Supporting Information for "Partially linear single-index generalized mean residual life models"

Peng Jin[1] and Mengling Liu*[1,2]

[1]Division of Biostatistics, Department of Population Health, New York University Grossman School of Medicine, New York, NY 10016, USA
[2]Department of Environmental Medicine, New York University Grossman School of Medicine, New York, NY 10016, USA

The Supporting Information contains additional simulation and data application results of Section 3.1 and 4 in the main manuscript, respectively. Regularity conditions and all technical proofs regarding Theorem 1 in Section 2.4 and the double robustness property in Section 2.5 are also available.

Web Appendix A: Additional simulation results

Table A.1 and A.2 present the simulation results of estimated single-index coefficients that referenced in Section 3.1 when censoring rate was 10%. In the case of a linear single-index function (S1), the estimated single-index coefficents under both GMRL model and PLSI GMRL model performed well. The efficiency loss of $\hat{\beta}$ from the PLSI GMRL model were small compared to the estimators from the GMRL model. When the true single-index function was a sine curve (S2), the estimators were low biased under both models, and the estimators from the PLSI GMRL were more efficient than the estimators from the GMRL model. In the case of a quadratic single-index function (S3), the estimators from the proposed PLSI GMRL model outperfromed those from the GMRL model in terms of biases, coverage probabilities, and efficiency.

Additional simulation results under a 20% censoring rate were summarized in Table A.3 for both $g(t) = \exp(t)$ and $g(t) = t$. Moreover, Table A.4 shows the simulation results under a 10% censoring rate with a baseline MRL function $m_0(t) = g^{-1}\{(-\frac{1}{3}t + 1)I(-\frac{1}{3}t + 1 \geq 0)\}$, mimicking an exponential distribution resulting from the right skewness for survival time when all covariates equal 0. We observed similar results as explained in Section 3.1.

Web Appendix B: Additional data application results

Figure B.1 visualizes the correlation matrix of Pearson correlation coefficients of five log-transformed biomarkers in the COVID-19 application, and Figure B.2 shows the standardized score process for model diagnosis. Figure B.3 shows the predicted recovery time within 30 days since hospitalization based on different procalcitonin levels.

Web Appendix C: Consistency and asymptotic normality of the estimators $\hat{\theta}$

Regularity conditions

One could use two types of asymptotic properties, increasing number of knots and fixed number of knots. For the reason of practicality, we assume the number of knots being fixed throughout the proof. The nonparametric single-index function $\psi(\cdot)$ is assumed to be a spline function with fixed knots. When the single-index function $\psi(\cdot)$ is smooth, the bias of $\hat{\psi}(\cdot)$ due to approximating $\psi(\cdot)$ by a spline is usually negligible compared with the variance of $\hat{\psi}(\cdot)$. To establish the asymptotic properties of the proposed estimators, we first outline the regularity conditions.

(C1) $H(t)$ converges almost surely to a nonrandom and bounded function $\tilde{H}(t)$ uniformly in $t \in [0, \tau]$.

(C2) V is bounded and G is continuous.

(C3) $m_0(t)$ is continuously differentiable on $[0,\tau]$.

(C4) $A := E\Big[\int_0^\tau I(T_i > t)\big[g'\{m_0(t)+\phi(V_i;\theta_*)\}\{\phi'(V_i;\theta_*)-\bar{v}(t)\}\{\phi'(V_i;\theta_*)\}^T - \frac{\partial\phi'(V_i;\theta_*)}{\partial\theta_*^T}\big[T_i - t - g\{m_0(t)+\phi(V_i;\theta_*)\}\big]\big]d\tilde{H}(t)\Big]$ is nonsingular where $\bar{v}(t)$ is the limit of $\bar{V}(t;\theta_*)$.

Proof of Theorem 1(i)

We first want to establish the consistency of the estimators $\hat{m}_0(t)$ and $\hat{\theta}$. Condition (C3) implies that $m_0(t)$ is of bounded variation on $[0, \tau]$. Therefore the processes

$$n^{-1}\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{G(T_i)}\big[T_i - t - g\{m_0(t) + s + \phi(V_i;\theta)\}$$

can be written as sums of products of monotone functions in t, s and all components of θ. Since monotone functions have pseudodimension 1, the processes are manageable (Bilias et al., 1997). It follows from the uniform strong law of large numbers (Lin et al., 2001) and the uniform consistency of $\hat{G}$ to G (Fleming and Harrington, 1991) that for any $\epsilon > 0$ and $r > 0$,

$$\begin{aligned} &n^{-1}\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t) + q + \phi(V_i;\theta)\}\big] \\ &\to E\big[F(t|V_i)(g\{m_*(t) + \phi(V_i;\theta_*)\} - g\{m_0(t) + q + \phi(V_i;\theta)\})\big] \end{aligned}$$

almost surely and uniformly in $q \in [0, r]$ and $\theta \in \Theta = \{\theta : \|\theta - \theta_*\| \le \epsilon\}$, where $F(t|V_i)$ denotes the survival function of T_i given V_i. Thus, except for a null set, for all large n, $t \in [0, \tau]$, $\theta \in \Theta$, and sufficiently large q,

$$n^{-1}\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t) + q + \phi(V_i;\theta)\}\big] < 0, \tag{1}$$

$$n^{-1}\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t) - q + \phi(V_i;\theta)\}\big] > 0. \tag{2}$$

By (1) and (2) and the monotonicity and continuity of $g(\cdot)$, any $t \in [0, \tau]$ and $\theta \in \Theta$, there exists a unique $\hat{m}_0(t;\theta)$ that satisfies

$$n^{-1}\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{\hat{m}_0(t;\theta) + \phi(V_i;\theta)\}\big] = 0. \tag{3}$$

To prove the existence and uniqueness of $\hat{m}_0(t)$ and $\hat{\theta}$, it suffices to show the existance of a unique solution to $U(\theta) = 0$. Take derivative of (3) with respect to θ, we have

$$\frac{\partial\hat{m}_0(t;\theta)}{\partial\theta} = -\frac{\sum_{i=1}^n \delta_i I(T_i > t)\hat{G}^{-1}(T_i)g'\{\hat{m}_0(t;\theta) + \phi(V_i;\theta)\}\phi'(V_i;\theta)}{\sum_{i=1}^n \delta_i I(T_i > t)\hat{G}^{-1}(T_i)g'\{\hat{m}_0(t;\theta) + \phi(V_i;\theta)\}} = -\bar{V}(t;\theta). \tag{4}$$

Let $\hat{A}(\theta) = -n^{-1}\partial U(\theta)/\partial\theta^T$. Based on (4), we have

$$\begin{aligned} \hat{A}(\theta) = n^{-1}\sum_{i=1}^n \int_0^\tau \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\Big[&g'\{\hat{m}_0(t;\theta) + \phi(V_i;\theta)\}\{\phi'(V_i;\theta) - \bar{V}(t;\theta)\}\{\phi'(V_i;\theta)\}^T \\ &- \frac{\partial\phi'(V_i;\theta)}{\partial\theta^T}\big[T_i - t - g\{\hat{m}_0(t;\theta) + \phi(V_i;\theta)\}\big]\Big]dH(t), \end{aligned}$$

which is nonnegative definite. By the uniform strong law of large numbers, it is easy to show $\hat{m}_0(t;\theta)$ converges almost surely to $m_0(t;\theta)$ uniformly in t and θ, and $m_0(t;\theta_*) = m_0(t)$. Similarity, $\bar{v}(t;\theta) = \lim_{n\to\infty} \bar{V}(t;\theta)$ exists and the convergence is almost surely and uniformly in t and θ. Thus, $\hat{A}(\theta)$ converges almost surely to a nonrandom $A(\theta)$ uniformly in θ, and it is obvious that $A(\theta_*) = A$, where

$$\begin{aligned} A(\theta) = E\bigg[\int_0^\tau F(t|V_i)\Big[& g'\{m_0(t;\theta)+\phi(V_i;\theta)\}\{\phi'(V_i;\theta)-\bar{v}(t;\theta)\}\{\phi'(V_i;\theta)\}^T \\ & - \big[T-t-g\{m_0(t;\theta)+\phi(V;\theta)\}\big]\frac{\partial\phi'(V_i;\theta)}{\partial\theta^T}\Big]d\tilde{H}(t)\bigg]. \end{aligned}$$

We can check that $n^{-1}U(\theta_*) \to 0$ almost surely and A is nonsingular by condition (C4). The uniform convergence of $\hat{A}(\theta)$ and the continuity of $A(\theta)$ imply that we can find a small neighborhood of θ_* where $\hat{A}(\theta)$ is positive definite when n is large enough. By the inverse function theorem (Rudin, 1976) that within a small neighborhood of θ_*, there exists a unique solution $\hat{\theta}$ to $U(\theta) = 0$ for every sufficiently large n. Condition (C4) also implies the global uniqueness of $\hat{\theta}$. Therefore, unique estimators $\hat{\theta}$ and $\hat{m}_0(t)(0 \le t \le \tau)$ exist. Note that above we showed $\hat{\theta}$ is strongly consistent. It then follows from the uniform convergence of $\hat{m}_0(t;\theta)$ to $m_*(t;\theta)$ that $\hat{m}_0(t) \doteq \hat{m}_0(t;\hat{\theta}) \to m_0(t;\theta_*) = m_0(t)$ almost surely uniformly in $[0,\tau]$.

Proof of Theorem 1(ii)

By the first order Taylor expansion, for any $t \in [0,\tau]$,

$$\begin{aligned} & \sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{\hat{m}_0(t;\theta_*)+\phi(V_i;\theta_*)\}\big] \\ = & \sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t)+\phi(V_i;\theta_*)\}\big] \\ & - \big[\hat{m}_0(t;\theta_*) - m_0(t)\big]\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)} g'\{m^*(t)+\phi(V_i;\theta_*)\}, \end{aligned}$$

where $m^*(t)$ lies between $\hat{m}_0(t|\theta_*)$ and $m_0(t)$. Then,

$$\hat{m}_0(t;\theta_*) - m_0(t) = \frac{\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t)+\phi(V_i;\theta_*)\}\big]}{\sum_{i=1}^n \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)} g'\{m^{**}(t)+\phi(V_i;\theta_*)\}}, \tag{5}$$

where $m^{**}(t)$ also lies between $\hat{m}_0(t;\theta_*)$ and $m_0(t)$. Also,

$$\begin{aligned} U(\theta_*) = \sum_{i=1}^n \int_0^\tau & \frac{\delta_i I(T_i > t)\phi'(V_i;\theta_*)}{\hat{G}(T_i)}\big[T_i - t - g\{m_0(t)+\phi(V_i;\theta_*)\} \\ & - g'\{m^*(t)+\phi(V_i;\theta_*)\}\{\hat{m}_0(t;\theta_*) - m_0(t)\}\big]dH(t) \end{aligned} \tag{6}$$

Combining (5) and (6), we get

$$\begin{aligned} U(\theta_*) &= \sum_{i=1}^n \int_0^\tau \frac{\delta_i I(T_i > t)}{\hat{G}(T_i)}\big[T_i - t - g\{m_*(t)+\phi(V_i;\theta_*)\}\big]\{\phi'(V_i;\theta_*) - \bar{V}^*(t)\}dH(t) \\ &= \sum_{i=1}^n \int_0^\tau M_i(t)\frac{G(T_i)}{\hat{G}(T_i)}\{\phi'(V_i;\theta_*) - \bar{V}^*(t)\}dH(t) \end{aligned}$$

where $\bar{V}^*(t) = \frac{\sum_{i=1}^n \frac{\delta_i I(T_i>t)}{\hat{G}(T_i)} g'\{m^*(t)+\phi(V_i;\theta_*)\}\phi'(V_i;\theta_*)}{\sum_{i=1}^n \frac{\delta_i I(T_i>t)}{\hat{G}(T_i)} g'\{m^{**}(t)+\phi(V_i;\theta_*)\}}$. Let $\bar{v}(t) := \bar{v}(t;\theta_*)$. Using the uniform consistency of $\hat{G}$ and $\hat{m}_0(t;\theta_*)$, and the uniform strong law of large numbers, we obtain that $\bar{V}^*(t)$ converges almost surely to $\bar{v}(t)$ uniformly in $t \in [0,\tau]$, and $\bar{v}(t)$ is a bounded function. Since $Pr\{C \ge \tau\} > 0$, we have

$$\sup_{0\le i\le n}\left|\frac{G(T_i)}{\hat{G}(T_i)} - 1\right| = O_p(n^{-1/2}).$$

Note that $\sum_{i=1}^n M_i(t) = O_p(n^{1/2})$ uniformly in $t \in [0,\tau]$ by the functional central limit theorem (Lin et al., 2000), then it follows from (C1) that

$$\begin{aligned}&\Big|\sum_{i=1}^n \int_0^\tau M_i(t)\{\bar{V}^*(t) - \bar{v}(t)\} dH(t)\Big| \\ &\le \int_0^\tau |\bar{V}^*(t) - \bar{v}(t)| \Big|\sum_{i=1}^n M_i(t)\Big| dH(t) = o_p(n^{1/2}),\end{aligned}$$

and

$$\Big|\sum_{i=1}^n \int_0^\tau M_i(t)\{\bar{V}^*(t) - \bar{v}(t)\}\Big(\frac{G(T_i)}{\hat{G}(T_i)} - 1\Big) dH(t)\Big| = o_p(n^{1/2}).$$

Therefore, we have

$$\sum_{i=1}^n \int_0^\tau M_i(t)\frac{G(T_i)}{\hat{G}(T_i)}\{\bar{V}^*(t) - \bar{v}(t)\} dH(t) = o_p(n^{1/2}), \tag{7}$$

and similarly,

$$\Big|\sum_{i=1}^n \int_0^\tau M_i(t)\frac{G(T_i)}{\hat{G}(T_i)}\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d[H - \tilde{H}](t)\Big| = o_p(n^{1/2}). \tag{8}$$

Thus, by (7) and (8), we have

$$\begin{aligned}U(\theta_*) &= \sum_{i=1}^n \int_0^\tau M_i(t)\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d\tilde{H}(t) \\ &\quad + \sum_{i=1}^n \int_0^\tau M_i(t)\Big(\frac{G(T_i)}{\hat{G}(T_i)} - 1\Big)\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d\tilde{H}(t) + o_p(n^{1/2}).\end{aligned} \tag{9}$$

The second term on the right-hand side of (9) equals

$$\begin{aligned}&\sum_{i=1}^n \int_0^\tau M_i(t)\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d\tilde{H}(t)\frac{G(T_i) - \hat{G}(T_i)}{G(T_i)} + o_p(n^{1/2}) \\ &= \sum_{i=1}^n \int_0^\tau \frac{Q^*(t)}{\hat{\pi}(t)} dM_i^c(t) + o_p(n^{1/2}),\end{aligned} \tag{10}$$

where

$$Q^*(t) = n^{-1}\sum_{i=1}^n I(T_i \ge t)\int_0^\tau M_i(u)\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d\tilde{H}(u).$$

By some calculations, it follows that

$$n^{-1/2}U(\theta_*) = n^{-1/2}\sum_{i=1}^n \xi_i + o_p(1),$$

where

$$\xi_i = \int_0^\tau M_i(t)\{\phi'(V_i;\theta_*) - \bar{v}(t)\} d\tilde{H}(t) + \int_0^\tau \frac{Q(t)}{\pi(t)} dM_i^c(t).$$

Utilizing the multivariate central limit theorem, $n^{-1/2}U(\theta_*)$ is aymptotically normal with mean zero and covariance matrix $\Sigma_{\theta_*} = E\{\xi_i^{\otimes 2}\}$. Thus, applying the Taylor expansion of $U(\hat{\theta})$ at θ_*, we have

$$n^{1/2}(\hat{\theta} - \theta_*) = A^{-1}n^{-1/2}U(\theta_*) + o_p(1) = A^{-1}n^{-1/2}\sum_{i=1}^n \xi_i + o_p(1).$$

It follows that $n^{1/2}(\hat{\theta} - \theta_*)$ is asymptotically zero-mean normal with covariance matrix $A^{-1}\Sigma_{\theta_*}(A^{-1})^T$, which can be estimated by $\hat{A}^{-1}\hat{\Sigma}_{\hat{\theta}}(\hat{A}^{-1})^T$ defined in Theorem 1. Thus, the proof is complete.

Web Appendix D: Double robustness property in Section 2.5

Here we prove $E[Y_{F,G}(\theta; T, \delta, V)|V] = 0$ when either F or G is correctly specified. For any survival functions F_1 and G_1, we have

$$Y_{F_1,G_1}(\theta; T, \delta, V) = \frac{\delta D(\theta; T, V)}{G_1(T)} - \frac{(1-\delta)Q_{F_1}(\theta; C, V)}{G_1(C)} = \int_0^T Q_{F_1}(\theta; c, V)\frac{dG_1(c)}{G_1(c)^2}. \tag{11}$$

Note that $E\big[\frac{\delta D(\theta;T,V)}{G_1(T)}|V\big] = -\int_0^\tau D(\theta; u, V)\frac{G(u)}{G_1(u)}dF(u|V)$. Also, we have

$$\begin{aligned}
E\big[\frac{(1-\delta)Q_{F_1}(\theta; C, V)}{G_1(C)}|V\big] &= E\big[\frac{F(C|V)}{G_1(C)}Q_{F_1}(\theta; C, V)|V\big] \\
&= E\big[\frac{F(C|V)}{F_1(C|V)}\int_C^\tau D(\theta; u, V)dF_1(u|V)G_1(C)^{-1}|V\big] \\
&= -\int_0^\tau D(\theta; u, V)\big[\int_0^u \frac{F(c|V)}{F_1(c|V)}G_1(c)^{-1}dG(c)\big]dF_1(u|V).
\end{aligned}$$

Finally, we observe that

$$\begin{aligned}
E\big[\int_0^T Q_{F_1}(\theta; c, V)\frac{dG_1(c)}{G_1(c)^2}|V\big] &= \int_0^\infty F(c|V)G(c)Q_{F_1}(\theta; c, V)\frac{dG_1(c)}{G_1(c)^2} \\
&= \int_0^\tau D(\theta; u, V)\big[\int_0^u \frac{F(c|V)}{F_1(c|Z)}\frac{G(c)}{G_1(c)^2}dG_1(c)\big]dF_1(u|V),
\end{aligned}$$

and $d[G(c)G_1(c)^{-1}] = G_1(c)^{-1}dG(c) - G(c)G_1(c)^{-2}dG_1(c)$. Therefore, we have

$$\begin{aligned}
E[Y_{F_1,G_1}(\theta; T, \delta, V)|V] = &-\int_0^\tau D(\theta; u, V)\frac{G(u)}{G_1(u)}dF(u|V) \\
&+ \int_0^\tau D(\theta; u, V)\big[\int_0^\tau \frac{F(c|V)}{F_1(c|V)}d[G(c)G_1(c)^{-1}]dF_1(u|V)\big]
\end{aligned} \tag{12}$$

When $F_1 = F$, then right side of (12) equals

$$\begin{aligned}
&-\int_0^\tau D(\theta; u, V)\big[\frac{G(u)}{G_1(u)} - \int_0^u d[G(c)G_1(c)^{-1}]\big]dF(u|V) \\
&= \int_0^\tau D(\theta; u, V)\big[\frac{G(u)}{G_1(u)} - \frac{G(u)}{G_1(u)} + 1\big]dF(u|V) \\
&= E[D(\theta; T, V)|V] = 0.
\end{aligned}$$

When $G_1 = G$, $d[G(c)G_1(c)^{-1}] = 0$, and right side of (12) is $-\int_0^\tau D(\theta; u, V)dF(u|V) = E[D(\theta; T, V)|V] = 0$, which completes the proof.

References

Bilias, Y., Gu, M., and Ying, Z. (1997). Towards a general asymptotic theory for Cox model with staggered entry. *The Annals of Statistics*, 25(2), 662-682.

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.

Lin, D. Y., Wei, L. J. and Ying, Z. (2001, June). Semiparametric Transformation Models for Point Processes. *Journal of the American Statistical Association* 96(454), 620-628.

Fleming, T. R. and Harrington, D. P. (1991). Counting Processes and Survival Analysis. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*, John Wiley and Sons Inc., New York.

Rudin, W. (1976): *Principles of Mathematical Analysis.* McGraw-Hill, New York, third edition.

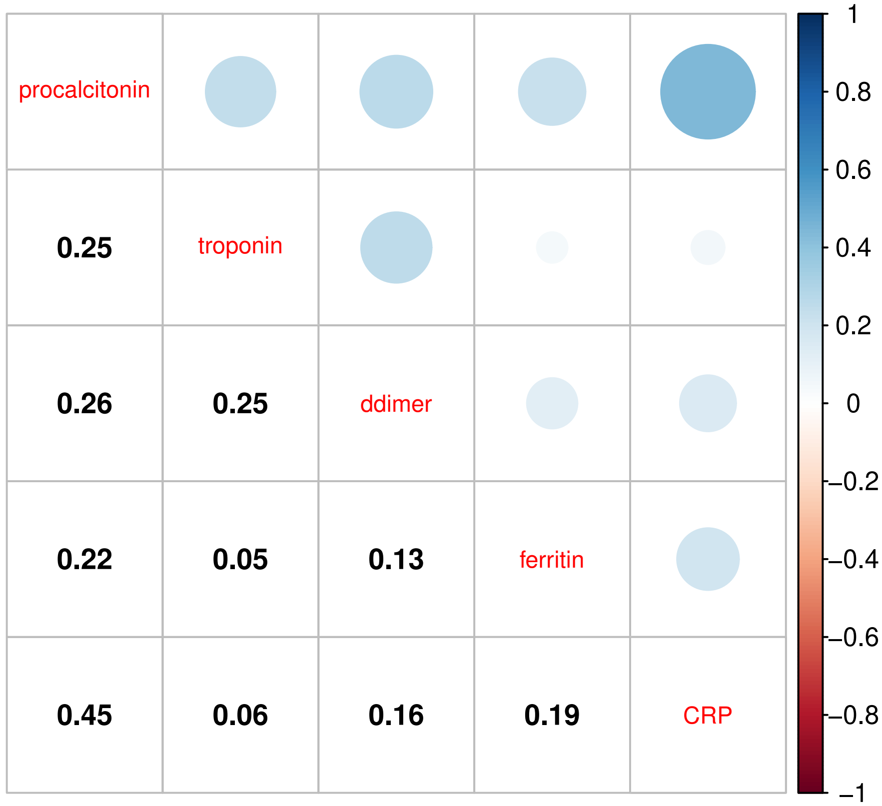

Figure B.1: Correlation matrix of Pearson correlation coefficients of five log-transformed biomarkers

Table A.1: Simulation results of estimated single-index coefficients under $g(t) = \exp(t)$ with independent censoring

	N	GMRL				PLSI GMRL			
		Bias	SD	SE	CP	Bias	SD	SE	CP
S1. Linear case									
β_1	2000	-0.005	0.066	0.064	92.8	-0.006	0.070	0.074	95.0
β_2	2000	-0.003	0.065	0.063	94.2	-0.003	0.070	0.074	95.6
β_3	2000	0.005	0.067	0.063	94.0	0.006	0.075	0.074	95.0
β_4	2000	0.004	0.065	0.064	94.0	0.005	0.069	0.073	97.2
β_1	4000	-0.001	0.044	0.045	95.4	-0.001	0.045	0.049	96.4
β_2	4000	-0.005	0.045	0.046	93.6	-0.005	0.047	0.049	96.6
β_3	4000	0.000	0.044	0.045	95.0	0.001	0.045	0.049	95.4
β_4	4000	0.001	0.044	0.045	95.8	0.002	0.046	0.048	96.8
S2. Sine curve case									
β_1	2000	-0.002	0.036	0.034	92.6	-0.002	0.031	0.031	93.6
β_2	2000	0.000	0.035	0.034	94.4	0.000	0.030	0.031	96.4
β_3	2000	0.002	0.034	0.034	95.8	0.001	0.030	0.031	96.2
β_4	2000	0.002	0.033	0.034	97.0	0.001	0.030	0.031	96.2
β_1	4000	-0.001	0.025	0.024	94.2	0.000	0.022	0.022	94.0
β_2	4000	-0.002	0.025	0.024	93.6	-0.001	0.022	0.022	95.4
β_3	4000	-0.001	0.024	0.024	94.6	0.000	0.022	0.022	95.4
β_4	4000	0.002	0.025	0.024	94.4	0.001	0.022	0.022	94.2
S3. Quadratic case									
β_1	2000	-0.095	0.256	0.244	90.2	-0.013	0.104	0.123	97.0
β_2	2000	-0.507	0.496	0.447	70.6	-0.003	0.104	0.135	99.2
β_3	2000	0.473	0.511	0.456	75.4	0.016	0.112	0.131	97.6
β_4	2000	0.497	0.515	0.456	71.6	0.015	0.113	0.131	97.8
β_1	4000	-0.079	0.262	0.246	90.6	-0.010	0.075	0.080	94.8
β_2	4000	-0.454	0.503	0.450	74.8	-0.006	0.078	0.081	96.0
β_3	4000	0.475	0.503	0.451	73.4	0.002	0.078	0.080	96.6
β_4	4000	0.450	0.495	0.447	73.8	0.007	0.074	0.080	94.8

GMRL: generalized mean residual life model; *PLSI GMRL*: partially linear single-index generalized mean residual life model; *SD*: sample standard deviation; *SE*: mean of estimated standard error; *CP*: empirical coverage probability of 95% confidence interval; Censoring rate was 10%.

Table A.2: Simulation results of estimated single-index coefficients under $g(t) = t$ with independent censoring

	N	GMRL				PLSI GMRL			
		Bias	SD	SE	CP	Bias	SD	SE	CP
S1. Linear case									
β_1	2000	-0.004	0.095	0.085	90.5	-0.011	0.104	0.114	96.2
β_2	2000	-0.006	0.092	0.087	92.5	-0.011	0.102	0.134	98.0
β_3	2000	-0.013	0.089	0.087	95.3	-0.011	0.102	0.133	98.0
β_4	2000	0.012	0.092	0.087	94.1	-0.014	0.102	0.139	96.8
β_1	4000	-0.002	0.066	0.061	94.0	-0.004	0.069	0.072	95.6
β_2	4000	-0.002	0.062	0.062	94.6	-0.004	0.065	0.075	95.8
β_3	4000	-0.004	0.069	0.062	92.2	-0.007	0.073	0.074	95.4
β_4	4000	0.008	0.062	0.062	95.6	0.004	0.065	0.077	97.4
S2. Sine curve case									
β_1	2000	-0.006	0.063	0.059	94.4	-0.014	0.064	0.073	96.2
β_2	2000	-0.005	0.062	0.060	93.2	-0.011	0.061	0.075	96.2
β_3	2000	0.000	0.060	0.059	95.2	-0.007	0.059	0.073	96.6
β_4	2000	0.005	0.065	0.059	93.2	-0.003	0.062	0.073	97.4
β_1	4000	-0.003	0.043	0.044	94.4	-0.006	0.043	0.045	95.8
β_2	4000	0.001	0.043	0.044	95.0	-0.002	0.042	0.047	96.6
β_3	4000	0.001	0.043	0.044	96.4	-0.001	0.042	0.047	96.8
β_4	4000	0.003	0.043	0.044	94.4	0.000	0.042	0.046	95.8
S3. Quadratic case									
β_1	2000	-0.092	0.258	0.239	90.8	-0.005	0.093	0.099	96.2
β_2	2000	-0.410	0.496	0.442	77.4	-0.002	0.096	0.105	97.6
β_3	2000	0.414	0.499	0.441	77.0	0.009	0.087	0.105	97.4
β_4	2000	0.401	0.499	0.443	80.0	0.016	0.088	0.102	98.0
β_1	4000	-0.050	0.276	0.229	89.4	-0.003	0.059	0.065	96.6
β_2	4000	-0.402	0.477	0.437	79.6	-0.008	0.060	0.065	95.6
β_3	4000	0.428	0.487	0.434	76.6	0.001	0.059	0.065	96.6
β_4	4000	0.407	0.484	0.437	79.2	0.002	0.061	0.064	96.8

GMRL: generalized mean residual life model; *PLSI GMRL*: partially linear single-index generalized mean residual life model; *SD*: sample standard deviation; *SE*: mean of estimated standard error; *CP*: empirical coverage probability of 95% confidence interval; Censoring rate was 10%.

Table A.3: Simulation results with independent censoring (20% censoring rate)

	N	α_1				α_2				$\omega(\beta_*, \hat{\beta})$	
		Bias	SD	SE	CP	Bias	SD	SE	CP	Mean	SD
						$g(t) = \exp(t)$					
S1. Linear case											
GMRL	2000	-0.004	0.099	0.097	94.2	0.001	0.095	0.097	94.2	7.54	3.14
PLSI GMRL	2000	-0.004	0.099	0.097	93.8	0.000	0.095	0.097	94.4	8.04	3.56
GMRL	4000	0.000	0.069	0.068	93.2	0.003	0.066	0.068	94.4	5.33	2.20
PLSI GMRL	4000	0.000	0.069	0.068	92.8	0.003	0.066	0.068	94.6	5.54	2.29
S2. Sine curve case											
GMRL	2000	0.004	0.088	0.088	93.8	0.001	0.087	0.088	94.4	4.09	1.76
PLSI GMRL	2000	0.005	0.086	0.086	94.0	0.001	0.086	0.086	94.6	3.64	1.55
GMRL	4000	0.003	0.061	0.062	96.0	-0.002	0.063	0.062	94.6	2.88	1.22
PLSI GMRL	4000	0.003	0.060	0.061	95.2	-0.002	0.063	0.061	94.0	2.56	1.07
S3. Quadratic case											
GMRL	2000	-0.001	0.098	0.096	94.0	-0.001	0.098	0.095	93.8	72.38	31.97
PLSI GMRL	2000	0.000	0.097	0.095	94.4	0.000	0.095	0.095	95.0	12.90	6.55
GMRL	4000	-0.004	0.071	0.067	92.2	0.004	0.065	0.067	94.4	75.42	30.80
PLSI GMRL	4000	-0.004	0.070	0.066	92.6	0.003	0.064	0.067	95.0	8.54	3.87
						$g(t) = t$					
S1. Linear case											
GMRL	2000	-0.006	0.131	0.119	91.1	-0.002	0.129	0.119	93.1	11.21	5.21
PLSI GMRL	2000	-0.006	0.130	0.119	91.9	-0.003	0.129	0.118	93.3	14.33	8.73
GMRL	4000	-0.003	0.090	0.088	93.2	0.001	0.090	0.088	93.6	8.22	3.67
PLSI GMRL	4000	-0.003	0.090	0.087	93.2	-0.001	0.090	0.088	94.0	9.24	4.79
S2. Sine curve case											
GMRL	2000	-0.018	0.282	0.292	91.3	-0.015	0.287	0.295	92.7	8.41	3.49
PLSI GMRL	2000	-0.016	0.280	0.288	90.9	-0.015	0.282	0.291	92.5	8.74	4.96
GMRL	4000	0.008	0.211	0.218	92.2	0.017	0.211	0.217	93.8	5.89	2.58
PLSI GMRL	4000	0.008	0.208	0.216	92.4	0.017	0.209	0.215	93.6	5.60	3.06
S3. Quadratic case											
GMRL	2000	-0.006	0.059	0.058	94.4	-0.006	0.063	0.058	92.2	63.87	34.22
PLSI GMRL	2000	-0.003	0.057	0.056	94.8	-0.005	0.058	0.055	93.8	11.91	6.33
GMRL	4000	-0.004	0.043	0.042	93.2	0.001	0.046	0.042	92.2	65.05	34.35
PLSI GMRL	4000	-0.001	0.041	0.040	94.2	0.004	0.043	0.039	92.6	7.30	3.64

GMRL: generalized mean residual life model; *PLSI GMRL*: partially linear single-index generalized mean residual life model; *SD*: sample standard deviation; *SE*: mean of estimated standard error;*CP*: empirical coverage probability of 95% confidence interval;$\omega(\beta_*, \hat{\beta})$ was calculated by $\arccos(\langle\beta_*, \hat{\beta}\rangle/\|\hat{\beta}\| \cdot \|\beta_*\|)$.

Table A.4: Simulation results under independent censoring with a baseline MRL function $m_0(t) = g^{-1}\{(-\frac{1}{3}t+1)I(-\frac{1}{3}t+1 \geq 0)\}$

	N	α_1				α_2				$\omega(\beta_*, \hat{\beta})$	
		Bias	SD	SE	CP	Bias	SD	SE	CP	Mean	SD
						$g(t) = \exp(t)$					
S1. Linear case											
GMRL	2000	0.004	0.121	0.123	95.2	-0.001	0.125	0.122	94.8	9.33	4.02
PLSI GMRL	2000	0.004	0.121	0.123	95.4	-0.002	0.126	0.123	94.8	10.45	4.82
GMRL	4000	0.002	0.087	0.087	94.4	0.007	0.089	0.087	94.6	6.61	2.74
PLSI GMRL	4000	0.002	0.087	0.087	94.6	0.007	0.089	0.087	94.6	6.98	3.15
S2. Sine curve case											
GMRL	2000	-0.004	0.120	0.119	94.8	0.004	0.119	0.120	95.2	5.27	2.34
PLSI GMRL	2000	-0.002	0.120	0.118	94.4	0.004	0.116	0.118	94.8	4.60	2.02
GMRL	4000	0.001	0.088	0.084	93.0	0.002	0.085	0.085	94.2	3.87	1.58
PLSI GMRL	4000	0.001	0.087	0.083	92.2	0.003	0.085	0.083	93.8	3.34	1.36
S3. Quadratic case											
GMRL	2000	0.001	0.122	0.121	94.9	-0.009	0.119	0.120	95.8	77.29	31.43
PLSI GMRL	2000	0.002	0.121	0.121	95.8	-0.010	0.119	0.120	95.2	17.91	8.63
GMRL	4000	-0.004	0.088	0.085	93.6	-0.001	0.082	0.086	95.8	72.43	32.06
PLSI GMRL	4000	-0.002	0.088	0.085	93.6	-0.002	0.081	0.085	96.6	12.29	6.03
						$g(t) = t$					
S1. Linear case											
GMRL	2000	-0.017	0.150	0.155	96.3	0.001	0.160	0.156	94.4	13.71	8.49
PLSI GMRL	2000	-0.017	0.151	0.156	95.7	0.001	0.160	0.156	94.0	17.87	8.82
GMRL	4000	0.001	0.117	0.110	91.8	0.001	0.109	0.109	94.4	9.93	4.08
PLSI GMRL	4000	0.001	0.117	0.110	91.8	0.001	0.110	0.110	94.6	11.78	5.85
S2. Sine curve case											
GMRL	2000	-0.023	0.298	0.298	95.6	-0.026	0.321	0.301	94.8	8.11	3.59
PLSI GMRL	2000	-0.022	0.299	0.297	94.4	-0.024	0.324	0.302	94.0	7.60	3.61
GMRL	4000	-0.004	0.216	0.214	94.2	-0.001	0.214	0.215	95.2	5.92	2.44
PLSI GMRL	4000	-0.003	0.217	0.214	93.8	0.001	0.212	0.215	95.4	5.56	2.47
S3. Quadratic case											
GMRL	2000	-0.015	0.144	0.139	95.1	-0.014	0.131	0.139	96.0	71.50	32.39
PLSI GMRL	2000	-0.014	0.144	0.138	95.3	-0.014	0.130	0.138	96.3	20.21	10.41
GMRL	4000	-0.002	0.094	0.098	96.1	0.004	0.100	0.098	95.0	69.72	31.46
PLSI GMRL	4000	0.002	0.094	0.097	95.9	0.005	0.099	0.097	94.8	12.61	6.39

GMRL: generalized mean residual life model; *PLSI GMRL*: partially linear single-index generalized mean residual life model; *SD*: sample standard deviation; *SE*: mean of estimated standard error; *CP*: empirical coverage probability of 95% confidence interval; $\omega(\beta_*, \hat{\beta})$ was calculated by $\arccos(\langle \beta_*, \hat{\beta} \rangle / \|\hat{\beta}\| \cdot \|\beta_*\|)$; Censoring rate was 10%.

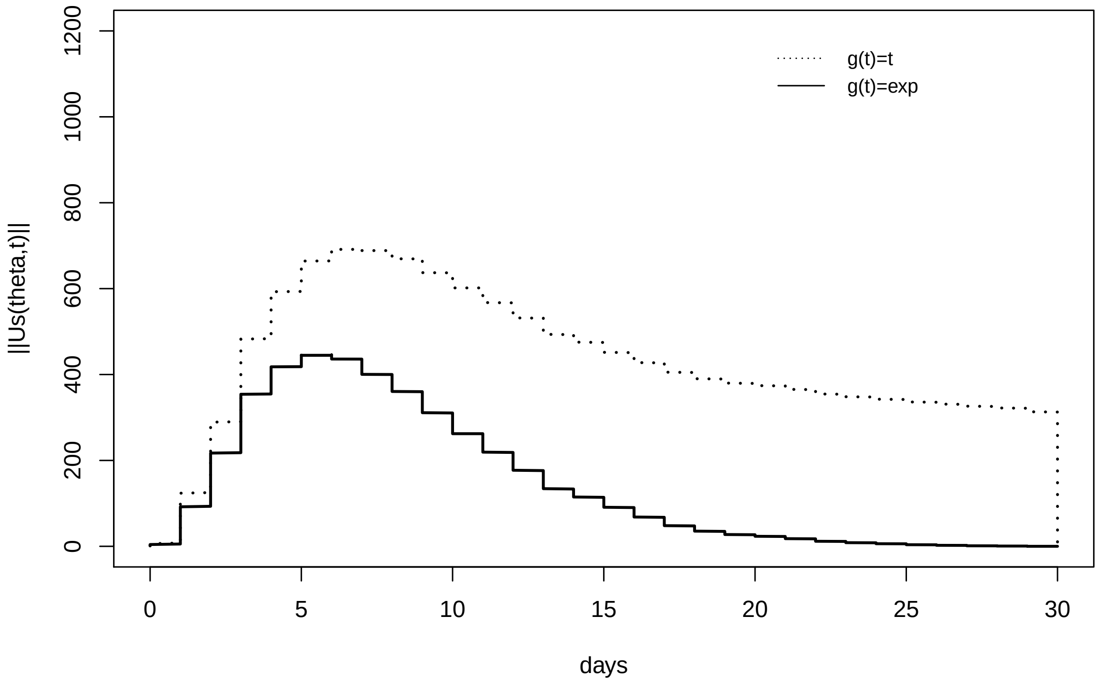

Figure B.2: Standardized score process for model diagnosis

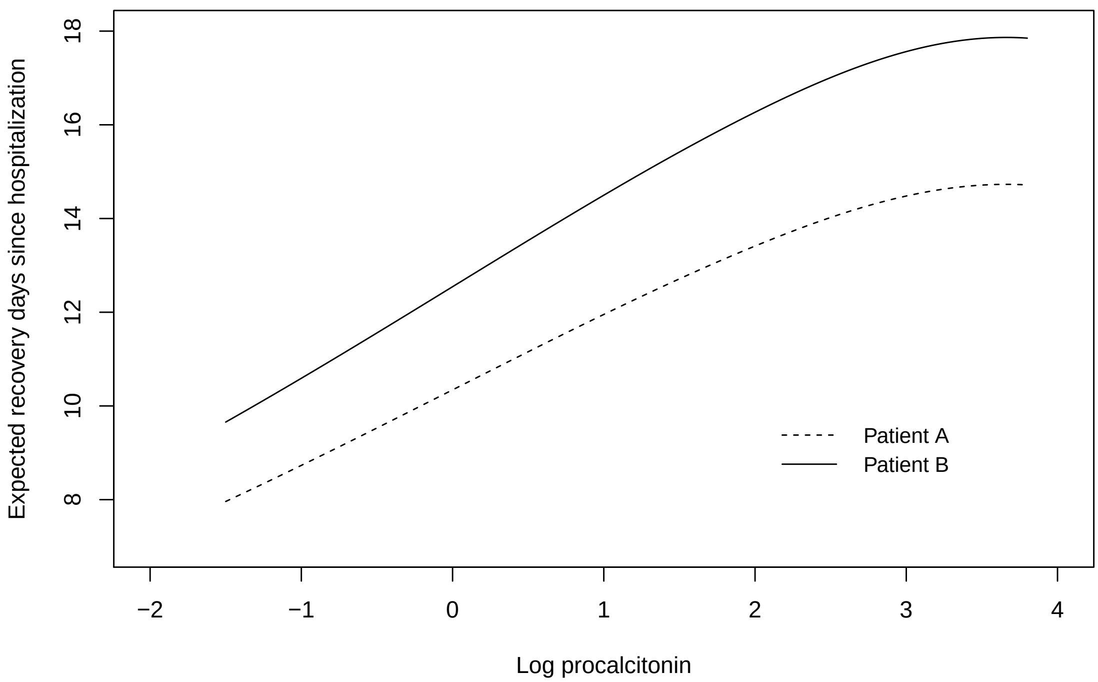

Figure B.3: Predicted recovery time within 30 days since hospitalization based on different procalcitonin levels. Patient A: a 35-year-old COVID-19 confirmed Caucasian female having 25 kg/m^2 BMI, 90% SpO_2, and 38.5 Celsius temperature with all log-transformed biomarkers at median values except for procalcitonin. Patient B: same profile as patient A but with a history of cardiac disease, diabetes, and malignancy.