-dependent covariates

References

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446):496–509.

Li, J., Scheike, T. H., and Zhang, M.-J. (2015). Checking fine and gray subdistribution hazards model with cumulative sums of residuals. *Lifetime Data Analysis*, 21:197–217.

Lin, D. Y., Wei, L. J., Yang, I., and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(4):711–730.

Spiekerman, C. F. and Lin, D. Y. (1998). Marginal regression models for multivariate failure time data. *Journal of the American Statistical Association*, 93(443):1164–1175.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak convergence and empirical processes: with applications to statistics*. Springer, New York.

Yin, G. (2007). Model checking for additive hazards model with multivariate survival data. *Journal of Multivariate Analysis*, 98(5):1018–1032.

Zhou, B., Fine, J., Latouche, A., and Labopin, M. (2012). Competing risks regression for clustered data. *Biostatistics*, 13(3):371–383.