REGRESSION ANALYSIS FOR CUMULATIVE INCIDENCE PROBABILITY UNDER COMPETING RISKS

Wei-Hwa Chang and Weijing Wang
Email: wjwang@stat.nctu.edu.tw
Institute of Statistics, National Chiao-Tung University
Hsin-Chu, Taiwan, R.O.C.

Supplementary Material

S1 Overview

The on-line supplement contains analysis of the severe acute respiratory syndrome (SARS) data, supplementary simulation reports and asymptotic analysis of the proposed methods. In the main text, two models are compared. The proposed model is given by

$$\begin{aligned} F_1(t \wedge \tau|\mathbf{Z}) &= \Pr(T \le \tau, \tilde{B} = 1|\mathbf{Z})\Pr(T \le t|T \le \tau, \tilde{B} = 1, \mathbf{Z}) \\ &= \pi(\mathbf{Z}^T\boldsymbol{\beta}(\tau))(1 - Q_{1,\mathbf{Z}}(t|\tau)). \end{aligned} \tag{S1.1}$$

The competitor of our model proposed by Fine and Gray (1999) and Fine (2001) can be written as

$$g(F_1(t|\mathbf{z})) = h(t) + \mathbf{z}^T\boldsymbol{\theta}. \tag{S1.2}$$

We will use the SARS example to compare the practical usefulness of the two models.

S2 Analysis of SARS data

S2.1 Data description

The Taiwan nationwide laboratory-confirmed SARS database was kindly provided by Dr. Mei-Shiang Ho and her colleagues in the institute of Biomedical Sciences, Academia Sinica. Patients with SARS had to be isolated in the hospital until recovery or death. The process can be described by the framework of competing risks. Here we define $\tilde{B} = 1$ to indicate that a patient was cured from disease (being discharged from the hospital and alive) and $\tilde{B} = 2$ to indicate that a patient was not cured (died during the isolation period). Because this infectious disease has been eventually under control in Taiwan, the database contains complete information about the two outcomes and the corresponding failure time. There are 258 infected patients in which 58 subjects were dead during the isolation period and 200 subjects were discharged from the hospital and alive.

Possible covariates include *age*, *gender*, *disease*, PCR, *viral load*, where *age* denotes a patient's age by years; *disease* is a binary variable indicating whether a patient had suffered from other diseases before getting infected of SARS (1: yes, 0: no); PCR is an indicator of whether the Polymerase Chain Reaction (PCR) test detected the SARS virus (1: yes, 0: no) and *viral load* measures the viral load detected by the PCR test. Note that if PCR equals 0, the individual had a negative virus titer, meaning that the patient has anti-body but zero viral load, and then the *viral load* is set as zero.

S2.2 Analysis of the original complete data

The function $F_1(t) = \Pr(\tilde{B} = 1, T \leq t)$ measures the probability of being discharged from the hospital (cured) by time t. We first present nonparametric analysis for each covariate group. Then we perform simple regression analysis for each covariate group using the LOGISTIC procedure in SAS.

Figures S.1-S.5 depict the empirical estimators of $F_1(t)$ based on the covariate groups. The continuous variable *age* was first divided into three groups, $age < 30$, $30 \leq age \leq 50$ and $age > 50$. Figure S.1 shows that the two younger groups ($age < 30$ and $30 \leq age \leq 50$) have similar patterns, while the older group ($age > 50$) has much lower chance of recovery at every t. At the end, the cure proportions of three age groups (from the youngest to the oldest) are 0.925, 0.879 and 0.443, respectively. The patterns of $F_1(t)$ for the *gender* groups and *disease* groups are similar such that the curves associated with different covariate values have no crossings. At the end, the female group (cure proportion = 0.842) had better recovery than the male group (cure proportion = 0.656).

Individuals without previous diseases (cure proportion = 0.845) also revealed better recovery than those in presence of other disease (cure proportion = 0.444).

The curves based on different groups of PCR and *viral load* behave differently from the former covariates. Note that *viral load*, originally measured continuously, was stratified into four groups: no viral load detected, $\leq 10^3$, $\in (10^3, 10^5]$, and $> 10^5$. The first group includes those with PCR = 0 (cure proportion = 0.946) and the last three groups are those with PCR = 1 (cure proportions equal 0.763, 0.648 and 0.526, respectively). At the end, the larger the level of viral load, the lower chance of recovery. However the four empirical curves have intersections in some middle time points.

We conducted several simple regression analyses based on the model

$$\text{logit}\,\{F_1(\tau_j)\} = \beta_{0,j} + \beta_{1,j}\ \text{z}, \tag{S2.1}$$

where z is a selected covariate and $\beta_{k,j}$ are simplifications of $\beta_k(\tau_j)$ $(k = 0, 1)$ for $j = 1, \ldots, 5$. In the analysis, *age* was divided into two groups (≤ 50 and > 50) and *viral load* was transformed into the scale of $\log_{10}$ to stabilize the effect caused by extreme large values. We set $\tau_1 = 14$, $\tau_2 = 21$, $\tau_3 = 28$, $\tau_4 = 35$ and τ_5 to be the maximum length of hospitalization for the cure satisfying $F_1(\tau_5) = \Pr(\tilde{B} = 1)$. The results are summarized in Table S.1. Here we discuss the effect of *age* for illustration. Treating the younger group ($age \leq 50$) as the baseline, the odds ratios along the time $e^{\hat{\beta}_{1,j}}$'s are 0.466, 0.221, 0.158, 0.146 and 0.090. This implies that the effect of *age* on the odds of $F_1(\tau_j)$ tends to be more influential as τ_j gets larger. Notice that *age* has substantial effect on $F_1(\tau_5)$, the final chance of recovery. For comparison, we analyze the data under model (S1.2) which assumes $\beta_{1,1} = \ldots = \beta_{1,5} = \theta$. The overall odds ratio $e^{\hat{\theta}}$ is 0.147 which seems very different from the separate odds ratios reported above. To formally examine whether the effect of *age* is time independent, a score test for assessing the difference between the reduced and the full model (with four degree of freedom) was performed. The resulting p-value is 0.006 which implies that model (S1.2) is not suitable for measuring the influence of *age* on $F_1(t)$. Table S.1 also shows that each covariate has a significant effect on $F_1(\tau_j)$ for larger τ_j. In general, younger females, who did not have other diseases and had lower viral load, had the best chance of recovery from SARS. Note that the effect of *gender* remained the same along the time. In fact, Figure S.2 shows that the curves for the male and the female do not intersect. However the two curves with different disease status have no crossing but the test of time homogeneity is rejected.

Although our paper does not study whether a covariate affects the latency distribution $Q_1(t|\tau)$, here we illustrate how to conduct further analysis if this is also of some interest. Let us use *age* again for illustration. We fit $Q_1(t|\tau_5)$ by the accelerated failure time model with a Weibull distribution, the estimated regression parameter for *age* is 0.276 (p-value = 0.004). The result implies that, for older patients ($age > 50$) who were eventually cured, they also needed longer time to get recovery (Figure S.6).

S2.3 Analysis of censored SARS data

As mentioned in the main paper, interim analysis based on incomplete data provides timely information for decision making. Although the original SARS dataset contains complete information about the value of $(T, \tilde{B})$, it is worthy to investigate how the proposed methods behave if this dataset is subject to further censoring. Here we generated a censoring variable which has a uniform distribution taking values from 0 to 70 making the censoring proportion to be around 30%.

Based on a censored version of the SARS data, we applied the proposed methods to fit a simple logistic regression for each covariate group and found that each covariate was statistically significant since time τ_2. Then we included all the covariates in the multiple logistic regression model which showed that the covariates *gender* and PCR became insignificant at all values of τ. The final fitted model is

$$logit(F_1(\tau)) = \beta_0(\tau) + \beta_1(\tau)\ age + \beta_2(\tau)\ disease + \beta_3(\tau)\ \log_{10}(viral\ load).$$

Table S.2 lists the detail results of the above analysis based on a single run using the artificial censoring scheme. Note that in the table we also report the previous results obtained from solving $\tilde{U}(\boldsymbol{\beta}) = 0$, the score function based on the original complete data. With the additional censoring, the proposed methods yield similar point estimates but larger standard deviations, as expected. Table S.3 list the average results by repeating the censoring scheme 300 times. The patterns are similar to that in a single run. Note that the proposed estimators also produce more precise results compared with the estimator of Fine (1999).

S3 Supplementary simulation results

S3.1 Data generation

Here we state the details of the data generation scheme. The covariate Z was generated from three distributions. For the discrete case, we set $Z \sim$ Bernoulli(0.5). For the continuous case, $Z \sim$ Normal$(0, 1)$ or $Z \sim$ Unif$(-3, 3)$. Let $\Delta_j = I(T \leq \tau, \tilde{B} = j)$ for $j = 1, 2$. Given Z, we set $\Delta_1 \sim$Bernoulli$(\pi(\beta_0 + \beta_1 Z))$ with

$$\pi(\beta_0 + \beta_1 Z) = \frac{\exp(\beta_0 + \beta_1 Z)}{1 + \exp(\beta_0 + \beta_1 Z)}.$$

If $\Delta_1 = 1$, we set $\Delta_2 = 0$; and if $\Delta_1 = 0$, we generated Δ_2 from a Bernoulli(p_2), where p_2 may depend on Z but its form is not of interest. Given (Δ_1, Δ_2), the failure time T

was generated from a distribution with the density function f_T which can be expressed as

$$f_T(t) = \begin{cases} f_1(t|\tau, Z) & \text{if} \quad (\Delta_1, \Delta_2) = (1, 0) \\ f_2(t|\tau, Z) & \text{if} \quad (\Delta_1, \Delta_2) = (0, 1) \\ f_3(t|\tau, Z) & \text{if} \quad (\Delta_1, \Delta_2) = (0, 0), \end{cases}$$

where $f_j(t|\tau, Z)$ $(j = 1, 2)$ are density functions with supports no greater than τ and $f_3(t|\tau, Z)$ is a density function whose value exceeds τ. In the simulations, we set

$$f_j(t|\tau, Z) = \frac{f_{Y_j}(t|Z)}{1 - S_{Y_j}(\tau|Z)} I(t \le \tau) \quad (j = 1, 2) \quad \text{and} \quad f_3(t|\tau, Z) = \frac{f_{Y_3}(t|Z)}{S_{Y_3}(\tau|Z)} I(t > \tau),$$

where $f_{Y_j}(t|Z)$ and $S_{Y_j}(t|Z)$ are the density and survival functions of Y_j which follows the accelerated failure-time model of the form,

$$\ln Y_j = \gamma_{0,j} + \gamma_{1,j} Z + \sigma_j \cdot W_j, \tag{S3.1}$$

where $\gamma_{0,j}$, $\gamma_{1,j}$ and σ_j are (nuisance) parameters and W_j is the error distribution.

The censoring variable was generated from Unif$(c_0, c_0 + c_1)$, where (c_0, c_1) are pre-specified constants making the censoring proportion to achieve the target value (i.e. 30% or 40%). Denoted $\{(\Delta_{1i}, \Delta_{2i}, T_i, Z_i, C_i)$ $(i = 1, \ldots, n)\}$ as a random sample of $(\Delta_1, \Delta_2, T, Z, C)$. Note that

$$I(X_i \le \tau, B_i = j) = \Delta_{ji} \cdot I(T_i \le C_i) \ (j = 1, 2),$$

where $X_i = T_i \wedge C_i$. The proposed methods can be implemented based on

$$\{X_i, I(T_i \le C_i), I(X_i \le \tau, B_i = 1), I(X_i \le \tau, B_i = 2)\}$$

for $i = 1, \ldots, n$. The value of τ is set to be 2.5. The sample size n was set to be 100 or 300.

S3.2 Analysis based on continuous covariates

In Tables S.4 and S.5, we report the simulation results when Z has the standard normal distribution and a uniform distribution respectively. Notice that, in Table S.4, $\hat{\boldsymbol{\beta}}_F$ still has large bias even when $n = 300$. We found that the IPCW technique, which utilizes $I(X_i \le \tau, B_i = 1)/\hat{G}(X_i-)$ as a proxy of $I(T_i \le \tau, \tilde{B}_i = 1)$, would make an observation with larger X_i to be more influential in the estimation. Our proposal by setting $V_{1i} = \pi(\mathbf{Z}_i^T \boldsymbol{\beta})(M_G - \pi(\mathbf{Z}_i^T \boldsymbol{\beta}))$ somewhat offset the influence of these observations. In contrast, Fine (1999) did not adjust the effect of censoring in his proposal of V_{1i} and hence $\hat{\boldsymbol{\beta}}_F$ was less stable.

Now we briefly describe the bootstrap procedure for variance estimation. Specifically 1000 sub-samples were drawn with replacement from the original sample, and for kth sub-sample, we obtained $\hat{\boldsymbol{\beta}}_{Ij}^{(k)}$ by solving $U_{Ij}(\boldsymbol{\beta}) = 0$ for $(j = 1, 2)$. Then the standard deviation of $\hat{\boldsymbol{\beta}}_{Ij}$ can be estimated by calculating the sample standard deviation of sub-estimates $\{\hat{\boldsymbol{\beta}}_{Ij}^{(k)} : k = 1, \ldots, 1000\}$ for $j = 1, 2$.

In Table S.6, we investigated whether the proposed methods remain robust when C actually depends on Z. We set $\ln C = \gamma_{0,c} + \gamma_{1,c}Z + \sigma_c W_c$, where Z may be binary or follow the standard normal distribution. In computation of the proposed estimators, we evaluated two estimators of $G(t) = \Pr(C > t)$. One is the Kaplan-Meier estimator and the other is a kernel-type smoothing estimator. Note that the former is based on the wrong assumption that C does not depend on Z. It turns out that the results based on the Kaplan-Meier estimator of $G(t)$ are biased while the kernel approach yields less biased estimators. All the proposed estimators are relatively more robust than $\hat{\boldsymbol{\beta}}_F$ under such a model mis-specification.

S4 Asymptotic analysis of the proposed methods

S4.1 Asymptotic properties of $U_{\mathrm{w}^*}(\boldsymbol{\beta})$

Assume that the true value $\boldsymbol{\beta}_0$ is located in the interior of the parameter space, which is a bounded convex region and $\pi_\phi(\cdot)$ is bounded. The estimating function $U_{\mathrm{w}^*}(\boldsymbol{\beta})$ in Section 2.2 of the main text can be written as

$$U_{\mathrm{w}^*}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[(V_{2i} - V_{3i})H_{1i} - (V_{1i} - V_{3i})H_{2i}\right] \frac{\pi_\phi(\mathbf{Z}_i^T\boldsymbol{\beta})}{V_{1i}V_{2i} - V_{3i}^2}\mathbf{Z}_i + B_{2n}(\boldsymbol{\beta}),$$

where

$$\begin{aligned} B_{2n}(\boldsymbol{\beta}) \;=\; & \sum_{i=1}^{n} \Bigg\{ \left[\frac{I(X_i \le \tau, B_i = 1)}{G(X_i-)} \frac{V_{2i} - V_{3i}}{V_{1i}V_{2i} - V_{3i}^2}\pi_\phi(\mathbf{Z}_i^T\boldsymbol{\beta})\mathbf{Z}_i\right] \frac{G(X_i-) - \hat{G}(X_i-)}{\hat{G}(X_i-)} \\ & - \left[\frac{I(X_i \le \tau, B_i = 2)}{G(X_i-)} \frac{V_{1i} - V_{3i}}{V_{1i}V_{2i} - V_{3i}^2}\pi_\phi(\mathbf{Z}_i^T\boldsymbol{\beta})\mathbf{Z}_i\right] \frac{G(X_i-) - \hat{G}(X_i-)}{\hat{G}(X_i-)} \\ & - \left[\frac{I(X_i > \tau)}{G(\tau)} \frac{V_{1i} - V_{3i}}{V_{1i}V_{2i} - V_{3i}^2}\pi_\phi(\mathbf{Z}_i^T\boldsymbol{\beta})\mathbf{Z}_i\right] \frac{G(\tau) - \hat{G}(\tau)}{\hat{G}(\tau)} \Bigg\}. \end{aligned}$$

To derive the asymptotic distribution of $n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta}_0)$, we first express the Kaplan-Meier estimator $\hat{G}(t)$ as the following integral form,

$$\frac{G(t)-\hat{G}(t)}{G(t)}=\sum_{i=1}^{n}\int_0^t\frac{\hat{G}(u-)}{G(u)}\frac{dM_{C,i}(u)}{\bar{Y}(u)},$$

where

$$M_{C,i}(u)=I(X_i\le u,B_i=0)-\int_0^u I(X_i\ge s)d\Lambda_C(s),$$

$\bar{Y}(u)=\sum_{i=1}^n I(X_i\ge u)$ and $\Lambda_C(s)$ is the cumulative hazard function of C. By the uniform convergence of the Kaplan-Meier estimator, we can write $n^{-1/2}B_{2n}(\boldsymbol{\beta}_0)$ as

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^{\infty}[q_1(t;\boldsymbol{\beta}_0)-q_2(t;\boldsymbol{\beta}_0)-q_3(t;\boldsymbol{\beta}_0)]\left(\frac{\bar{Y}(t)}{n}\right)^{-1}dM_{C,i}(t)+o_p(1),$$

where

$$q_1(t;\boldsymbol{\beta}_0)=\frac{1}{n}\sum_{k=1}^{n}I(X_k\ge t)\left[\frac{I(X_k\le\tau,B_k=1)}{G(X_k-)}\right]\frac{\mathrm{v}_{2k}-\mathrm{v}_{3k}}{\mathrm{v}_{1k}\mathrm{v}_{2k}-\mathrm{v}_{3k}^2}\pi_\phi(\mathbf{Z}_k^T\boldsymbol{\beta}_0)\mathbf{Z}_k, \tag{S4.1}$$

$$q_2(t;\boldsymbol{\beta}_0)=\frac{1}{n}\sum_{k=1}^{n}I(X_k\ge t)\left[\frac{I(X_k\le\tau,B_k=2)}{G(X_k-)}\right]\frac{\mathrm{v}_{1k}-\mathrm{v}_{3k}}{\mathrm{v}_{1k}\mathrm{v}_{2k}-\mathrm{v}_{3k}^2}\pi_\phi(\mathbf{Z}_k^T\boldsymbol{\beta}_0)\mathbf{Z}_k, \tag{S4.2}$$

$$q_3(t;\boldsymbol{\beta}_0)=\frac{1}{n}\sum_{k=1}^{n}I(\tau\ge t)\left[\frac{I(X_k>\tau)}{G(\tau)}\right]\frac{\mathrm{v}_{1k}-\mathrm{v}_{3k}}{\mathrm{v}_{1k}\mathrm{v}_{2k}-\mathrm{v}_{3k}^2}\pi_\phi(\mathbf{Z}_k^T\boldsymbol{\beta}_0)\mathbf{Z}_k, \tag{S4.3}$$

$\mathrm{v}_{1k}=\pi(\mathbf{Z}_k^T\boldsymbol{\beta}_0)(\tilde{M}-\pi(\mathbf{Z}_k^T\boldsymbol{\beta}_0))$, $\mathrm{v}_{2k}=\bar{\pi}(\mathbf{Z}_k^T\boldsymbol{\beta}_0)(\tilde{M}-\bar{\pi}(\mathbf{Z}_k^T\boldsymbol{\beta}_0))$, $\mathrm{v}_{3k}=\bar{\pi}(\mathbf{Z}_k^T\boldsymbol{\beta}_0)\pi(\mathbf{Z}_k^T\boldsymbol{\beta}_0)$ and $\tilde{M}$ is the median of the random variable $1/G(X-)$.

Therefore $n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta}_0)$ can be expressed as $n^{-1/2}\sum_{i=1}^n\xi_i+o_p(1)$, where

$$\begin{aligned}\xi_i &= \left\{\left[\frac{I(X_i\le\tau,B_i=1)}{G(X_i-)}-\pi(\mathbf{Z}_i^T\boldsymbol{\beta}_0)\right](\mathrm{v}_{2i}-\mathrm{v}_{3i})\right.\\ &- \left.\left[\frac{I(X_i>\tau)}{G(\tau)}+\frac{I(X_i\le\tau,B_i=2)}{G(X_i-)}-\bar{\pi}(\mathbf{Z}_i^T\boldsymbol{\beta}_0)\right](\mathrm{v}_{1i}-\mathrm{v}_{3i})\right\}\frac{\pi_\phi(\mathbf{Z}_i^T\boldsymbol{\beta}_0)}{\mathrm{v}_{1i}\mathrm{v}_{2i}-\mathrm{v}_{3i}^2}\mathbf{Z}_i\\ &+ \int_0^{\infty}\frac{q(t;\boldsymbol{\beta}_0)}{y(t)}dM_{C,i}(t),\end{aligned}$$

$y(t)=\lim_{n\to\infty}\bar{Y}(t)/n$ and $q(t;\boldsymbol{\beta}_0)=\lim_{n\to\infty}[q_1(t;\boldsymbol{\beta}_0)-q_2(t;\boldsymbol{\beta}_0)-q_3(t;\boldsymbol{\beta}_0)]$. Since $\{\xi_i\ (i=1,...,n)\}$ are zero-mean independent random variables, by the multivariate central limit theorem, $n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta}_0)$ has an asymptotic normal distribution with mean 0 and covariance matrix $\Gamma_{\mathrm{w}^*}=\lim_{n\to\infty}n^{-1}\sum_{i=1}^n\xi_i\xi_i^T$.

S4.2 Asymptotic properties of $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$

Recall that $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$ is the solution to $U_{\mathrm{w}^*}(\boldsymbol{\beta}) = 0$. Since $U_{\mathrm{w}^*}(\boldsymbol{\beta})$ is differentiable with respect to $\boldsymbol{\beta}$ and has a bounded derivative, consistency of $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$ follows. By a Taylor expansion of $n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}_0$, we can write

$$0 = n^{-1/2}U_{\mathrm{w}^*}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}) = n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta}_0) - A_{\mathrm{w}^*}(\boldsymbol{\beta}_0)\; n^{1/2}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*} - \boldsymbol{\beta}_0) + o_p(1),$$

where

$$A_{\mathrm{w}^*}(\boldsymbol{\beta}_0) = -\lim_{n\to\infty} \frac{1}{n}\frac{\partial U_{\mathrm{w}^*}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^T}\bigg|_{\boldsymbol{\beta}=\boldsymbol{\beta}_0}.$$

It follows that

$$n^{1/2}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*} - \boldsymbol{\beta}_0) = [A_{\mathrm{w}^*}(\boldsymbol{\beta}_0)]^{-1}\, n^{-1/2}U_{\mathrm{w}^*}(\boldsymbol{\beta}_0) + o_p(1). \tag{S4.4}$$

Hence $n^{1/2}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*} - \boldsymbol{\beta}_0)$ has an asymptotically normal distribution with mean 0 and covariance matrix $V_{\mathrm{w}^*} = [A_{\mathrm{w}^*}(\boldsymbol{\beta}_0)]^{-1}\,\Gamma_{\mathrm{w}^*}\,[A_{\mathrm{w}^*}(\boldsymbol{\beta}_0)]^{-1}$.

Replacing $\boldsymbol{\beta}_0$, G, $y(t)$ and $d\Lambda_C(t)$ by the corresponding estimates, $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$, $\hat{G}$, $\bar{Y}(t)/n$ and $dN_C(t)/\bar{Y}(t)$, where $N_C(t) = \sum_k I(X_k \le t, B_k = 0)$, respectively, $\hat{\xi}_i$ equals

$$\begin{aligned}
&\left\{\left[\frac{I(X_i \le \tau, B_i = 1)}{\hat{G}(X_i-)} - \pi(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\right](\hat{\mathrm{v}}_{2i} - \hat{\mathrm{v}}_{3i})\right.\\
&\left. - \left[\frac{I(X_i > \tau)}{\hat{G}(\tau)} + \frac{I(X_i \le \tau, B_i = 2)}{\hat{G}(X_i-)} - \bar{\pi}(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\right](\hat{\mathrm{v}}_{1i} - \hat{\mathrm{v}}_{3i})\right\}\frac{\pi_\phi(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})}{\hat{\mathrm{v}}_{1i}\hat{\mathrm{v}}_{2i} - \hat{\mathrm{v}}_{3i}^2}\mathbf{Z}_i\\
&+\frac{nI(B_i = 0)\hat{q}(X_i;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})}{\sum_{k=1}^n I(X_k \ge X_i)} - \sum_{j=1}^n \frac{nI(B_j = 0, X_i \ge X_j)\hat{q}(X_j;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})}{\left(\sum_{k=1}^n I(X_k \ge X_j)\right)^2},
\end{aligned}$$

where $\hat{\mathrm{v}}_{1i} = \pi(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})(M_G - \pi(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}))$, $\hat{\mathrm{v}}_{2i} = \bar{\pi}(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})(M_G - \bar{\pi}(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}))$,

$$\hat{\mathrm{v}}_{3i} = \bar{\pi}(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\pi(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}),\;\; \hat{q}(t;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}) = \hat{q}_1(t;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}) - \hat{q}_2(t;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}) - \hat{q}_3(t;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}),$$

and $\hat{q}_j(t;\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})$ $(j = 1, 2, 3)$ are obtained by using $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$, $\hat{G}$ and $(\hat{\mathrm{v}}_{1k}, \hat{\mathrm{v}}_{2k}, \hat{\mathrm{v}}_{3k})$ instead of $\boldsymbol{\beta}_0$, G and $(\mathrm{v}_{1k}, \mathrm{v}_{2k}, \mathrm{v}_{3k})$ in (S4.1)−(S4.3). It follows that the covariance matrix Γ_{w^*} can be estimated by $\hat{\Gamma}_{\mathrm{w}^*} = n^{-1}\sum_{i=1}^n \hat{\xi}_i\hat{\xi}_i^T$ and then

$$\hat{V}_{\mathrm{w}^*} = \left[\hat{A}_{\mathrm{w}^*}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\right]^{-1}\hat{\Gamma}_{\mathrm{w}^*}\left[\hat{A}_{\mathrm{w}^*}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\right]^{-1}$$

where

$$\hat{A}_{\mathrm{w}^*}(\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}) = \sum_{i=1}^n \frac{1}{n}\left[\frac{\hat{\mathrm{v}}_{1i} + \hat{\mathrm{v}}_{2i} - 2\hat{\mathrm{v}}_{3i}}{\hat{\mathrm{v}}_{1i}\hat{\mathrm{v}}_{2i} - \hat{\mathrm{v}}_{3i}^2}\pi_\phi^2(\mathbf{Z}_i^T\hat{\boldsymbol{\beta}}_{\mathrm{w}^*})\mathbf{Z}_i\mathbf{Z}_i^T\right].$$

S4.3 Previous nonparametric results of Wang (2003)

Modifying the idea of Wang (2003), we can estimate $p_j(x) = \Pr(T \leq \tau, \tilde{B} = j | T > x)$ by

$$\hat{p}_j(x) = \frac{1}{n\hat{S}(x)} \sum_{i=1}^{n} \frac{I(x < X_i \leq \tau, B_i = j)}{\hat{G}(X_i-)},$$

where $\hat{S}(x)$ is the Kaplan-Meier estimator of $S(x)$ which, according to Satten and Datta (2001), can be re-expressed as an average of inverse probability of censoring given by

$$\frac{1}{n} \sum_{i=1}^{n} \left[\frac{I(X_i > x, B_i \neq 0)}{\hat{G}(X_i-)} + \frac{I(X_i > X_{(m)})}{\hat{G}(X_{(m)})} \right],$$

where $X_{(m)}$ denotes the largest observed failure time. Based on Wang's idea, $Q_j(t|\tau)$ can be estimated by

$$\prod_{u \leq t} \left\{ 1 - \frac{\sum_{i=1}^{n} I(u = X_i \leq \tau, B_i = j)}{\sum_{i=1}^{n} \left[I(u \leq X_i \leq \tau, B_i = j) + I(u \leq X_i \leq \tau, B_i = 0)\hat{p}_j(X_i) \right]} \right\}.$$

S4.4 Asymptotic properties of $U_{I1}(\boldsymbol{\beta})$

Suppose that $\mathbf{Z}$ takes K distinct values, $z_1, \ldots, z_K$. Original data are partitioned into K mutually exclusive subsets, $\left\{ \left(\Delta_{1k}^j, X_k^j, B_k^j \right) \ (k = 1, \ldots, n_j) \right\}$, which corresponds to the set of $\{i : (\Delta_{1i}, X_i, B_i, \mathbf{Z}_i = z_j) \ (i = 1, \ldots, n)\}$ and $n_j = \sum_{i=1}^{n} I(\mathbf{Z}_i = z_j)$. We have $p_{z_j}(X_k^j) = E(\Delta_{1k}^j \big| X_k^j, B_k^j = 0, \mathbf{Z} = z_j)$, which can be estimated by

$$\hat{p}_{z_j}(X_k^j) = \frac{1}{n_j \hat{S}_{z_j}(X_k^j)} \sum_{h=1}^{n_j} \frac{I(X_k^j < X_h^j \leq \tau, B_h^j = 1)}{\hat{G}_{z_j}(X_h^j-)},$$

where $\hat{S}_{z_j}(t)$ and $\hat{G}_{z_j}(t)$ are Kaplan-Meier estimators of $S_{z_j}(t) = \Pr(T > t | \mathbf{Z} = z_j)$ and $G_{z_j}(t) = \Pr(C > t | \mathbf{Z} = z_j)$. The estimating equation $U_{I1}(\boldsymbol{\beta})$ can be re-expressed as

$$U_{I1}(\boldsymbol{\beta}) = \sum_{j=1}^{K} \left\{ \sum_{k=1}^{n_j} \left[\hat{\Delta}_{1k}^j - \pi(z_j^T \boldsymbol{\beta}) \right] \frac{\pi_\phi(z_j^T \boldsymbol{\beta})}{\pi(z_j^T \boldsymbol{\beta}) \bar{\pi}(z_j^T \boldsymbol{\beta})} z_j \right\},$$

where $\hat{\Delta}_{1k}^j = I(B_k^j = 1, X_k^j \leq \tau) + I(B_k^j = 0, X_k^j \leq \tau)\hat{p}_{z_j}(X_k^j)$.

To derive asymptotic distribution of $n^{-1/2} U_{I1}(\boldsymbol{\beta}_0)$, we first express it as sum of the

following two terms,

$$\begin{aligned}\frac{1}{\sqrt{n}}U_{I1}(\boldsymbol{\beta}_0) &= \sum_{j=1}^{K}\sqrt{\frac{n_j}{n}}\left\{\frac{1}{\sqrt{n_j}}\sum_{k=1}^{n_j}\left[E_k^j-\pi(z_j^T\boldsymbol{\beta}_0)\right]\Psi_{z_j}(\boldsymbol{\beta}_0)\right\}\\ &+ \sum_{j=1}^{K}\sqrt{\frac{n_j}{n}}\left\{\frac{1}{\sqrt{n_j}}\sum_{k=1}^{n_j}I(B_k^j=0)\left[\hat{p}_{z_j}(X_k^j)-p_{z_j}(X_k^j)\right]\Psi_{z_j}(\boldsymbol{\beta}_0)\right\}\end{aligned} \quad \text{(S4.5)}$$

where

$$E_k^j = E\left(\Delta_{1k}^j\middle|\, X_k^j, B_k^j; \mathbf{Z}=z_j\right) = I(B_k^j=1, X_k^j\le\tau) + I(B_k^j=0, X_k^j\le\tau)p_{z_j}(X_k^j)$$

and

$$\Psi_{z_j}(\boldsymbol{\beta}_0) = \frac{\pi_\phi(z_j^T\boldsymbol{\beta}_0)}{\pi(z_j^T\boldsymbol{\beta}_0)\bar{\pi}(z_j^T\boldsymbol{\beta}_0)}z_j.$$

Denote the last part of (S4.5) by $C_2(\boldsymbol{\beta}_0)$, by the strong consistency of Kaplan-Meier estimators, we have

$$C_2(\boldsymbol{\beta}_0) = \sum_{j=1}^{K}\left\{\sqrt{\frac{n_j}{n}}\Psi_{z_j}(\boldsymbol{\beta}_0)\left[C_{2.1}^j+C_{2.2}^j\right]\right\}+o_p(1),$$

where

$$C_{2.1}^j = \frac{1}{\sqrt{n_j}}\sum_{k=1}^{n_j}\left[\frac{I(B_k^j=0)}{n_jS_{z_j}(X_k^j)}\sum_{h=1}^{n_j}I(X_k^j<X_h^j\le\tau, B_h^j=1)\left(\frac{1}{\hat{G}_{z_j}(X_h^j-)}-\frac{1}{G_{z_j}(X_h^j-)}\right)\right],$$

$$C_{2.2}^j = \frac{1}{\sqrt{n_j}}\sum_{k=1}^{n_j}\left[I(B_k^j=0)\left(\frac{1}{\hat{S}_{z_j}(X_k^j)}-\frac{1}{S_{z_j}(X_k^j)}\right)\frac{1}{n_j}\sum_{h=1}^{n_j}\left(\frac{I(X_k^j<X_h^j\le\tau, B_h^j=1)}{G_{z_j}(X_h^j-)}\right)\right].$$

Interchanging the summations in $C_{2.1}^j$, we get

$$C_{2.1}^j = \frac{1}{\sqrt{n_j}}\sum_{h=1}^{n_j}\left[D(X_h^j)\frac{I(X_h^j\le\tau, B_h^j=1)}{G_{z_j}(X_h^j-)}\frac{\hat{G}_{z_j}(X_h^j-)-G_{z_j}(X_h^j-)}{G_{z_j}(X_h^j-)}\right]+o_p(1)$$

where

$$D(X_h^j) = \lim_{n_j\to\infty}\frac{1}{n_j}\sum_{k=1}^{n_j}\frac{I(B_k^j=0, X_k^j<X_h^j)}{S_{z_j}(X_k^j)}.$$

One can write

$$\frac{\hat{G}_{z_j}(t)-G_{z_j}(t)}{G_{z_j}(t)} = \sum_{l=1}^{n_j}\int_0^t\frac{\hat{G}_{z_j}(u-)}{G_{z_j}(u)}\frac{dM_{C,l}^j(u)}{\bar{Y}^j(u)}$$

where

$$\bar{Y}^j(u) = \sum_{i=1}^{n_j} I(X_i^j \geq u), \quad M_{C,l}^j(u) = I(X_l^j \leq u, B_l^j = 0) - \int_0^u I(X_l^j \geq s) d\Lambda_C^j(s),$$

and $\Lambda_C^j(s)$ is the cumulative hazard function of C given $\mathbf{Z} = z_j$. It follows that

$$C_{2.1}^j = \frac{1}{\sqrt{n_j}} \sum_{l=1}^{n_j} \int_0^\infty \frac{q^j(u)}{p^j(u)} dM_{C,l}^j(u) + o_p(1),$$

where

$$q^j(u) = \lim_{n_j \to \infty} \frac{1}{n_j} \sum_{h=1}^{n_j} D(X_h^j) \frac{I(u \leq X_h^j \leq \tau, B_h^j = 1)}{G_{z_j}(X_h^j -)} \quad \text{and} \quad p^j(u) = \lim_{n_j \to \infty} \frac{\bar{Y}^j(u)}{n_j}.$$

Similarly, one can write

$$C_{2.2}^j = \frac{1}{\sqrt{n_j}} \sum_{l=1}^{n_j} \int_0^\infty \frac{r^j(u)}{p^j(u)} \, dM_{T,l}^j(u) + o_p(1),$$

where

$$r^j(u) = \lim_{n_j \to \infty} \frac{1}{n_j} \sum_{k=1}^{n_j} \frac{I(B_k^j = 0, X_k^j \geq u) P_{z_j}(X_k^j)}{S_{z_j}(X_k^j)},$$

$$P_{z_j}(X_k^j) = \lim_{n_j \to \infty} \frac{1}{n_j} \sum_{h=1}^{n_j} \frac{I(B_h^j = 1, X_k^j < X_h^j \leq \tau)}{G_{z_j}(X_h^j -)},$$

and

$$M_{T,l}^j(u) = I(X_l^j \leq u, B_l^j \neq 0) - \int_0^u I(X_l^j \geq s) d\Lambda_T^j(s),$$

$\Lambda_T^j(s)$ is the cumulative hazard function of T given $\mathbf{Z} = z_j$.

In summary, we have

$$\frac{1}{\sqrt{n}} U_{I1}(\boldsymbol{\beta}_0) = \sum_{j=1}^{K} \sqrt{\frac{n_j}{n}} \left(\frac{1}{\sqrt{n_j}} \sum_{k=1}^{n_j} \zeta_k^j \right) \Psi_{z_j}(\boldsymbol{\beta}_0) + o_p(1)$$

where

$$\zeta_k^j = E_k^j - \pi(z_j^T \boldsymbol{\beta}_0) + \int_0^\infty \frac{q^j(u)}{p^j(u)} dM_{C,k}^j(u) + \int_0^\infty \frac{r^j(u)}{p^j(u)} dM_{T,k}^j(u).$$

Notice that $(\zeta_1^j, \ldots, \zeta_{n_j}^j)$ are zero-mean independent random variables for each j where $j = 1, \ldots, K$. By the multivariate central limit theorem, $\frac{1}{\sqrt{n}} U_{I1}(\boldsymbol{\beta}_0)$ has an asymptotical normal distribution with mean 0 and covariance matrix

$$\Gamma_{I1} = \lim_{n\to\infty} n^{-1} \sum_{j=1}^{K} \sum_{k=1}^{n_j} (\zeta_k^j)^2 \ \Psi_{z_j}(\boldsymbol{\beta}_0) \Psi_{z_j}^T(\boldsymbol{\beta}_0).$$

Let $\hat{\boldsymbol{\beta}}_{I1}$ be the solution of $U_{I1}(\boldsymbol{\beta}) = 0$. Asymptotic properties of $\hat{\boldsymbol{\beta}}_{I1}$ can be obtained as of $\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$ stated in section S4.2. According to (S4.4), $n^{1/2}(\hat{\boldsymbol{\beta}}_{I1} - \boldsymbol{\beta}_0)$ has an asymptotically normal distribution with mean 0 and covariance matrix $V_{I1} = [A_{I1}(\boldsymbol{\beta}_0)]^{-1} \Gamma_{I1} [A_{I1}(\boldsymbol{\beta}_0)]^{-1}$ where

$$A_{I1}(\boldsymbol{\beta}_0) = E\left[\frac{\pi_\phi^2(\mathbf{Z}^T\boldsymbol{\beta}_0)}{\pi(\mathbf{Z}^T\boldsymbol{\beta}_0)\bar{\pi}(\mathbf{Z}^T\boldsymbol{\beta}_0)} \mathbf{Z}\mathbf{Z}^T \right].$$

S4.5 Testing homogeneity of covariate effect over time

Now we illustrate how to use the results obtained for model (S1.1) to verify the assumption of model (S1.2) or help choosing time-dependent covariates in that model. Let $\boldsymbol{\beta}_0(\tau)$ be the true value of $\boldsymbol{\beta}(\tau)$. The assumption of model (S1.2) can be verified by testing $H_0 : \tilde{\boldsymbol{\beta}}_0(\tau_1) = \tilde{\boldsymbol{\beta}}_0(\tau_2)$ for any $\tau_1 \neq \tau_2$, where $\tilde{\boldsymbol{\beta}}_0(\tau)$ is the last p components of $\boldsymbol{\beta}_0(\tau)$. Specifically let

$$R = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}_{p\times(p+1)}.$$

Applying (S4.4), we can approximate the distribution of $\sqrt{n} R\{\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}(\tau_2) - \hat{\boldsymbol{\beta}}_{\mathrm{w}^*}(\tau_1)\}$ under H_0 by

$$\frac{1}{\sqrt{n}} R \ \{A_{\mathrm{w}^*}^{-1}(\mathbf{Z}^T\boldsymbol{\beta}_0(\tau_2)) \cdot U_{\mathrm{w}^*}(\boldsymbol{\beta}_0(\tau_2)) - A_{\mathrm{w}^*}^{-1}(\mathbf{Z}^T\boldsymbol{\beta}_0(\tau_1)) \cdot U_{\mathrm{w}^*}(\boldsymbol{\beta}_0(\tau_1))\},$$

which converges to a mean-zero p-dimensional normal random variable. The corresponding covariance matrix can be estimated as illustrated in section S4.2.

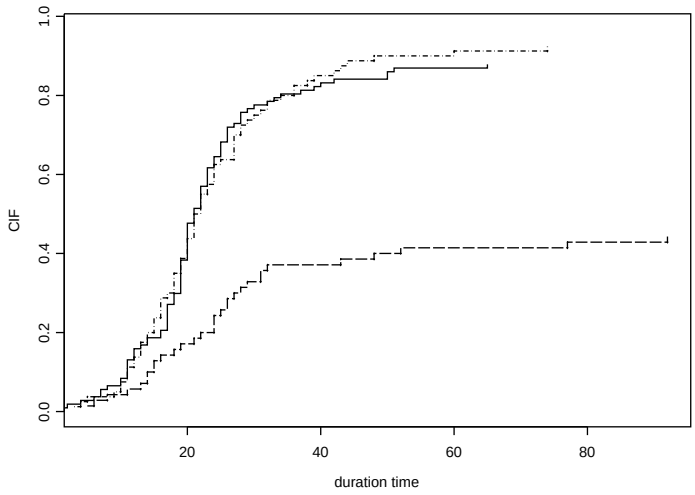

Figure S.1: The cumulative incidence function of cure for three age groups: $age < 30$ ($-\cdot-\cdot-$), $30 \leq age \leq 50$ (——) and $50 < age$ ($- - -$).

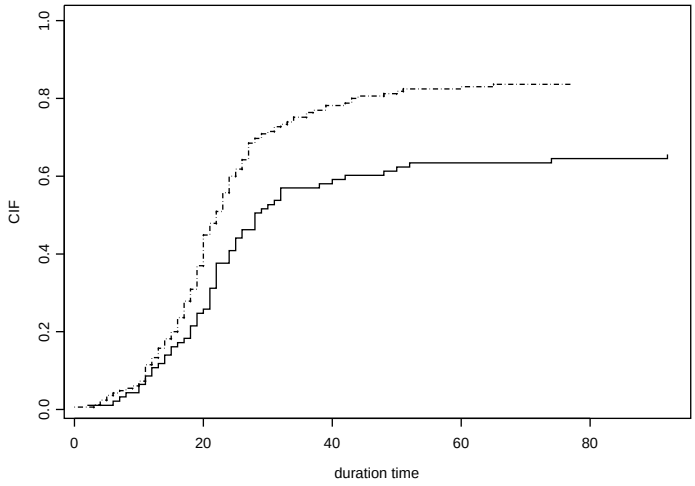

Figure S.2: The cumulative incidence function of cure for two groups of different gender: $female$ ($-\cdot-\cdot-$) and $male$ (——).

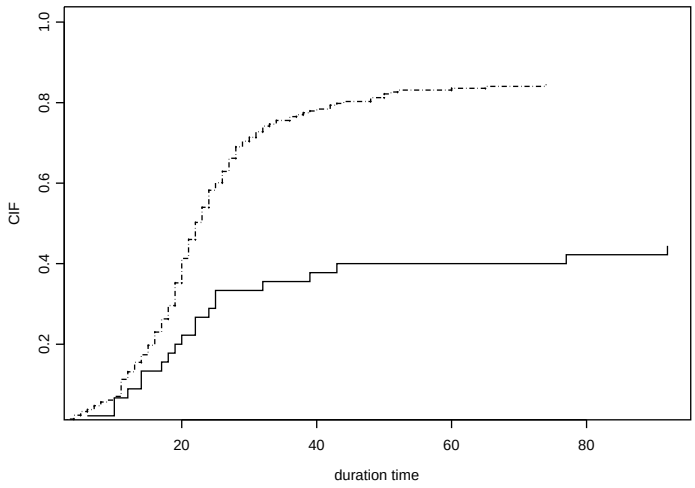

Figure S.3: The cumulative incidence function of cure for two groups with/without other diseases: "without disease" (−·−·−) and "with disease" (——).

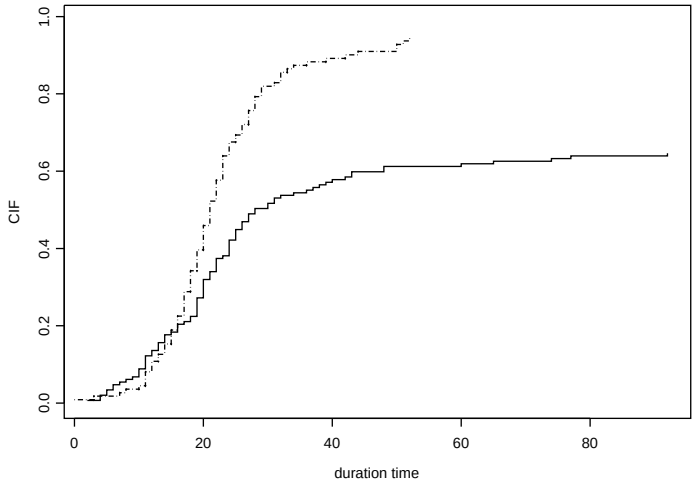

Figure S.4: The cumulative incidence function of cure for two groups according to whether the PCR test detected the SARS virus or not: "yes" (——) and "no" (−·−·−).

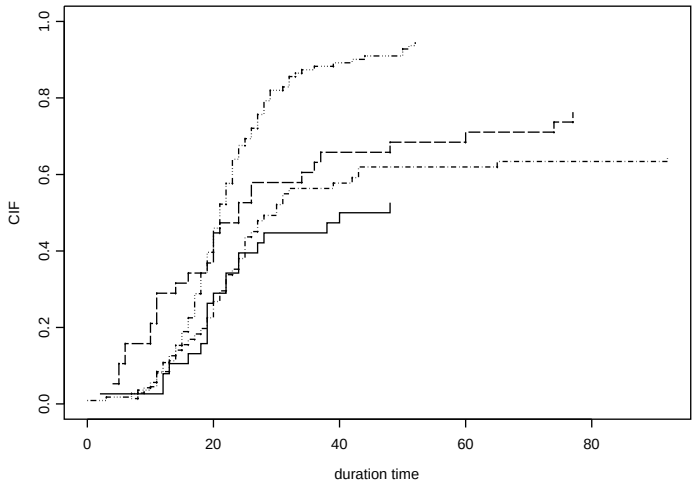

Figure S.5: The cumulative incidence function of cure for groups with different level of virus load (vl): vl = 0 (−···−), $0 <$ vl $< 10^3$ (−−−), $10^3 <$ vl $< 10^5$ (−·−·−) and $10^5 <$ vl (——).

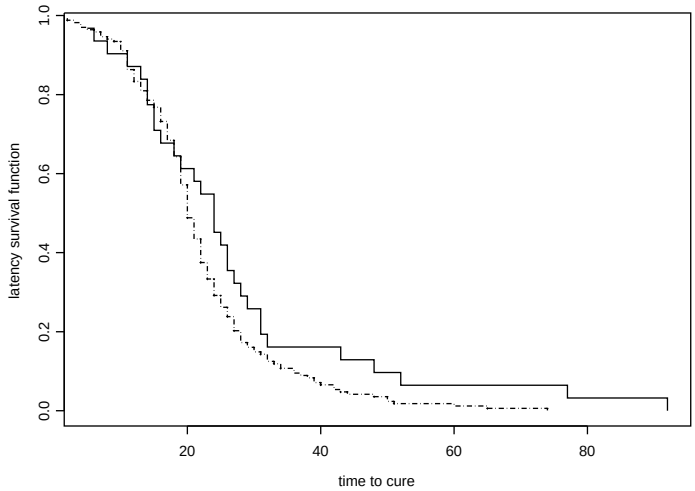

Figure S.6: The latency survival function of cure for two age groups: $age \leq 50$ (−·−·−) and $age > 50$ (——).

Covariate	$\hat{\beta}_{1,1}$	$\hat{\beta}_{1,2}$	$\hat{\beta}_{1,3}$	$\hat{\beta}_{1,4}$	$\hat{\beta}_{1,5}$	$\hat{\theta}$	p-value of testing time homogeneity
age							
> 50	-0.763	-1.510^a	-1.843^a	-1.926^a	-2.409^a	-1.914^a	0.006
	(0.439)	(0.340)	(0.307)	(0.308)	(0.341)	(0.276)	
gender							
male	-0.313	-0.707^a	-0.811^a	-0.825^a	-1.031^a	-0.755^a	0.476
	(0.361)	(0.273)	(0.268)	(0.276)	(0.305)	(0.234)	
disease							
with disease	-0.312	-1.093^a	-1.494^a	-1.725^a	-1.920^a	-1.568^a	0.007
	(0.474)	(0.384)	(0.349)	(0.350)	(0.355)	(0.312)	
PCR							
positive	0.172	-0.753^a	-1.328^a	-1.758^a	-2.260^a	-0.934^a	0.000
	(0.341)	(0.258)	(0.286)	(0.330)	(0.454)	(0.229)	
$\log_{10}$(viral load)	-0.064	-0.207^a	-0.293^a	-0.370^a	-0.461^a	-0.264^a	0.000
	(0.075)	(0.059)	(0.061)	(0.066)	(0.078)	(0.051)	

Table S.1: *Simple logistic regression analysis for assessing the effect of a covariate on the cumulative probability of cure by time τ_j for $j = 1, \ldots, 5$. The last column reports the p-value for testing equivalence of $\beta_{1,j}$ ($j = 1, \ldots, 5$). Items with p-values < 0.05 are marked by a.*

	Covariate	τ_2	τ_3	τ_4	τ_5
$\tilde{U}$	age > 50	-1.240 (0.370)^a	-1.478 (0.343)^a	-1.486 (0.355)^a	-2.066 (0.414)^a
	with disease	-0.400 (0.435)	-0.654 (0.417)	-0.904 (0.427)^a	-0.854 (0.470)^b
	$\log_{10}$(viral load)	-0.158 (0.061)^a	-0.251 (0.065)^a	-0.338 (0.072)^a	-0.464 (0.093)^a
U_{w^*}	age > 50	-1.237 (0.462)^a	-1.647 (0.430)^a	-1.566 (0.435)^a	-1.885 (0.464)^a
	with disease	-0.581 (0.521)	-0.633 (0.485)	-0.805 (0.460)^b	-0.837 (0.507)^b
	$\log_{10}$(viral load)	-0.181 (0.068)^a	-0.286 (0.073)^a	-0.396 (0.082)^a	-0.473 (0.090)^a
U_{I_1}	age > 50	-1.276 (0.435)^a	-1.646 (0.373)^a	-1.568 (0.414)^a	-1.925 (0.488)^a
	with disease	-0.500 (0.567)	-0.757 (0.462)	-1.110 (0.509)^a	-1.113 (0.530)^a
	$\log_{10}$(viral load)	-0.173 (0.078)^a	-0.251 (0.067)^a	-0.360 (0.090)^a	-0.398 (0.089)^a
U_{I_2}	age > 50	-1.277 (0.435)^a	-1.655 (0.376)^a	-1.569 (0.410)^a	-1.947 (0.479)^a
	with disease	-0.488 (0.556)	-0.767 (0.468)	-1.015 (0.502)^a	-1.050 (0.527)^a
	$\log_{10}$(viral load)	-0.179 (0.077)^a	-0.259 (0.064)^a	-0.361 (0.087)^a	-0.408 (0.091)^a
U_F	age > 50	-1.222 (0.459)^a	-1.364 (0.449)^a	-1.291 (0.521)^a	-1.714 (0.639)^a
	with disease	-0.399 (0.516)	-0.505 (0.517)	-0.869 (0.568)	-0.910 (0.675)
	$\log_{10}$(viral load)	-0.215 (0.075)^a	-0.327 (0.092)^a	-0.443 (0.123)^a	-0.497 (0.181)^a

Table S.2: *Multiple logistic regression analysis for SARS data subject to a single run of artificial censoring. In each cell, the estimated parameter and the estimated standard error (in parenthesis) are given. Items with p-value* < 0.05 *are marked by a and with p-value* < 0.1 *are marked by b.*

	Covariate	τ_2	τ_3	τ_4	τ_5
$\tilde{U}$	age > 50	-1.240 (0.370)	-1.478 (0.343)	-1.486 (0.355)	-2.066 (0.414)
	with disease	-0.400 (0.435)	-0.654 (0.417)	-0.904 (0.427)	-0.854 (0.470)
	$\log_{10}$(viral load)	-0.158 (0.061)	-0.251 (0.065)	-0.338 (0.072)	-0.464 (0.093)
U_{w^*}	age > 50	-1.279 (0.450)	-1.507 (0.417)	-1.530 (0.440)	-2.116 (0.506)
	with disease	-0.393 (0.487)	-0.672 (0.454)	-0.941 (0.478)	-1.170 (0.537)
	$\log_{10}$(viral load)	-0.158 (0.069)	-0.250 (0.075)	-0.339 (0.085)	-0.500 (0.102)
U_{I_1}	age > 50	-1.269 (0.486)	-1.502 (0.442)	-1.538 (0.462)	-2.115 (0.520)
	with disease	-0.402 (0.557)	-0.699 (0.509)	-0.993 (0.526)	-1.174 (0.585)
	$\log_{10}$(viral load)	-0.156 (0.071)	-0.247 (0.076)	-0.331 (0.088)	-0.442 (0.100)
U_{I_2}	age > 50	-1.270 (0.487)	-1.500 (0.438)	-1.540 (0.463)	-2.114 (0.515)
	with disease	-0.404 (0.550)	-0.685 (0.498)	-0.989 (0.521)	-1.173 (0.585)
	$\log_{10}$(viral load)	-0.157 (0.071)	-0.241 (0.076)	-0.333 (0.086)	-0.433 (0.099)
U_F	age > 50	-1.265 (0.469)	-1.483 (0.480)	-1.500 (0.553)	-2.365 (1.099)
	with disease	-0.410 (0.535)	-0.711 (0.550)	-0.990 (0.602)	-1.246 (0.955)
	$\log_{10}$(viral load)	-0.156 (0.075)	-0.250 (0.098)	-0.345 (0.126)	-0.587 (0.346)

Table S.3: *Multiple logistic regression analysis for SARS data by repeating artificial censoring 300 times. In each cell, the average of the parameter estimates and the average of the standard-error estimates are reported.*

Sample size	% censored	Estimators	Comparison criteria					
			BS	SD	ASD	CP (%)	MSE	RE
100	30	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.134	0.498	0.446	94.1	0.266	2.317
		$\hat{\boldsymbol{\beta}}_{I1}$	-0.070	0.455	0.474	94.0	0.212	2.909
		$\hat{\boldsymbol{\beta}}_{I2}$	-0.073	0.455	0.444	93.4	0.212	2.906
		$\hat{\boldsymbol{\beta}}_{F}$	0.134	0.774	0.841	93.4	0.617	1
100	40	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.138	0.541	0.479	94.3	0.312	3.742
		$\hat{\boldsymbol{\beta}}_{I1}$	-0.079	0.475	0.476	95.3	0.231	5.044
		$\hat{\boldsymbol{\beta}}_{I2}$	-0.077	0.474	0.481	95.7	0.231	5.062
		$\hat{\boldsymbol{\beta}}_{F}$	0.157	1.069	1.549	96.1	1.167	1
300	30	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.038	0.257	0.252	94.4	0.067	4.689
		$\hat{\boldsymbol{\beta}}_{I1}$	-0.025	0.249	0.247	95.7	0.063	5.035
		$\hat{\boldsymbol{\beta}}_{I2}$	-0.025	0.248	0.251	96.0	0.062	5.098
		$\hat{\boldsymbol{\beta}}_{F}$	0.097	0.553	0.566	93.8	0.316	1
300	40	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.048	0.280	0.270	95.1	0.081	6.940
		$\hat{\boldsymbol{\beta}}_{I1}$	-0.065	0.251	0.252	93.5	0.067	8.322
		$\hat{\boldsymbol{\beta}}_{I2}$	-0.064	0.253	0.258	94.0	0.068	8.212
		$\hat{\boldsymbol{\beta}}_{F}$	0.101	0.742	0.792	96.2	0.560	1

Table S.4: *Finite-sample comparison for four estimators of* $\beta_1 = 1.8$ *when the covariate* Z *follows the standard normal distribution. The label* BS *denotes the average bias,* SD *denotes the sample standard deviation,* ASD *denotes the average of the standard deviation estimates,* CP *denotes the empirical coverage probabilities of nominal* 95% *confidence intervals,* MSE *denotes the mean squared errors and* RE *denotes the relative efficiency defined as the ratio of the* MSE *of* $\hat{\boldsymbol{\beta}}_F$ *to that of the others.*

Sample size	% censored	Estimators	Comparison criteria					
			BS	SD	ASD	CP (%)	MSE	RE
100	30	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.041	0.291	0.261	93.3	0.086	3.330
		$\hat{\boldsymbol{\beta}}_{I1}$	0.037	0.259	0.264	97.0	0.068	4.211
		$\hat{\boldsymbol{\beta}}_{I2}$	0.032	0.258	0.270	97.2	0.068	4.259
		$\hat{\boldsymbol{\beta}}_{F}$	0.090	0.529	0.607	93.2	0.288	1
100	40	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.073	0.347	0.308	93.5	0.126	3.568
		$\hat{\boldsymbol{\beta}}_{I1}$	0.049	0.311	0.317	96.4	0.099	4.521
		$\hat{\boldsymbol{\beta}}_{I2}$	0.044	0.308	0.313	96.0	0.097	4.626
		$\hat{\boldsymbol{\beta}}_{F}$	0.127	0.657	0.891	93.3	0.448	1
300	30	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.019	0.143	0.146	95.1	0.021	4.303
		$\hat{\boldsymbol{\beta}}_{I1}$	-0.028	0.124	0.140	96.5	0.016	5.543
		$\hat{\boldsymbol{\beta}}_{I2}$	-0.030	0.122	0.131	96.1	0.016	5.664
		$\hat{\boldsymbol{\beta}}_{F}$	0.059	0.293	0.292	94.2	0.089	1
300	40	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	0.028	0.172	0.163	94.2	0.030	6.892
		$\hat{\boldsymbol{\beta}}_{I1}$	0.015	0.167	0.168	95.6	0.028	7.443
		$\hat{\boldsymbol{\beta}}_{I2}$	0.015	0.163	0.170	95.5	0.027	7.777
		$\hat{\boldsymbol{\beta}}_{F}$	0.099	0.447	0.495	94.6	0.209	1

Table S.5: *Finite-sample comparison for four estimators of* $\beta_1 = 0.86$ *when the covariate* Z *follows a uniform distribution. The label* BS *denotes the average bias,* SD *denotes the sample standard deviation,* ASD *denotes the average of the standard deviation estimates,* CP *denotes the empirical coverage probabilities of nominal* 95% *confidence intervals,* MSE *denotes the mean squared errors and* RE *denotes the relative efficiency defined as the ratio of the* MSE *of* $\hat{\boldsymbol{\beta}}_F$ *to that of the others.*

Sample size = 300, % censored = 30						
			Estimators of β_1			
type of covariate	$\hat{G}$	Criteria	$\hat{\boldsymbol{\beta}}_{\mathrm{w}^*}$	$\hat{\boldsymbol{\beta}}_{I1}$	$\hat{\boldsymbol{\beta}}_{I2}$	$\hat{\boldsymbol{\beta}}_{F}$
Binary	Kernel-type estimator	BS	-0.029	-0.015	-0.001	-0.037
		SD	0.319	0.316	0.315	0.323
		MSE	0.103	0.100	0.099	0.105
	Kaplan-Meier estimator	BS	0.092	0.083	0.081	-0.989
		SD	0.326	0.318	0.318	0.471
		MSE	0.114	0.108	0.108	1.199
Standard Normal	Kernel-type estimator	BS	-0.073	-0.066	-0.064	0.098
		SD	0.254	0.245	0.244	0.260
		MSE	0.070	0.064	0.064	0.077
	Kaplan-Meier estimator	BS	-0.191	-0.106	-0.103	2.775
		SD	0.258	0.248	0.246	1.058
		MSE	0.103	0.073	0.071	8.823

Table S.6: *Robustness analysis when the censoring variable depends on Z. The label* BS *denotes the average bias,* SD *denotes the sample standard deviation and* MSE *denotes the mean squared errors.*