Biostatistics (2017), **0**, 0, *pp.* 1–21
doi:10.1093/biostatistics/"Supplementary materials cure model˙revision"

Semiparametric Model and Inference for Spontaneous Abortion Data with a Cured Proportion and Biased Sampling Supplementary Materials

Jin Piao[1], Jing Ning[*2], Christina D. Chambers[3,4], Ronghui Xu[4,5],

[1] *Department of Biostatistics, The University of Texas School of Public Health, Houston, TX 77030, USA*

[2] *Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA*

[3] *Department of Pediatrics,* [4] *Department of Family Medicine and Public Health,* [5] *Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA*

[*] jning@mdanderson.org

1. Regularity conditions

The regularity conditions are summarized as shown below.

1. The parameters $\{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$ are in a compact set $\mathcal{A}$ that contains $\{\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0\}$. The parameter set $\mathcal{B}$ for the baseline hazard function contains all nondecreasing functions Λ that satisfy $\Lambda(0) = 0$ and $\Lambda(\tau) < \infty$, where τ is the upper bounds for the support of $\tilde{T}$.

2. The cumulative hazard function $\Lambda_0(\cdot)$ is continuously differentiable. The parametric model for the truncation time $\{h(\cdot|\boldsymbol{\theta}) : \boldsymbol{\theta} \in \mathcal{A}\}$ is identifiable.

3. The covariate vector $\boldsymbol{Z}$ is bounded, and $E_0||\boldsymbol{Z}||^2$, $E_0||e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_1}||$ and $E_0||e^{\boldsymbol{\beta}^T \boldsymbol{Z}}||$ are bounded for every $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ in $\mathcal{A}$.

4. The functions $\dot{h}(\cdot|\boldsymbol{\theta})$ and $h(\cdot|\boldsymbol{\theta})$ satisfy a Lipschitz condition in $\boldsymbol{\theta}$, i.e., there exist functions F_1 and F_2 such that
$|\dot{h}(a|\boldsymbol{\theta}_1) - \dot{h}(a|\boldsymbol{\theta}_2)| \leqslant ||\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2||_2 F_1(a) \quad \text{and} \quad |h(a|\boldsymbol{\theta}_1) - h(a|\boldsymbol{\theta}_2)| \leqslant ||\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2||_2 F_2(a)$
where $E_0\left\{F_1(\tilde{A})^2\right\} < \infty$ and $E_0\left\{F_2(\tilde{A})^2\right\} < \infty$.

5. The matrix J_{11} is positive definite, where J_{11} is the Fisher information matrix for $(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0)$.

2. Proof of Theorem 1

The technical details of the consistency proof are similar to those in the literature for maximum likelihood estimation under semiparametric models (Murphy, 1995; Parner *and others*, 1998). We provide only a sketch of the proof. The first step is to show that $\widehat{\boldsymbol{\xi}}_n = \left\{\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\beta}}_n, \widehat{\Lambda}_n\right\}$ is bounded. As $\left\{\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\beta}}_n\right\}$ is found in a compact set, we can find a convergence subsequence of $\left\{\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\beta}}_n\right\}$. Then we only need to show that $\widehat{\Lambda}_n$ is bounded, i.e., $\overline{\lim}_n \widehat{\Lambda}_n < \infty$. We follow the idea of using contradiction construction from Parner *and others* (1998) as follows. Assuming that $\widehat{\Lambda}_n$ diverges, we can construct some sequence $\{\bar{\boldsymbol{\theta}}_n, \bar{\boldsymbol{\alpha}}_n, \bar{\boldsymbol{\beta}}_n, \bar{\Lambda}_n\}$ such that the empirical Kullback-Leibler distance $l_n(\widehat{\boldsymbol{\xi}}_n) - l_n(\bar{\boldsymbol{\xi}}_n)$ would become negative infinity. This contradicts the fact that $\widehat{\boldsymbol{\xi}}_n$ maximizes the log-likelihood function and $l_n(\widehat{\boldsymbol{\xi}}_n) - l_n(\bar{\boldsymbol{\xi}}) \geqslant 0$ for every set of $\bar{\boldsymbol{\xi}}_n$ in the parameter set.

We next apply Helly's selection principle to find a convergent subsequence of $\widehat{\boldsymbol{\xi}}_{n_k}$ for an arbitrary subsequence from the sequence indexed by $\{1, 2, \ldots, n\}$. By the strong law of large numbers, such a convergent subsequence must converge to $\boldsymbol{\xi}_0$ using the classical Kullback-Leibler information approach. For any given subsequence $\{n_k\}$, we can identify a further subsequence of $\widehat{\boldsymbol{\xi}}_{n_k}$ that converges to $\boldsymbol{\xi}_0$. Helly's selection theorem implies that the entire sequence $\widehat{\boldsymbol{\xi}}_n =$

$\left\{\widehat{\boldsymbol{\theta}}_n, \widehat{\boldsymbol{\alpha}}_n, \widehat{\boldsymbol{\beta}}_n, \widehat{\Lambda}_n(t)\right\}$ must converge to $\boldsymbol{\xi}_0 = \{\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_0(t)\}$ for each t. By the assumption that $\Lambda_0(\cdot)$ is continuous, the convergence $\widehat{\Lambda}_n(t)$ is also uniform at each t. The convergence is also almost certain by carrying out the proof for a fixed ω in the underlying probability space Ω, and applying the law of large numbers countable many times.

3. Proof of Theorem 2

The score functions of $\{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$ are calculated by taking the derivative of $l_n(\boldsymbol{\xi})$ with respect to $\{\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$,

$$\begin{aligned} U_{1n}(\boldsymbol{\xi}) = \sum_{i=1}^n \Bigg[& Y_i \int_0^\tau \left\{ \frac{\dot{h}(A_i|\boldsymbol{\theta})}{h(A_i|\boldsymbol{\theta})} - \int_0^\tau Q_1(t|\boldsymbol{Z}_i)\dot{h}(t|\boldsymbol{\theta})dt \right\} dN_i(u) \\ & - \int_0^\tau \left\{ \dot{h}(A_i|\boldsymbol{\theta}) - h(A_i|\boldsymbol{\theta}) \int_0^\tau Q_1(t|\boldsymbol{Z}_i)\dot{h}(t|\boldsymbol{\theta})dt \right\} Q_2(u|\boldsymbol{Z}_i) e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1}} dN_i(u) \\ & + \left\{ \dot{h}(A_i|\boldsymbol{\theta}) - h(A_i|\boldsymbol{\theta}) \int_0^\tau Q_1(t|\boldsymbol{Z}_i)\dot{h}(t|\boldsymbol{\theta})dt \right\} Q_2(X_i|\boldsymbol{Z}_i) e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1}} \Bigg], \end{aligned} \tag{3.1}$$

$$\begin{aligned} U_{2n}(\boldsymbol{\xi}) = \sum_{i=1}^n \Bigg[& Y_i \int_0^\tau \boldsymbol{Z}_{i1} dN_i(u) - \int_0^\tau Q_2(u|\boldsymbol{Z}_i) h(A_i|\boldsymbol{\theta}) \exp(\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1}) \boldsymbol{Z}_{i1} dN_i(u) \\ & + \frac{-\int_0^\tau S(v|\boldsymbol{Z}_i) h_\theta(v|\boldsymbol{Z}_i) dv + S(X_i|\boldsymbol{Z}_i) h(A_i|\boldsymbol{\theta})}{A(X_i|\boldsymbol{Z}_i)} \frac{e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1}} \boldsymbol{Z}_{i1}}{1 + e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1}}} \Bigg], \end{aligned}$$

$$\begin{aligned} U_{3n}(\boldsymbol{\xi}) = \sum_{i=1}^n \Bigg\{ & Y_i \int_0^\tau \boldsymbol{Z}_i dN_i(u) - Y_i \int_0^\tau \int_v^\tau e^{\boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i dN_i(u) d\Lambda(v) \\ & + Y_i \int_0^\tau \int_v^\tau \int_0^\tau Q_1(w|\boldsymbol{Z}_i) h(w|\boldsymbol{\theta}) e^{\boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i dN_i(u) dw d\Lambda(v) \\ & + \int_0^\tau \int_v^\tau Q_2(u|\boldsymbol{Z}_i) h(A|\boldsymbol{\theta}) e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i dN_i(u) d\Lambda(v) \\ & - \int_0^\tau \int_v^\tau \int_0^\tau Q_2(u|\boldsymbol{Z}_i) Q_1(w|\boldsymbol{Z}_i) h(w|\boldsymbol{\theta}) dN_i(u) dw e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i h(A_i|\boldsymbol{\theta}) d\Lambda(v) \\ & - \int_0^\tau Q_2(X_i|\boldsymbol{Z}_i) e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i M_i(v) h(A_i|\boldsymbol{\theta}) d\Lambda(v) \\ & + \int_0^\tau \int_v^\tau Q_2(X_i|\boldsymbol{Z}_i) Q_1(u|\boldsymbol{Z}_i) h(u|\boldsymbol{\theta}) e^{\boldsymbol{\alpha}^T \boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T \boldsymbol{Z}_i} \boldsymbol{Z}_i h(A_i|\boldsymbol{\theta}) du d\Lambda(v) \Bigg\}, \end{aligned} \tag{3.2}$$

where $Q_1(t|\boldsymbol{Z}) = \frac{\exp\left\{-\int_0^t e^{\boldsymbol{\beta}^T\boldsymbol{Z}} d\Lambda(v)\right\}}{\int_0^\tau S(v|\boldsymbol{Z})h(v|\boldsymbol{\theta})dv}$, $A(t|\boldsymbol{Z}) = \int_0^\tau S(v|\boldsymbol{Z})h(v|\boldsymbol{\theta})dv + \exp(\boldsymbol{\alpha}^T\boldsymbol{Z})\exp\left\{-\int_0^t e^{\boldsymbol{\beta}^T\boldsymbol{Z}} d\Lambda(v)\right\} h(A|\boldsymbol{\theta})$ and $Q_2(t|\boldsymbol{Z}) = \exp\left\{-\int_0^t e^{\boldsymbol{\beta}^T\boldsymbol{Z}} d\Lambda(v)\right\}/A(t|\boldsymbol{Z})$. For the infinite dimensional parameter $\Lambda(\cdot)$, we construct a submodel $d\Lambda_\eta(\cdot) = \{1+\eta\phi(\cdot)\}\,d\Lambda_\eta(\cdot)$, where $\phi(\cdot)$ is a bounded and integrable function and η is a positive constant. By taking the derivative of $l_n(\boldsymbol{\xi})$ with respect to η and evaluating at $\eta = 0$, the score operator for $\Lambda(\cdot)$, $U_{4n}(t,\boldsymbol{\xi})$, has the following form

$$
\begin{aligned}
&\sum_{i=1}^n \Bigg\{ Y_i \int_0^t dN_i(u) - Y_i \int_0^t \int_v^\tau e^{\boldsymbol{\beta}^T\boldsymbol{Z}_i} dN_i(u) d\Lambda(v) + Y_i \int_0^t \int_v^\tau \int_0^\tau Q_1(w|\boldsymbol{Z}_i) h(w|\boldsymbol{\theta}) e^{\boldsymbol{\beta}^T\boldsymbol{Z}_i} dN_i(u) dw d\Lambda(v) \\
&+ \int_0^t \int_v^\tau Q_2(u|\boldsymbol{Z}_i) h(A|\boldsymbol{\theta}) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T\boldsymbol{Z}_i} dN_i(u) d\Lambda(v) - \int_0^t Q_2(X_i|\boldsymbol{Z}_i) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T\boldsymbol{Z}_i} M_i(v) h(A_i|\boldsymbol{\theta}) d\Lambda(v) \\
&- \int_0^t \int_v^\tau \int_0^\tau Q_2(u|\boldsymbol{Z}_i) Q_1(w|\boldsymbol{Z}_i) h(w|\boldsymbol{\theta}) dN_i(u) dw e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T\boldsymbol{Z}_i} h(A_i|\boldsymbol{\theta}) d\Lambda(v) \\
&+ \int_0^t \int_v^\tau Q_2(X_i|\boldsymbol{Z}_i) Q_1(u|\boldsymbol{Z}_i) h(u|\boldsymbol{\theta}) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_{i1} + \boldsymbol{\beta}^T\boldsymbol{Z}_i} h(A_i|\boldsymbol{\theta}) du d\Lambda(v) \Bigg\}.
\end{aligned}
$$

To apply the Z-theorem to the infinite-dimensional estimating equations for the asymptotic normality, we need to verify three main conditions: the Fréchet derivative, weak convergence of $\sqrt{n}U_n(\boldsymbol{\xi}_0)$ and stochastic approximation of the estimating equations. We first confirm that the estimating equations, $U_0(\boldsymbol{\xi}) = \{U_{10}(\boldsymbol{\xi}), U_{20}(\boldsymbol{\xi}), U_{30}(\boldsymbol{\xi}), U_{40}(\cdot,\boldsymbol{\xi})\}$, are Fréchet differentiable and the Fréchet derivative is continuously invertible, where

$$
\begin{aligned}
U_{10}(\boldsymbol{\xi}) = {} & E_0\left[Y \int_0^\tau \left\{ \frac{\dot{h}(A|\boldsymbol{\theta})}{h(A|\boldsymbol{\theta})} - \int_0^\tau Q_1(t|\boldsymbol{Z}) \dot{h}(t|\boldsymbol{\theta}) dt \right\} dN(u) \right] \\
& - \int_0^\tau E_0\left[\left\{ \dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta}) \int_0^\tau Q_1(t|\boldsymbol{Z}) \dot{h}(t|\boldsymbol{\theta}) dt \right\} Q_2(u|\boldsymbol{Z}) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1} dN(u) \right] \\
& + E_0\left[\left\{ \dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta}) \int_0^\tau Q_1(t|\boldsymbol{Z}) \dot{h}(t|\boldsymbol{\theta}) dt \right\} Q_2(X|\boldsymbol{Z}) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1} \right],
\end{aligned}
$$

$$
\begin{aligned}
U_{20}(\boldsymbol{\xi}) = {} & E_0\left[\int_0^\tau Y \boldsymbol{Z}_1 dN(u) \right] - \int_0^\tau E_0\left[Q_2(u|\boldsymbol{Z}) h(A|\boldsymbol{\theta}) e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1} \boldsymbol{Z}_1 dN(u) \right] \\
& + E_0\left[\frac{-\int_0^\tau S_2(v|\boldsymbol{Z}) h_\theta(v|\boldsymbol{\theta}) dv + S(X|\boldsymbol{Z}) h(A|\boldsymbol{\theta})}{A(X|\boldsymbol{Z})} \frac{e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1} \boldsymbol{Z}_1}{1 + e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}} \right],
\end{aligned}
$$

$$U_{30}(\boldsymbol{\xi}) = E_0\left[Y\int_0^\tau \boldsymbol{Z}dN(u)\right] - \int_0^\tau E_0\left[\int_v^\tau\int_0^\tau Q_2(u|\boldsymbol{Z})Q_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})dN(u)dwe^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}h(A|\boldsymbol{\theta})d\Lambda(v)\right]$$
$$+\int_0^\tau E_0\left[\int_v^\tau\int_0^\tau YQ_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}dN(u)dwd\Lambda(v)\right] + \int_0^\tau E_0\left[\int_v^\tau Q_2(u|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}dN(u)d\Lambda(v)\right]$$
$$-\int_0^\tau E_0\left[\int_v^\tau Ye^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}dN(u)d\Lambda(v)\right] - \int_0^\tau E_0\left[Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}M(v)h(A|\boldsymbol{\theta})d\Lambda(v)\right]$$
$$+\int_0^\tau E_0\left[\int_v^\tau Q_2(X|\boldsymbol{Z})Q_1(u|\boldsymbol{Z})h(u|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}h(A|\boldsymbol{\theta})dud\Lambda(v)\right],$$

$$U_{40}(t,\boldsymbol{\xi}) = E_0\left[Y\int_0^t dN(u)\right] - \int_0^t E_0\left[\int_v^\tau Ye^{\boldsymbol{\beta}^T\boldsymbol{Z}}dN(u)d\Lambda(v)\right] - \int_0^t E_0\left[Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}M(v)h(A|\boldsymbol{\theta})d\Lambda(v)\right]$$
$$+\int_0^t E_0\left[\int_v^\tau\int_0^\tau YQ_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\beta}^T\boldsymbol{Z}}dN(u)dwd\Lambda(v)\right] + \int_0^t E_0\left[\int_v^\tau Q_2(u|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}dN(u)d\Lambda(v)\right]$$
$$-\int_0^t E_0\left[\int_v^\tau\int_0^\tau Q_2(u|\boldsymbol{Z})Q_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})dN(u)dwe^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}h(A|\boldsymbol{\theta})d\Lambda(v)\right]$$
$$+\int_0^t E_0\left[\int_v^\tau Q_2(X|\boldsymbol{Z})Q_1(u|\boldsymbol{Z})h(u|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}h(A|\boldsymbol{\theta})dud\Lambda(v)\right].$$

The Fréchet differentiability of $U_0(\boldsymbol{\xi})$ at $\boldsymbol{\xi} = \boldsymbol{\xi}_0$ can be verified by the definition, and the derivation can be calculated by using the Gâteaux variation of $U_0(\boldsymbol{\xi})$. Specifically, we can take the differentiation of $U_0(\boldsymbol{\xi}_\eta)$ with respect to η, where $\boldsymbol{\xi}_\eta = (\boldsymbol{\theta}_\eta, \boldsymbol{\alpha}_\eta, \boldsymbol{\beta}_\eta, \Lambda_\eta) = (\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_0) + \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \Lambda)$, and evaluated at $\eta = 0$. The Gâteaux derivative of $U_{10}(\boldsymbol{\xi})$ evaluated at $\boldsymbol{\xi}_0$ is

$$-\left\{s_{11}(\boldsymbol{\theta}) + s_{12}(\boldsymbol{\alpha}) + s_{13}(\boldsymbol{\beta}) + s_{14}(\Lambda)\right\},$$

where

$$s_{11}(\boldsymbol{\theta}) = \frac{\partial}{\partial\eta}U_{10}(\boldsymbol{\theta}_\eta, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\theta}\kappa_{11},$$
$$s_{12}(\boldsymbol{\alpha}) = \frac{\partial}{\partial\eta}U_{10}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_\eta, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\alpha}^T\kappa_{12},$$
$$s_{13}(\boldsymbol{\beta}) = \frac{\partial}{\partial\eta}U_{10}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_\eta, \Lambda_0)|_{\eta=0} = \boldsymbol{\beta}^T\int_0^\tau \kappa_{13}^{(1)}(v)d\Lambda_0(v),$$
$$s_{14}(\Lambda) = \frac{\partial}{\partial\eta}U_{10}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_\eta)|_{\eta=0} = \int_0^\tau\int_v^\tau \kappa_{13}^{(0)}(u)dud\Lambda(v),$$

$$
\begin{aligned}
\kappa_{11} = &-E_0\Bigg\{Y\int_0^\tau \Bigg[\frac{\ddot{h}(A|\boldsymbol{\theta})}{h(A|\boldsymbol{\theta})} - \Big\{\frac{\dot{h}(A|\boldsymbol{\theta})}{h(A|\boldsymbol{\theta})}\Big\}^2 - \int_0^\tau Q_1(t|\boldsymbol{Z})\ddot{h}(t|\boldsymbol{\theta})dt + \Big\{\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}^2\Bigg]dN(u)\Bigg\} \\
&+ \int_0^\tau E_0\Bigg\{\Bigg[\ddot{h}(A|\boldsymbol{\theta}) - \dot{h}(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\ddot{h}(t|\boldsymbol{\theta})dt \\
&\qquad + h(A|\boldsymbol{\theta})\Big\{\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}^2\Bigg]Q_2(u|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}dN(u)\Bigg\} \\
&+ \int_0^\tau E_0\Bigg[\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_3(u|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}dN(u)\Bigg] \\
&- E_0\Bigg\{\Bigg[\ddot{h}(A|\boldsymbol{\theta}) - \dot{h}(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\ddot{h}(t|\boldsymbol{\theta})dt \\
&\qquad + h(A|\boldsymbol{\theta})\Big\{\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}^2\Bigg]Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\Bigg\} \\
&- E_0\Bigg[\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_3(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{12} = &\int_0^\tau E_0\Bigg[\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(v|\boldsymbol{Z})\dot{h}(v|\boldsymbol{\theta})dv\Big\}\Big\{Q_2(u|\boldsymbol{Z})\boldsymbol{Z}_1 + Q_4(u|\boldsymbol{Z})\Big\}\exp(\boldsymbol{\alpha}^T\boldsymbol{Z}_1)dN(u)\Bigg] \\
&- E_0\Bigg[\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(v|\boldsymbol{Z})\dot{h}(v|\boldsymbol{\theta})dv\Big\}\Big\{Q_2(X|\boldsymbol{Z})\boldsymbol{Z} + Q_4(X|\boldsymbol{Z})\Big\}e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{13}^{(l)}(v) = &-E_0\Bigg[Y\int_0^\tau e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Big\{\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt - \int_0^\tau\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})Q_1(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1dt\Big\}dN(u)\Bigg] \\
&- E_0\Bigg[\int_v^\tau\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_2(u|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Big\{1 - Q_2(u)e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}h(A|\boldsymbol{\theta}))\Big\}dN(u)\Bigg] \\
&+ E_0\Bigg[\int_0^\tau\int_v^\tau\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_2(u|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Big\{\frac{S(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})}{A(u|\boldsymbol{Z})}\Big\}du_1dN(u)\Bigg] \\
&+ E_0\Bigg[\int_0^\tau h(A|\boldsymbol{\theta})Q_2(u|\boldsymbol{Z})e^{\boldsymbol{\beta}^T\boldsymbol{Z}+\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}^l\Big\{\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt - \int_0^\tau\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})Q_1(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1dt\Big\}dN(u)\Bigg] \\
&- E_0\Bigg[\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Big\{1 - Q_2(X)e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}h(A|\boldsymbol{\theta}))\Big\}\Bigg] \\
&+ E_0\Bigg[\int_v^\tau\Big\{\dot{h}(A|\boldsymbol{\theta}) - h(A|\boldsymbol{\theta})\int_0^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt\Big\}Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Big\{\frac{S(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})}{A(X|\boldsymbol{Z})}\Big\}du_1\Bigg] \\
&+ E_0\Bigg[h(A|\boldsymbol{\theta})Q_2(X|\boldsymbol{Z})e^{\boldsymbol{\beta}^T\boldsymbol{Z}+\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}^l\Big\{\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})dt - \int_0^\tau\int_v^\tau Q_1(t|\boldsymbol{Z})\dot{h}(t|\boldsymbol{\theta})Q_1(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1dt\Big\}\Bigg],
\end{aligned}
$$

$Q_3(t|\boldsymbol{Z}) = -Q_2(t|\boldsymbol{Z})\Big\{\frac{\int_0^\tau S(u_1|\boldsymbol{Z})\dot{h}(u_1|\boldsymbol{\theta})du_1}{A(t|\boldsymbol{Z})} + e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}Q_2(t|\boldsymbol{Z})\dot{h}(A|\boldsymbol{\theta})\Big\}$ and $Q_4(t|\boldsymbol{Z}) = -Q_2(t|\boldsymbol{Z})^2 h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}_1$.

The Gâteaux derivative of $U_{20}(\boldsymbol{\xi})$ evaluated at $\boldsymbol{\xi}_0$ is

$$-\left\{s_{21}(\boldsymbol{\theta}) + s_{22}(\boldsymbol{\alpha}) + s_{23}(\boldsymbol{\beta}) + s_{24}(\Lambda)\right\},$$

where

$$\begin{aligned}
s_{21}(\boldsymbol{\theta}) &= \frac{\partial}{\partial\eta}U_{20}(\boldsymbol{\theta}_\eta, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\theta}\kappa_{12}, \\
s_{22}(\boldsymbol{\alpha}) &= \frac{\partial}{\partial\eta}U_{20}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_\eta, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\alpha}^T\kappa_{22}, \\
s_{23}(\boldsymbol{\beta}) &= \frac{\partial}{\partial\eta}U_{20}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_\eta, \Lambda_0)|_{\eta=0} = \boldsymbol{\beta}^T\int_0^\tau \kappa_{23}^{(2)}(v)d\Lambda_0(v), \\
s_{24}(\Lambda) &= \frac{\partial}{\partial\eta}U_{20}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_\eta)|_{\eta=0} = \int_0^\tau\int_v^\tau \kappa_{23}^{(1)}(u)dud\Lambda(v),
\end{aligned}$$

$$\begin{aligned}
\kappa_{22} = &\int_0^\tau E_0\left[Q_2(u|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}_1^2 dN(u)\right] + \int_0^\tau E_0\left[Q_4(u|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}_1 dN(u)\right] \\
&- E_0\left[\frac{-\int_0^\tau S(v|\boldsymbol{Z})h_\theta(v)dv + S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})}{A(X|\boldsymbol{Z})}\frac{e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\boldsymbol{Z}_1^2}{\left\{1+e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}\right\}^2}\right] \\
&+ E_0\left[\frac{\left\{-\int_0^\tau S(v|\boldsymbol{Z})h_\theta(v)dv + S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})\right\}\left\{S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})\right\}}{A(X|\boldsymbol{Z})^2}\frac{\exp(2\boldsymbol{\alpha}^T\boldsymbol{Z}_1)\boldsymbol{Z}_1^2}{1+e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}}\right],
\end{aligned}$$

$$\begin{aligned}
\kappa_{23}^{(l)}(v) = &E_0\Bigg\{\int_v^\tau\Big[-Q_2(u|\boldsymbol{Z})\Big\{1 - Q_2(u)e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}h(A)\Big\}\Big]h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}_1\boldsymbol{Z}^l dN(u)\Bigg\} \\
&- E_0\left[\left\{\int_v^\tau \frac{S(t|\boldsymbol{Z})h(t|\boldsymbol{\theta})}{A(X|\boldsymbol{Z})}dt - \frac{S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})M(v)}{A(X|\boldsymbol{Z})}\right\}\frac{e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}_1\boldsymbol{Z}^l}{1+e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}}\right] \\
&- \left\{\int_v^\tau S(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1 + S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}M(v)\right\}\frac{e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}_1\boldsymbol{Z}^l}{1+e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}} \\
&\qquad\frac{\left\{-\int_0^\tau S(u_1|\boldsymbol{Z})h_\theta(u_1)du_1 + S(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})\right\}}{A(X|\boldsymbol{Z})^2}.
\end{aligned}$$

The Gâteaux derivative of $U_{30}(\boldsymbol{\xi})$ evaluated at $\boldsymbol{\xi}_0$ is

$$-\left\{s_{31}(\boldsymbol{\theta}) + s_{32}(\boldsymbol{\alpha}) + s_{33}(\boldsymbol{\beta}) + s_{34}(\Lambda)\right\},$$

where

$$
\begin{aligned}
s_{31}(\boldsymbol{\theta}) &= \frac{\partial}{\partial\eta}U_{30}(\boldsymbol{\theta}_\eta, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\theta}^T \int_0^\tau \kappa_{13}^{(1)}(v)d\Lambda_0(v),\\
s_{32}(\boldsymbol{\alpha}) &= \frac{\partial}{\partial\eta}U_{30}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_\eta, \boldsymbol{\beta}_0, \Lambda_0)|_{\eta=0} = \boldsymbol{\alpha}^T \int_0^\tau \kappa_{23}^{(1)}(v)d\Lambda_0(v),\\
s_{33}(\boldsymbol{\beta}) &= \frac{\partial}{\partial\eta}U_{30}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_\eta, \Lambda_0)|_{\eta=0} = \boldsymbol{\beta}^T \Big\{ \int_0^\tau \kappa_{331}^{(2)}(u)d\Lambda_0(u) + \int_0^\tau\int_0^\tau \kappa_{332}^{(2)}(u,v)d\Lambda_0(v)d\Lambda_0(u)\Big\},\\
s_{34}(\Lambda) &= \frac{\partial}{\partial\eta}U_{30}(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0, \Lambda_\eta)|_{\eta=0} = \int_0^\tau \kappa_{331}^{(1)}(u)d\Lambda(u) + \int_0^\tau\int_0^\tau \kappa_{332}^{(1)}(u,v)d\Lambda_0(v)d\Lambda(u),
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{331}^{(l)}(v) &= E_0\Bigg[\int_v^\tau Y e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l dN(u)\Bigg] - E_0\Bigg[\int_v^\tau\int_0^\tau YQ_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l dN(u)dw\Bigg]\\
&- E_0\Bigg[\int_v^\tau Q_2(u|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l dN(u)\Bigg] + E_0\Bigg[\int_v^\tau\int_0^\tau Q_2(u|\boldsymbol{Z})Q_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l h(A|\boldsymbol{\theta})dN(u)dw\Bigg]\\
&+ E_0\Bigg[Q_2(X|\boldsymbol{Z})h(A|\boldsymbol{\theta})M(v)e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Bigg] - E_0\Bigg[\int_v^\tau Q_2(X|\boldsymbol{Z})Q_1(w|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l h(A|\boldsymbol{\theta})dw\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
\kappa_{332}^{(l)}(u,v) &= -E_0\Bigg[\int_u^\tau\int_0^\tau YQ_5(w,u,v|\boldsymbol{Z})h(w|\boldsymbol{\theta})e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l dN(u_1)dw\Bigg]\\
&- E_0\Bigg[\int_u^\tau Q_6(u_1,u,v|\boldsymbol{Z})h(A|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l dN(u_1)\Bigg] + E_0\Bigg[Q_6(X,u,v|\boldsymbol{Z})h(A|\boldsymbol{\theta})M(u)e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l\Bigg]\\
&+ E_0\Bigg[\int_u^\tau\int_0^\tau \Big\{Q_6(u_1,u,v|\boldsymbol{Z})Q_1(w|\boldsymbol{Z}) + Q_2(u_1|\boldsymbol{Z})Q_5(w,u,v|\boldsymbol{Z})\Big\}h(w|\boldsymbol{\theta})dN(u_1)dw e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l h(A|\boldsymbol{\theta})\Bigg]\\
&- E_0\Bigg[\int_u^\tau \Big\{Q_6(X,u,v|\boldsymbol{Z})Q_1(w|\boldsymbol{Z}) + Q_2(X|\boldsymbol{Z})Q_5(w,u,v|\boldsymbol{Z})\Big\}h(w|\boldsymbol{\theta})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1+\boldsymbol{\beta}^T\boldsymbol{Z}}\boldsymbol{Z}^l h(A|\boldsymbol{\theta})dw\Bigg],
\end{aligned}
$$

$$
\begin{aligned}
Q_5(z,u,v|\boldsymbol{Z}) &= \Big\{-I(z \geqslant v) + \int_u^\tau Q_1(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1\Big\} Q_1(z|\boldsymbol{Z})e^{\boldsymbol{\beta}^T\boldsymbol{Z}},\\
Q_6(z,u,v|\boldsymbol{Z}) &= -Q_2(z|\boldsymbol{Z})e^{\boldsymbol{\beta}^T\boldsymbol{Z}}\Big\{I(z \geqslant v) - \frac{\int_u^\tau S(u_1|\boldsymbol{Z})h(u_1|\boldsymbol{\theta})du_1}{A(z|\boldsymbol{Z})} - Q_2(z|\boldsymbol{Z})e^{\boldsymbol{\alpha}^T\boldsymbol{Z}_1}h(A|\boldsymbol{\theta})I(z \geqslant v)\Big\}.
\end{aligned}
$$

Similarly, we can derive the Gâteaux derivative of $U_{40}(t, \boldsymbol{\xi})$ evaluated at $\boldsymbol{\xi}_0$, denoted as

$$
-\left\{s_{41}(\boldsymbol{\theta})(t) + s_{42}(\boldsymbol{\alpha})(t) + s_{43}(\boldsymbol{\beta})(t) + s_{44}(\Lambda)(t)\right\},
$$

where

$$s_{41}(\boldsymbol{\theta})(t)=\frac{\partial}{\partial\eta}U_{40}(t,\boldsymbol{\theta}_\eta,\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0,\Lambda_0)|_{\eta=0}=\boldsymbol{\theta}\int_0^t\kappa_{13}^{(0)}(v)d\Lambda_0(v),$$
$$s_{42}(\boldsymbol{\alpha})(t)=\frac{\partial}{\partial\eta}U_{40}(t,\boldsymbol{\theta}_0,\boldsymbol{\alpha}_\eta,\boldsymbol{\beta}_0,\Lambda_0)|_{\eta=0}=\boldsymbol{\alpha}^T\int_0^t\kappa_{23}^{(0)}(v)d\Lambda_0(v),$$
$$s_{43}(\boldsymbol{\beta})(t)=\frac{\partial}{\partial\eta}U_{40}(t,\boldsymbol{\theta}_0,\boldsymbol{\alpha}_0,\boldsymbol{\beta}_\eta,\Lambda_0)|_{\eta=0}=\boldsymbol{\beta}^T\Big\{\int_0^t\kappa_{331}^{(1)}(u)d\Lambda_0(u)+\int_0^t\int_0^\tau\kappa_{332}^{(1)}(u,v)d\Lambda_0(v)d\Lambda_0(u)\Big\}$$
$$s_{44}(\Lambda)(t)=\frac{\partial}{\partial\eta}U_{40}(t,\boldsymbol{\theta}_0,\boldsymbol{\alpha}_0,\boldsymbol{\beta}_0,\Lambda_\eta)|_{\eta=0}=\int_0^t\kappa_{331}^{(0)}(u)d\Lambda(u)+\int_0^t\int_0^\tau\kappa_{332}^{(0)}(u,v)d\Lambda(v)d\Lambda_0(u).$$

The Fréchet derivative of $\dot U_0(\boldsymbol{\xi})$ has the following form

$$\dot U_0(\boldsymbol{\xi})=\begin{pmatrix}s_{11}&s_{12}&s_{13}&s_{14}\\ s_{21}&s_{22}&s_{23}&s_{24}\\ s_{31}&s_{32}&s_{33}&s_{34}\\ s_{41}&s_{42}&s_{43}&s_{44}\end{pmatrix}\begin{pmatrix}\boldsymbol{\theta}\\ \boldsymbol{\alpha}\\ \boldsymbol{\beta}\\ \Lambda\end{pmatrix}=\begin{pmatrix}s_{11}(\boldsymbol{\theta})&s_{12}(\boldsymbol{\alpha})&s_{13}(\boldsymbol{\beta})&s_{14}(\Lambda)\\ s_{21}(\boldsymbol{\theta})&s_{22}(\boldsymbol{\alpha})&s_{23}(\boldsymbol{\beta})&s_{24}(\Lambda)\\ s_{31}(\boldsymbol{\theta})&s_{32}(\boldsymbol{\alpha})&s_{33}(\boldsymbol{\beta})&s_{34}(\Lambda)\\ s_{41}(\boldsymbol{\theta})&s_{42}(\boldsymbol{\alpha})&s_{43}(\boldsymbol{\beta})&s_{44}(\Lambda)\end{pmatrix}.$$

Denote the finite parameters $\boldsymbol{\mu}=(\boldsymbol{\theta},\boldsymbol{\alpha},\boldsymbol{\beta})^T$ and define

$$J_{11}=\begin{pmatrix}\kappa_{11}&\kappa_{12}&\Big\{\int_0^\tau\kappa_{13}^{(1)}(v)d\Lambda_0(v)\Big\}^T\\ \kappa_{12}&\kappa_{22}&\Big\{\int_0^\tau\kappa_{23}^{(1)}d\Lambda_0(v)\Big\}^T\\ \int_0^\tau\kappa_{13}^{(1)}(v)d\Lambda_0(v)&\int_0^\tau\kappa_{23}^{(1)}(v)d\Lambda_0(v)&\int_0^\tau\kappa_{331}^{(2)}(v)d\Lambda_0(v)+\int_0^\tau\int_0^\tau\kappa_{332}^{(2)}(u,v)d\Lambda_0(v)d\Lambda_0(u)\end{pmatrix},$$

$$J_{21}(t)=J_{12}(v)^T=\begin{pmatrix}\int_0^t\kappa_{13}^{(0)}(v)d\Lambda_0(v)&\int_0^t\kappa_{23}^{(0)}(v)d\Lambda_0(v)&\int_0^t\kappa_{331}^{(1)}(u)d\Lambda_0(u)+\int_0^t\int_0^\tau\kappa_{332}^{(1)}(u,v)d\Lambda_0(v)d\Lambda_0(u)\end{pmatrix}.$$

The Fréchet derivative of $\dot U_0(\boldsymbol{\xi})$ has the following form

$$\dot U_0(\boldsymbol{\xi})=\begin{pmatrix}\sigma_{11}&\sigma_{12}\\ \sigma_{21}&\sigma_{22}\end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}\\ \Lambda\end{pmatrix}=\begin{pmatrix}\sigma_{11}(\boldsymbol{\mu})+\sigma_{12}(\Lambda)\\ \sigma_{21}(\boldsymbol{\mu})+\sigma_{22}(\Lambda)\end{pmatrix},$$

where

$$\sigma_{11}(\boldsymbol{\mu})=-\begin{pmatrix}s_{11}(\boldsymbol{\theta})+s_{12}(\boldsymbol{\alpha})+s_{13}(\boldsymbol{\beta})\\ s_{21}(\boldsymbol{\theta})+s_{22}(\boldsymbol{\alpha})+s_{23}(\boldsymbol{\beta})\\ s_{31}(\boldsymbol{\theta})+s_{32}(\boldsymbol{\alpha})+s_{33}(\boldsymbol{\beta})\end{pmatrix}=J_{11}\boldsymbol{\mu},$$
$$\sigma_{21}(\boldsymbol{\mu})(t)=-\begin{pmatrix}s_{41}(\boldsymbol{\theta})(t)+s_{42}(\boldsymbol{\alpha})(t)+s_{43}(\boldsymbol{\beta})(t)\end{pmatrix}=J_{21}(t)\boldsymbol{\mu},$$
$$\sigma_{12}(\Lambda)=\begin{pmatrix}s_{14}(\Lambda)\\ s_{24}(\Lambda)\\ s_{34}(\Lambda)\end{pmatrix}=\int_0^\tau J_{12}(v)d\Lambda(v),\ \text{and}\ \sigma_{22}(\Lambda)=s_{44}(\Lambda).$$

If the inverse of $\dot U_{\boldsymbol{\xi}_0}$ exists, then it must have the following form

$$\dot U_{\boldsymbol{\xi}_0}^{-1}(\boldsymbol{\xi})=\begin{pmatrix}\sigma_{11}^{-1}+\sigma_{11}^{-1}\sigma_{12}\Phi^{-1}\sigma_{21}\sigma_{11}^{-1}&-\sigma_{11}^{-1}\sigma_{12}\Phi^{-1}\\ \Phi^{-1}\sigma_{21}\sigma_{11}^{-1}&\Phi^{-1}\end{pmatrix}\begin{pmatrix}\boldsymbol{\mu}\\ \Lambda\end{pmatrix},$$

where $\Phi = \sigma_{22} - \sigma_{21}\sigma_{11}^{-1}\sigma_{12}$. Thus, to show $\dot{U}_{\boldsymbol{\xi}_0}$ is continuously invertible, we only need to show that σ_{11} and Φ are continuously invertible. Note that σ_{11} is identical to J_{11}, which is the Fisher information for $(\boldsymbol{\theta}_0, \boldsymbol{\alpha}_0, \boldsymbol{\beta}_0)$ when the baseline hazard functions Λ_0 are known. It is reasonable to assume that the information matrix J_{11} is positive definite and invertible with inverse J_{11}^{-1}. Thus, σ_{11} is continuously invertible with inverse $\sigma_{11}^{-1}(\boldsymbol{\mu}) = J_{11}^{-1}\boldsymbol{\mu}$.

To show that Φ is continuously invertible, it is equivalent to showing that there exists a unique solution to the operator equation $\Phi(\Lambda) = \sigma_{22}(\Lambda) - \sigma_{21}\sigma_{11}^{-1}\sigma_{12}(\Lambda) = \check{\Lambda}$ for each bounded function $\check{\Lambda}$.

$$\int_0^t \kappa_{331}^{(0)}(u)d\Lambda(u) + \int_0^t \int_0^\tau \kappa_{332}^{(0)}(u,v)d\Lambda(v)d\Lambda_0(u) - \int_0^\tau J_{21}(t)J_{11}^{-1}J_{12}(u)d\Lambda(u) = \check{\Lambda}(t).$$

Taking the derivative with respect to t on both sides, we have

$$\kappa_{331}^{(0)}(t)d\Lambda(t) + \lambda_0(t)\int_0^\tau \kappa_{332}^{(0)}(t,v)d\Lambda(v) - \int_0^\tau \dot{J}_{21}(t)J_{11}^{-1}J_{12}(u)d\Lambda(u) = d\check{\Lambda}(t).$$

We can rewrite the equation as a Fredholm integral equation of the second type

$$d\Lambda(t) - \int_0^\tau K(t,u)d\Lambda(u) = \frac{d\check{\Lambda}(t)}{\kappa_{331}^{(0)}(t)},$$

where

$$K(t,u) = \frac{\lambda_0(t)}{\kappa_{331}^{(0)}(t)}\left\{\begin{pmatrix} \kappa_{13}^{(0)}(t) \\ \kappa_{23}^{(0)}(t) \\ \kappa_{331}^{(1)}(t) + \int_0^\tau \kappa_{332}^{(1)}(t,v)d\Lambda_0(v) \end{pmatrix}^T J_{11}^{-1}J_{12}(u) - \kappa_{332}^{(0)}(t,u)\right\}.$$

The existence and uniqueness of the solution to the linear integral equation for a given value of $\check{\Lambda}(\cdot)$ are well known. Under the regularity conditions, we have

$$\int_0^\tau \left|\frac{\check{\lambda}(t)}{\kappa_{331}^{(0)}(t)}\right|^2 dt < \infty, \quad \int_0^\tau \int_0^\tau |K(t,u)|^2 \, dtdu < \infty.$$

By the classical existence and uniqueness theorem for Fredholm integral equations, there exists one and only one solution to the deterministic equation. It follows that the inverse operator Φ^{-1} exists and has the form

$$\Phi^{-1}(\check{\Lambda}) = \int_0^t \frac{d\check{\Lambda}(u)}{\kappa_{331}^{(0)}(u)} + \int_0^\tau \int_0^t \frac{R(v,u)}{\kappa_{331}^{(0)}(u)} dvd\check{\Lambda}(u),$$

where $R(t,u)$ is independent of $\check{\Lambda}(\cdot)$ and satisfies the equation

$$R(t,u) = K(t,u) + \int K(t,v)R(v,t)dv.$$

As the true parameter values satisfy $U_0(\boldsymbol{\xi}_0) = 0$, we have

$$\begin{aligned}\sqrt{n}U_n(\boldsymbol{\xi}_0) &= \sqrt{n}\{U_n(\boldsymbol{\xi}_0) - U_0(\boldsymbol{\xi}_0)\} \\ &= \sqrt{n}\left\{U_{1n}(\boldsymbol{\xi}_0) - U_{10}(\boldsymbol{\xi}_0), U_{2n}(\boldsymbol{\xi}_0) - U_{20}(\boldsymbol{\xi}_0), U_{3n}(\boldsymbol{\xi}_0) - U_{30}(\boldsymbol{\xi}_0), U_{4n}(t,\boldsymbol{\xi}_0) - U_{40}(t,\boldsymbol{\xi}_0)\right\}.\end{aligned}$$

Notice that $\sqrt{n}\{U_{1n}(\boldsymbol{\xi}_0) - U_{10}(\boldsymbol{\xi}_0), U_{2n}(\boldsymbol{\xi}_0) - U_{20}(\boldsymbol{\xi}_0), U_{3n}(\boldsymbol{\xi}_0) - U_{30}(\boldsymbol{\xi}_0)\}$ converges in law to $\mathcal{W}_1$ from the multivariate central limit theorem, as it can be rewritten as the summation of independently and identically distributed (i.i.d) random vectors. The process $\sqrt{n}\{U_{4n}(t,\boldsymbol{\xi}_0) - U_{40}(t,\boldsymbol{\xi}_0)\}$ is a sum of i.i.d. processes of bounded variation. From a lemma for the central limit theorem for processes of bounded variation (Van Der Vaart and Wellner (1996), Example 2.11.16), it converges to a tight Gaussian process $\mathcal{W}_2$ if the second moment is finite. Finally, by the continuous mapping theorem, the weak convergence of $\sqrt{n}U_n(\boldsymbol{\xi}_0)$ follows.

The final step is to show the stochastic approximation of

$$\begin{aligned}&\sqrt{n}\left\{(U_n - U_0)(\boldsymbol{\xi}_n) - (U_n - U_0)(\boldsymbol{\xi}_0)\right\} \\ &= \sqrt{n}\left[\{U_n(\cdot,\boldsymbol{\xi}_n) - U_0(\cdot,\boldsymbol{\xi}_n)\} - \{U_n(\cdot,\boldsymbol{\xi}_0) - U_0(\cdot,\boldsymbol{\xi}_0)\}\right] = o_{P*}(1).\end{aligned}$$

Denote $\mathcal{P}_n$ as the empirical measure and

$$\dot{l}(t,\boldsymbol{\xi},\mathcal{O}) = \left\{\dot{l}_1(\boldsymbol{\xi},\mathcal{O}), \dot{l}_2(\boldsymbol{\xi},\mathcal{O}), \dot{l}_3(\boldsymbol{\xi},\mathcal{O}), \dot{l}_4(t,\boldsymbol{\xi},\mathcal{O})\right\}.$$

Then we have

$$U_n(\boldsymbol{\xi}) = \mathcal{P}_n\dot{l}(\cdot,\boldsymbol{\xi},\mathcal{O}) = \frac{1}{n}\sum_{i=1}^{n}\dot{l}(\cdot,\boldsymbol{\xi},\mathcal{O}_i).$$

Denote the empirical process by $\mathcal{G}_n f = \sqrt{n}(\mathcal{P}_n f - \mathcal{P}_0 f)$, where $\mathcal{P}_0$ denotes the expectation under the true value $\boldsymbol{\xi}_0$. Then, $\sqrt{n}(U_n - U_0)(\boldsymbol{\xi}) = \mathcal{G}_n\dot{l}(\cdot,\boldsymbol{\xi},\mathcal{O})$ is the empirical process index by the class of functions

$$\left\{\dot{l}(t,\boldsymbol{\xi},\mathcal{O}), \boldsymbol{\xi} \in \mathcal{A}\times\mathcal{B}, t \in [0,\tau)\right\}.$$

Let the norm $||\cdot||_{\mathcal{H}}$ on $\mathcal{H} = \mathcal{A}\times\mathcal{B}$ be defined as $||\boldsymbol{\xi}||_{\mathcal{H}} = |\boldsymbol{\theta}| + |\boldsymbol{\alpha}| + |\boldsymbol{\beta}| + |\Lambda|_\infty$. Then the stochastic condition is

$$||\mathcal{G}_n\dot{l}(t, \widehat{\boldsymbol{\xi}}_n, \mathcal{O}) - \mathcal{G}_n\dot{l}(t, \boldsymbol{\xi}_0, \mathcal{O})||_{\mathcal{H}} = o_{P*}(1).$$

We can use the lemma from Van Der Vaart and Wellner (1996) to prove the stochastic approximation. First, it can be shown that $\left\{\dot{l}(t, \boldsymbol{\xi}, \mathcal{O}) - \dot{l}(t, \boldsymbol{\xi}_0, \mathcal{O}) : t \in [0, \tau), ||\boldsymbol{\xi} - \boldsymbol{\xi}_0||_{\mathcal{H}} < \delta\right\}$ is P_0-Donsker using the fact that the functions are the sum, products and continuous transformations of P_0-Donsker. Furthermore, $\dot{l}(t, \boldsymbol{\xi}, \mathcal{O})$ converges to $\dot{l}(t, \boldsymbol{\xi}_0, \mathcal{O})$ for each t and t when $||\boldsymbol{\xi} - \boldsymbol{\xi}_0||_{\mathcal{H}} \to 0$. The convergence also holds for the square moment by the dominated convergence theorem. We have

$$sup_{t\in[0,\tau)} E_0||\dot{l}(t, \boldsymbol{\xi}, \mathcal{O}) - \dot{l}(t, \boldsymbol{\xi}_0, \mathcal{O})||^2_{\mathcal{H}} \to 0.$$

The stochastic approximation of $(U_n - U_0)(\widehat{\boldsymbol{\xi}}_n) - (U_n - U_0)(\boldsymbol{\xi}_0)$ follows by Lemma 3.3.5 of Van Der Vaart and Wellner (1996) .

4. Sensitivity studies about model misspecification of the truncation time

We have conducted some sensitivity studies to evaluate the robustness of the proposed method with model misspecification of the truncation time. Specifically, we considered three scenarios: (i) the truncation time $\tilde{A}$ is generated from a truncated Gamma distribution while a truncated Weibull distribution is assumed when applying the proposed method; (ii) $\tilde{A}$ is generated from a truncated generalized Gamma distribution while a truncated Weibull distribution is assumed when applying the proposed method; and (iii) $\tilde{A}$ is generated from a truncated Gamma distribution while a truncated generalized Gamma distribution is assumed when applying the proposed method. The other aspects of data generation were the same as described in Section 5. The density function of a truncated generalized Gamma distribution is $h(t|\boldsymbol{\theta}) = g(t|\boldsymbol{\theta})/G(20|\boldsymbol{\theta})$ with $g(t|\boldsymbol{\theta}) = \theta_1 t^{\theta_1\theta_3 - 1} e^{-(t/\theta_2)^{\theta_1}} / \{\Gamma(\theta_3)\theta_2^{\theta_1\theta_3}\}$, and $G(t|\boldsymbol{\theta})$ is the cumulative density function, where $(\theta_1, \theta_2, \theta_3) > 0$. The generalized Gamma distribution degenerates to the Weibull distribution if

$\theta_3 = 1$, and degenerates to the Gamma distribution if $\theta_1 = 1$. The censoring rates were set to be 0%, 10% and 30%, respectively. We set n = 600 and used 500 replicates. For comparison, we also performed a naive analysis by ignoring the unique data structure. Specifically, we fitted a logistic regression model by excluding subjects with unknown values of Y_i, and then employed a Cox proportional hazards model for the left-truncated data by using subjects with $Y_i = 1$. Tables S1, S2 and S3 summarize the average estimates, empirical standard errors and average EM-aided standard errors for scenarios (i), (ii) and (iii), respectively.

As shown in the three tables, although the parametric truncation models under the three scenarios were misspecified, the estimated regression coefficients had small empirical biases. Under the setting without right censoring, the empirical biases were smaller than 4%. With increasing censoring rates, the empirical biases were stable and remained within a reasonable range ($\leqslant 5\%$). The variance estimation was also robust to the model misspecification on the truncation time; the standard errors estimated by the EM-aided procedure approximated the empirical standard errors well.

In summary, the proposed estimation method for the parameters of interest has robust performance with violations of the parametric model assumptions on the truncation time. Also, the simulation studies confirm that the proposed method works well under the rich parametric assumption. In practice, one can start with a rich family of parameters, e.g., the truncated generalized Gamma distribution, for the truncation time, which includes the Weibull distribution and the Gamma distribution as special cases.

5. Comparison of the length-biased method and the general left-truncated method

We have compared the small sample performance of the estimation method for the length-biased data and general left-truncated data when the data satisfy the stationarity assumption. We generated the truncation time $\tilde{A}$ from the uniform distribution with a truncation rate of 50%. We set n = 300, 600 or 1000, with 500 replications. Table S4 summarizes the average estimates, empirical standard errors and average EM-aided standard errors. We further calculated the relative efficiency (RE) of the two methods, defined as the mean squared error (MSE) from the general left-truncated data method, divided by that from the length-biased data method.

We made several observations from the simulation results. First, both methods performed well and had small empirical biases for the regression coefficients under the logistic model and Cox model. The estimated standard errors by both methods approximated the empirical standard errors well for estimated regression coefficients. Second, the method for the length-biased data was more statistically efficient, with the RE ranging from 1.00 to 1.38. Third, the estimated parameter θ in the truncated Weibull distribution obtained by the method for the general left-truncated data was very large, suggesting the estimated distribution of the truncation time was very close to the Uniform distribution (Figure 1). Also, variance estimation of this parameter was not stable due to the extreme value of the parameter. Last, the method for the length-biased data was more computationally efficient. For example, in a 100-run simulation with sample size 600 and 10% censoring rate using a 3.30GHz CPU desktop, the CPU times of the methods for the length-biased data and general left-truncated method were 2.12 hours and 5.40 hours, respectively. In practice, we suggested conducting a test for the stationary assumption (Addona and Wolfson, 2006; Asgharian *and others*, 2006). If the stationary assumption holds, the method for the length-biased data is preferred for improving statistical efficiency and computational simplicity.

6. Simulation studies with sample size 1000

Tables S5 and S6 summarize the simulation results with a sample size of 1000. The simulation results have the same patterns as those with smaller sample sizes.

7. Details about derivations of some equations

Derivation of $S_o(t|\boldsymbol{Z})$: Given that the population is a mixture of cured and uncured components, the marginal survival function of the observed time T is

$$\begin{aligned} S_o(t|\boldsymbol{Z}) &= P(T > t|\boldsymbol{Z}, \widetilde{Y} = 1)P(\widetilde{Y} = 1|\boldsymbol{Z}) + P(T > t|\boldsymbol{Z}, \widetilde{Y} = 0)P(\widetilde{Y} = 0|\boldsymbol{Z}) \\ &= \frac{S(t|\boldsymbol{Z})}{P(\widetilde{T} > \widetilde{A}|\boldsymbol{Z})} P(\widetilde{Y} = 1|\boldsymbol{Z}) + P(\widetilde{Y} = 0|\boldsymbol{Z}), \end{aligned}$$

where $S(t|\boldsymbol{Z}) = \exp\left\{-\Lambda(t)\exp(\boldsymbol{\beta}'\boldsymbol{Z})\right\}$ and $0 < t < 20$. Note in above equation, we have $P(T > t|\boldsymbol{Z}, \widetilde{Y} = 0) = 1$ because for the cured subgroup ($\widetilde{Y} = 0$), the time to the event of interest is not subject to truncation and the value is always 20.

Derivation of Equation 3.3:

To derive the likelihood of the observed data, we consider three scenarios:

(1) For an uncensored patient belonging to the uncured group ($\delta_i = 1, Y_i = 1$), the corresponding likelihood component is

$$\frac{P(\widetilde{Y}_i = 1|\boldsymbol{Z}_i, \boldsymbol{\alpha}) f(X_i|\boldsymbol{Z}_i, \boldsymbol{\beta}, \boldsymbol{\lambda})}{P(\widetilde{T}_i > \widetilde{A}_i|\boldsymbol{Z}_i, \boldsymbol{\psi})},$$

based on the conditional density function conditional on the sampling constraint $\tilde{T} > \tilde{A}$.

(2) For an uncensored patient belonging to the cured group ($\delta_i = 1, Y_i = 0$), the corresponding likelihood component is $P(\widetilde{Y}_i = 0|\boldsymbol{Z}_i, \boldsymbol{\alpha})$ because the actual time to the event of interest is not subject to truncation and the value of $\widetilde{T}_i$ is always 20.

(3) For a censored patient ($\delta_i = 0$), the group indicator is unobserved, and then the corresponding likelihood component is $S_o(X_i|\boldsymbol{Z}_i, \boldsymbol{\psi})$.

Equation 3.3 follows.

Derivation of Equation 3.4: For censored subject i, we are not able to observe its SAB status

(Y_i), and its expectation conditional on the observed data (O_i) is calculated as below.

$$
\begin{aligned}
&E(Y_i|\mathcal{O}_i,\boldsymbol{\psi}) = P(Y_i=1|\mathcal{O}_i,\boldsymbol{\psi}) = P(\widetilde{Y}_i=1|\widetilde{T}_i > x_i, \boldsymbol{Z}_i, \boldsymbol{\psi}) \\
&= \frac{P(\widetilde{Y}_i=1|\boldsymbol{Z}_i,\boldsymbol{\alpha})P(\widetilde{T}_i > x_i|\widetilde{Y}_i=1,\boldsymbol{Z}_i,\boldsymbol{\beta},\boldsymbol{\lambda})}{P(\widetilde{Y}_i=0|\boldsymbol{Z}_i,\boldsymbol{\alpha})P(\widetilde{T}_i > x_i|\widetilde{Y}_i=0,\boldsymbol{Z}_i,\boldsymbol{\beta},\boldsymbol{\lambda}) + P(\widetilde{Y}_i=1|\boldsymbol{Z}_i,\boldsymbol{\alpha})P(\widetilde{T}_i > x_i|\widetilde{Y}_i=1,\boldsymbol{Z}_i,\boldsymbol{\beta},\boldsymbol{\lambda})} \\
&= \frac{P(\widetilde{Y}_i=1|\boldsymbol{Z}_i,\boldsymbol{\alpha})S(x_i|\boldsymbol{Z}_i,\boldsymbol{\beta},\boldsymbol{\lambda})}{P(\widetilde{Y}_i=0|\boldsymbol{Z}_i,\boldsymbol{\alpha}) + P(\widetilde{Y}_i=1|\boldsymbol{Z}_i,\boldsymbol{\alpha})S(x_i|\boldsymbol{Z}_i,\boldsymbol{\beta},\boldsymbol{\lambda})}.
\end{aligned}
$$

Note that in the second line of this equation, we have $P(\widetilde{T}_i > x_i|\widetilde{Y}_i = 0, \boldsymbol{Z}_i, \boldsymbol{\beta}, \boldsymbol{\lambda}) = 1$ since for subjects in the cured group ($\widetilde{Y}_i = 0$), the actual time to the event of interest is fixed ($\widetilde{T}_i = 20$) and is always greater than x_i. Equation 3.4 follows.

References

Addona, Vittorio and Wolfson, David B. (2006). A formal test for the stationarity of the incidence rate using data from a prevalent cohort study with follow-up. *Lifetime data analysis* **12**(3), 267–284.

Asgharian, Masoud, Wolfson, David B and Zhang, Xun. (2006). Checking stationarity of the incidence rate using prevalent cohort survival data. *Statistics in medicine* **25**(10), 1751–1767.

Murphy, Susan A. (1995). Asymptotic theory for the frailty model. *The Annals of Statistics*, 182–198.

Parner, Erik *and others*. (1998). Asymptotic theory for the correlated gamma-frailty model. *The Annals of Statistics* **26**(1), 183–214.

Van Der Vaart, Aad W and Wellner, Jon A. (1996). *Weak Convergence and Empirical Processes*. Springer, New York.

Fig. 1. Truncated Weibull distribution with parameter values $\boldsymbol{\theta} = (1.02, 335752)$

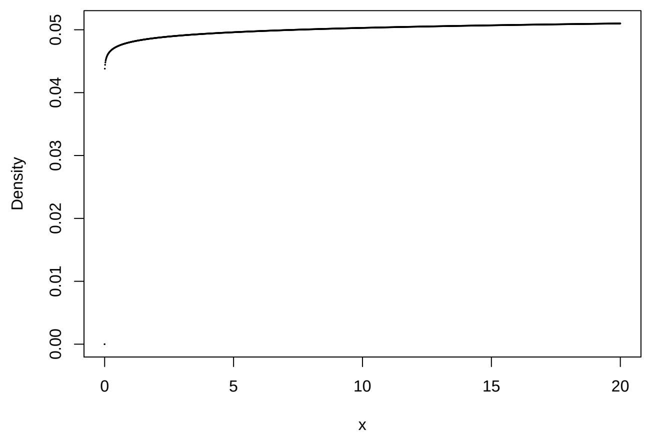

Table S1: Summary of sensitivity analysis for scenario (i). EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

			Proposed Method			**Naive Method**	
CENSOR	PARA	TRUE	EST	SD	ESE	EST	SD
0%	α_0	1.2	1.16	0.11	0.12	1.03	0.10
	α_1	1	1.03	0.21	0.18	1.09	0.20
	α_2	1	0.94	0.33	0.38	0.82	0.33
	β_1	-0.5	-0.50	0.10	0.10	-0.50	0.10
	β_2	1	1.01	0.18	0.17	1.01	0.18
10%	α_0	1.2	1.17	0.13	0.14	1.89	0.16
	α_1	1	1.04	0.25	0.22	1.00	0.31
	α_2	1	0.94	0.40	0.43	0.99	0.51
	β_1	-0.5	-0.50	0.13	0.12	-0.40	0.11
	β_2	1	1.01	0.22	0.21	0.81	0.20
30%	α_0	1.2	1.20	0.18	0.20	2.31	0.23
	α_1	1	1.05	0.35	0.29	0.92	0.43
	α_2	1	0.95	0.50	0.58	1.16	0.67
	β_1	-0.5	-0.50	0.15	0.15	-0.35	0.12
	β_2	1	1.01	0.26	0.25	0.71	0.21

Table S2: Summary of sensitivity analysis for scenario (ii). EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

			Proposed Method			**Naive Method**	
CENSOR	PARA	TRUE	EST	SD	ESE	EST	SD
0%	α_0	1.20	1.17	0.10	0.13	0.99	0.10
	α_1	1.00	1.02	0.20	0.19	1.10	0.20
	α_2	1.00	0.98	0.34	0.40	0.81	0.33
	β_1	-0.5	-0.50	0.10	0.10	-0.50	0.10
	β_2	1.00	1.01	0.19	0.17	1.00	0.19
10%	α_0	1.20	1.18	0.12	0.14	1.70	0.14
	α_1	1	1.04	0.24	0.22	1.04	0.29
	α_2	1.00	0.95	0.37	0.43	0.95	0.46
	β_1	-0.50	-0.50	0.12	0.12	-0.41	0.11
	β_2	1.00	1.01	0.20	0.20	0.84	0.19
30%	α_0	1.2	1.18	0.14	0.16	1.97	0.17
	α_1	1	1.04	0.26	0.24	1.00	0.33
	α_2	1	0.98	0.43	0.50	1.04	0.56
	β_1	-0.5	-0.50	0.14	0.13	-0.38	0.11
	β_2	1	0.99	0.24	0.22	0.75	0.20

Table S3: Summary of sensitivity analysis for scenario (iii). EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

			Proposed Method			**Naive Method**	
CENSOR	PARA	TRUE	EST	SD	ASE	EST	SD
0%	α_0	1.20	1.17	0.11	0.08	0.97	0.10
	α_1	1.00	1.02	0.20	0.20	1.11	0.20
	α_2	1.00	0.96	0.33	0.29	0.77	0.33
	β_1	-0.5	-0.50	0.10	0.10	-0.50	0.10
	β_2	1.00	1.01	0.18	0.18	1.00	0.18
10%	α_0	1.20	1.18	0.12	0.10	1.62	0.14
	α_1	1	1.03	0.23	0.23	1.03	0.27
	α_2	1.00	0.96	0.37	0.33	0.92	0.43
	β_1	-0.50	-0.50	0.12	0.12	-0.41	0.11
	β_2	1.00	1.01	0.21	0.21	0.83	0.19
30%	α_0	1.2	1.20	0.15	0.12	1.99	0.18
	α_1	1	1.04	0.29	0.29	0.98	0.36
	α_2	1	0.97	0.45	0.41	1.07	0.58
	β_1	-0.5	-0.49	0.14	0.14	-0.36	0.11
	β_2	1	1.01	0.25	0.24	0.74	0.21

Table S4: Summary of simulations for efficiency comparison: EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

			General Left-truncated Data Method			**Length-biased Data Method**		
N	PARA	TRUE	EST	SD	ESE	EST	SD	ESE
300	α_0	1.2	1.08	0.15	0.11	1.13	0.15	0.12
	α_1	1	1.08	0.24	0.23	1.07	0.24	0.23
	α_2	1	0.98	0.47	0.38	1.00	0.47	0.39
	β_1	-0.5	-0.49	0.15	0.14	-0.49	0.15	0.14
	β_2	1	1.06	0.25	0.25	1.04	0.24	0.25
	θ_1	-	1.06	-	-	-	-	-
	θ_2	-	146724.35	-	-	-	-	-
600	α_0	1.2	1.14	0.11	0.08	1.16	0.10	0.08
	α_1	1	1.04	0.19	0.16	1.03	0.19	0.16
	α_2	1	0.94	0.32	0.27	0.94	0.32	0.27
	β_1	-0.5	-0.50	0.10	0.10	-0.50	0.10	0.10
	β_2	1	1.01	0.17	0.18	1.00	0.17	0.17
	θ_1	-	1.02	-	-	-	-	-
	θ_2	-	335752.49	-	-	-	-	-
1000	α_0	1.2	1.18	0.09	0.06	1.17	0.08	0.06
	α_1	1	1.02	0.14	0.12	1.02	0.14	0.12
	α_2	1	0.97	0.25	0.21	0.97	0.25	0.21
	β_1	-0.5	-0.50	0.08	0.08	-0.50	0.08	0.08
	β_2	1	1.01	0.13	0.13	1.00	0.13	0.13
	θ_1	-	1.02	-	-	-	-	-
	θ_2	-	213558.81	-	-	-	-	-

Table S5: Summary of simulation studies with length-biased data. EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

				Proposed Method			**Naive Method**	
N	CENSOR	PARA	TRUE	EST	SD	ESE	EST	SD
1000	0%	α_0	1.2	1.17	0.08	0.06	0.49	0.07
		α_1	1	1.02	0.14	0.12	1.25	0.14
		α_2	1	0.97	0.25	0.21	0.50	0.24
		β_1	-0.5	-0.50	0.08	0.08	-0.50	0.09
		β_2	1	1.00	0.13	0.13	1.00	0.15
	10%	α_0	1.2	1.17	0.08	0.07	0.69	0.08
		α_1	1	1.02	0.15	0.13	1.21	0.15
		α_2	1	0.97	0.26	0.22	0.59	0.27
		β_1	-0.5	-0.50	0.09	0.08	-0.44	0.09
		β_2	1	1.00	0.14	0.14	0.89	0.15
	30%	α_0	1.2	1.18	0.09	0.08	0.78	0.10
		α_1	1	1.03	0.17	0.15	1.18	0.18
		α_2	1	0.97	0.30	0.25	0.67	0.31
		β_1	-0.5	-0.50	0.10	0.09	-0.40	0.10
		β_2	1	1.01	0.16	0.16	0.83	0.17

Table S6: Summary of simulation studies with general left-truncated data. EST: empirical mean; SD: empirical standard deviation; ESE: average of asymptotic standard error estimates.

				Proposed Method			**Naive Method**	
N	CENSOR	PARA	TRUE	EST	SD	ESE	EST	SD
1000	0%	α_0	1.2	1.20	0.08	0.09	1.04	0.08
		α_1	1	1.01	0.15	0.14	1.07	0.15
		α_2	1	0.99	0.25	0.29	0.85	0.26
		β_1	-0.5	-0.50	0.08	0.08	-0.50	0.08
		β_2	1	1.01	0.12	0.13	1.00	0.13
		θ_1	1	1.00	0.03	0.03	1.01	0.02
		θ_2	2.8	2.79	0.12	0.14	2.47	0.08
	10%	α_0	1.2	1.20	0.09	0.10	1.85	0.11
		α_1	1	1.02	0.18	0.15	0.99	0.22
		α_2	1	0.99	0.29	0.33	1.00	0.38
		β_1	-0.5	-0.50	0.09	0.09	-0.41	0.08
		β_2	1	1.00	0.15	0.15	0.82	0.13
		θ_1	1	1.00	0.03	0.03	1.01	0.02
		θ_2	2.8	2.79	0.12	0.14	2.47	0.08
	30%	α_0	1.2	1.21	0.12	0.13	2.38	0.17
		α_1	1	1.02	0.24	0.20	0.92	0.33
		α_2	1	0.99	0.39	0.43	1.19	0.54
		β_1	-0.5	-0.49	0.11	0.11	-0.34	0.09
		β_2	1	0.99	0.19	0.19	0.68	0.15
		θ_1	1	1.00	0.03	0.03	1.01	0.02
		θ_2	2.8	2.80	0.12	0.14	2.47	0.08