Journal of Statistical Planning and Inference 140 (2010) 3425–3441

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

A class of mean residual life regression models with censored survival data

Liuquan Sun [a,*], Qiang Zhao [b]

[a] Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, PR China
[b] Department of Mathematics, Texas State University, San Marcos, TX 78666, USA

ARTICLE INFO

Article history:
Received 10 March 2008
Received in revised form
22 September 2009
Accepted 12 May 2010
Available online 28 May 2010

Keywords:
Censored data
Estimating equation
Failure time
Goodness-of-fit test
Mean residual life

ABSTRACT

When describing a failure time distribution, the mean residual life is sometimes preferred to the survival or hazard rate. Regression analysis making use of the mean residual life function has recently drawn a great deal of attention. In this paper, a class of mean residual life regression models are proposed for censored data, and estimation procedures and a goodness-of-fit test are developed. Both asymptotic and finite sample properties of the proposed estimators are established, and the proposed methods are applied to a cancer data set from a clinic trial.

1. Introduction

The mean residual life function (MRLF) is of interest in many fields such as actuarial studies, reliability and survival analysis, etc. For example, it may be more informative to tell a prostate cancer patient how long he can survive or live without disease recurrence, on average, given his current situation (which of course assumes the fact that he has "survived" or lived without the disease so far). In another example, a customer may be interested in knowing how much longer his or her computer can last, given that the computer has worked normally for, say, t years. For a nonnegative survival time T with finite expectation, the MRLF at time $t \geq 0$ is

$$m(t) = E(T - t | T > t).$$

To assess the effects of covariates on the mean residual life, the proportional mean residual life model by Oakes and Dasu (1990) may be used:

$$m(t|Z) = m_0(t)\exp(\beta_0' Z), \tag{1}$$

where $m(t|Z)$ is the MRLF corresponding to the p-vector covariate Z, $m_0(t)$ is some unknown baseline MRLF when Z=0, and β_0 is an unknown vector of regression parameters.

Previous work on the MRLF has focused on single-sample and two-sample cases (Oakes and Dasu, 1990). For regression analysis, Maguluri and Zhang (1994) used the underlying proportional hazards structure of the model to develop estimation procedures for β_0 in model (1), and Yuen et al. (2003) proposed a goodness-of-fit test for model (1), when there was no censoring involved. In the presence of censoring, Chen and Cheng (2005) used counting process theory to develop

* Corresponding author.
E-mail addresses: slq@amt.ac.cn (L. Sun), qiang.zhao@txstate.edu (Q. Zhao).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2010.05.012

semiparametric inference procedures for β_0 in model (1), and Chen et al. (2005) extended the estimation procedure of Maguluri and Zhang (1994) to censored survival data using inverse probability of censoring weighting techniques (Robins and Rotnitzky, 1992). Recently, Chen and Cheng (2006) and Chen (2007) proposed a new class of additive mean residual life model and discussed various estimation procedures with or without right censoring. However, other regression forms may be more natural or descriptive in some applications.

In this paper, we consider a more general class of mean residual life regression models given by

$$m(t|Z) = m_0(t)g(\beta_0'Z), \tag{2}$$

where $g(t) \geq 0$ is a pre-specified link function and assumed to be continuous almost everywhere and twice differentiable. Examples of potential link function include $g(x)=1+x$, $g(x)=e^x$ and $g(x)=\log(1+e^x)$. Selection of an appropriate link function may be based on prior data or the resulting interpretation of the regression parameters.

In the next section, we will first discuss the situation where the censoring time is independent of T and Z, and a general inference procedure based on estimating functions is proposed. The procedure can be easily implemented numerically and the asymptotic properties of the proposed estimates of regression parameters are established. Section 3 generalizes the methods to the situation where the censoring time may depend on Z through the proportional hazards model. In Section 4, we develop test procedures for checking the adequacy of model (2) under both independent and covariate-dependent censoring scenarios based on an appropriate stochastic process which is asymptotically Gaussian. Section 5 reports some results from simulation studies conducted for evaluating the proposed methods. In Section 6, we apply the methodology to a data set from a cancer clinic trial and some concluding remarks are given in Section 7.

2. Inference with independent censoring times

In this section, let C be the potential censoring time, and assume that C is independent of T and Z. To avoid lengthy technical discussion of the tail behavior of the limiting distributions, we further assume that $\Pr\{C \geq \tau\} > 0$, where $0 < \tau = \inf\{t : \Pr(T \geq t) = 0\} < \infty$. Let $\{T_i, C_i, Z_i;\ i=1,\ldots,n\}$ be independent replicates of $\{T, C, Z\}$ and suppose that we observe $\{X_i, \delta_i, Z_i; i = 1, \ldots, n\}$, where $X_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$. Here $I(\cdot)$ is the indicator function. Define

$$M_i(t) = N_i(t) - \int_0^t Y_i(u)\, d\Lambda(u|Z_i), \quad i = 1, \ldots, n, \tag{3}$$

where $N_i(t) = I(X_i \leq t, \delta_i = 1)$, $Y_i(t) = I(X_i \geq t)$, and $\Lambda(t|Z_i)$ is the cumulative hazard function of T_i given Z_i. It is well known that $M_i(t)$, $i=1,\ldots,n$, are zero-mean martingale with respect to the $\sigma-$filtration $\sigma\{N_i(u), Y_i(u+), Z_i : 0 \leq u \leq t, i = 1, \ldots, n\}$.

Note that the survival function of T given Z is

$$S(t|Z) = \frac{m(0|Z)}{m(t|Z)} \exp\left\{ -\int_0^t \frac{du}{m(u|Z)} \right\}.$$

Then under model (2), we have

$$m_0(t)\, d\Lambda(t|Z_i) = g(\beta_0'Z_i)^{-1}\, dt + dm_0(t). \tag{4}$$

Thus, in view of (3) and (4), for given β, a reasonable estimator for $m_0(t)$ is the solution to

$$\sum_{i=1}^n [m_0(t)\, dN_i(t) - Y_i(t)\{g(\beta'Z_i)^{-1}\, dt + dm_0(t)\}] = 0, \quad 0 \leq t \leq \tau. \tag{5}$$

Denote this estimator by $\hat{m}_{a0}(t;\beta)$. Straightforward algebra on (5) leads to

$$\hat{m}_{a0}(t;\beta) = \Phi_n(t)^{-1} \int_t^\tau \frac{\Phi_n(u)\sum_{i=1}^n Y_i(u) g(\beta'Z_i)^{-1}}{\sum_{i=1}^n Y_i(u)}\, du, \tag{6}$$

where $\Phi_n(t) = \exp\{-\int_0^t \sum_{i=1}^n dN_i(u)/\sum_{i=1}^n Y_i(u)\}$, which is the Nelson–Aalen estimator of the survival function for the pooled observations with independent censoring times. To estimate β_0, using the generalized estimating equation methods (Liang and Zeger, 1986; Cai and Schaubel, 2004; Chen and Cheng, 2005), we propose the following class of estimating equations for β_0:

$$\sum_{i=1}^n \int_0^\tau \frac{g^{(1)}(\beta'Z_i)}{g(\beta'Z_i)} Z_i [\hat{m}_{a0}(t;\beta)\, dN_i(t) - Y_i(t)\{g(\beta'Z_i)^{-1}\, dt + d\hat{m}_{a0}(t;\beta)\}] = 0,$$

where $g^{(1)}(x)=dg(x)/dx$. In view of (5), the above estimating equations are equivalent to

$$U_a(\beta) = n^{-1} \sum_{i=1}^n \int_0^\tau \{h(\beta'Z_i)Z_i - \overline{Z}_a(t;\beta)\}[\hat{m}_{a0}(t;\beta)\, dN_i(t) - Y_i(t) g(\beta'Z_i)^{-1}\, dt] = 0, \tag{7}$$

where $h(x)=g^{(1)}(x)/g(x)$, and

$$\overline{Z}_a(t;\beta) = \frac{\sum_{i=1}^n Y_i(t) h(\beta'Z_i) Z_i}{\sum_{i=1}^n Y_i(t)}.$$

Let $\hat{\beta}_a$ denote the solution to $U_a(\beta)=0$ and $\hat{m}_{a0}(t) \equiv \hat{m}_{a0}(t;\hat{\beta}_a)$, the corresponding estimator of the unknown baseline MRLF $m_0(t)$. Following the arguments of Chen and Cheng (2005) and Lin et al. (2001), we can check that both $\hat{\beta}_a$ and $\hat{m}_{a0}(t)$ always exist and are unique and consistent. To study the asymptotic distribution of $\hat{\beta}_a$, we show in Appendix A.1 that $n^{1/2}U_a(\beta_0)$ is asymptotically normal with mean zero and covariance matrix that can be consistently estimated by $\hat{\Sigma}_a$, where

$$\hat{\Sigma}_a = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\hat{\beta}_a' Z_i)Z_i - \hat{\mu}(t)\}^{\otimes 2}\hat{m}_{a0}(t)^2\, dN_i(t),$$

$$\hat{\mu}(t) = \overline{Z}_a(t;\hat{\beta}_a) + \frac{\Phi_n(t)}{\pi_n(t)}\int_0^t n^{-1}\sum_{i=1}^{n}[h(\hat{\beta}_a' Z_i)Z_i - \overline{Z}_a(u;\hat{\beta}_a)]\frac{dN_i(u)}{\Phi_n(u)},$$

and $\pi_n(t)=n^{-1}\sum_{i=1}^n Y_i(t)$. Here for a vector v, $v^{\otimes 0}=1$, $v^{\otimes 1}=v$ and $v^{\otimes 2}=vv'$. Then it follows that $n^{1/2}(\hat{\beta}_a-\beta_0)$ is asymptotically normal with zero mean and covariance matrix that can be consistently estimated by $\hat{A}^{-1}\hat{\Sigma}_a\hat{A}^{-1}$, where

$$\hat{A} = n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\hat{\beta}_a' Z_i)Z_i - \overline{Z}_a(t;\hat{\beta}_a)\}^{\otimes 2}Y_i(t)g(\hat{\beta}_a' Z_i)^{-1}\, dt.$$

We also show in Appendix A.2 that $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$ $(0 \le t \le \tau)$ converges weakly to a zero-mean Gaussian process whose covariance function at (s,t) can be estimated consistently by $\hat{\Gamma}_a(s,t)=n^{-1}\sum_{i=1}^n \hat{\varphi}_i(s)\hat{\varphi}_i(t)$, where

$$\hat{\varphi}_i(t) = -\frac{1}{\Phi_n(t)}\int_t^{\tau}\frac{\Phi_n(u)}{\pi_n(u)}[\hat{m}_{a0}(u)\, dN_i(u) - Y_i(u)\{g(\hat{\beta}_a' Z_i)^{-1}\, du + d\hat{m}_{a0}(u)\}]$$
$$+\hat{m}_{a0}(t)\overline{Z}_a(t;\hat{\beta}_a)'\hat{A}^{-1}\int_0^{\tau}\{h(\hat{\beta}_a' Z_i)Z_i - \hat{\mu}(u)\}[\hat{m}_{a0}(u)\, dN_i(u) - Y_i(u)\{g(\hat{\beta}_a' Z_i)^{-1}\, du + d\hat{m}_{a0}(u)\}].$$

The asymptotic normality for $\hat{m}_{a0}(t)$, together with the consistent variance estimator $\hat{\Gamma}_a(t,t)$, enables us to construct pointwise confidence intervals for $m_0(t)$. Since $m_0(t)$ is nonnegative, one may want to use the log transformation for the construction of its confidence intervals. To construct simultaneous confidence bands for $m_0(t)$ over a time interval of interest $[t_1,t_2]$ $(0 < t_1 < t_2 \le \tau)$, we need to evaluate the distribution of the supremum of a related process over $[t_1,t_2]$. It is not possible to evaluate such distributions analytically because the limiting process of $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$ does not have an independent increments structure. To handle this problem, we use a resampling scheme to approximate the distribution of $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$. Define

$$\hat{W}_a(t) = n^{-1/2}\sum_{i=1}^{n}\hat{\varphi}_i(t)\Omega_i,$$

where $(\Omega_1,\ldots,\Omega_n)$ are independent standard normal variables which are independent of the data $\{X_i,\delta_i,Z_i; i=1,\ldots,n\}$. According to the arguments of Lin et al. (2000), the distribution of the process $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$ can be approximated by that of the zero-mean Gaussian process $\hat{W}_a(t)$. To approximate the distributions of $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$, we obtain a large number of realizations from $\hat{W}_a(t)$ by repeatedly generating the normal random sample $(\Omega_1,\ldots,\Omega_n)$ while fixing the data $\{X_i,\delta_i,Z_i; i=1,\ldots,n\}$ at their observed values. Using this simulation method, we may determine an approximate $1-\alpha$ simultaneous confidence bands for $m_0(t)$ over a time interval of interest $[t_1,t_2]$.

3. Inference with covariate-dependent censoring times

Now we consider the situation where T, C and Z may depend on each other, but given Z, we assume that T is independent of C. Also we assume that the hazard function of C given Z has the form

$$\lambda_c(t|Z) = \lambda_0(t)\exp\{\gamma_0' Z\}, \tag{8}$$

where $\lambda_0(t)$ is an unspecified baseline hazard function and γ_0 is a vector of unknown regression parameters. Note that the model from Section 2 is the model of Section 3 with $\gamma_0=0$. Of course, γ_0 is usually unknown. A natural estimate of γ_0, which is efficient under model (8), is given by the maximum partial likelihood estimate defined as the solution to (Cox, 1972)

$$U_r(\gamma) = \sum_{i=1}^{n}\int_0^{\tau}\{Z_i - \overline{Z}_r(t;\gamma)\}\, dN_i^c(t) = 0, \tag{9}$$

where $N_i^c(t)=I(X_i \le t, \delta_i=0)$, and $\overline{Z}_r(t;\gamma)=S^{(1)}(t;\gamma)/S^{(0)}(t;\gamma)$, $S^{(k)}(t;\gamma)=\sum_{i=1}^n Y_i(t)Z_i^{\otimes k}\exp\{\gamma' Z_i\}$ for k=0,1,2. Let $\hat{\gamma}$ denote the estimator given by $U_r(\gamma)=0$, and $\hat{\Lambda}_0(t)$ be the Breslow estimator of $\Lambda_0(t)=\int_0^t \lambda_0(u)\, du$, where

$$\hat{\Lambda}_0(t) = \sum_{i=1}^{n}\int_0^{t}\frac{dN_i^c(u)}{\sum_{i=1}^n Y_i(u)\exp\{\hat{\gamma}' Z_i\}}.$$

Consider a hypothetical equilibrium renewal process formed by renewals following the same survival distribution as $S(t|Z)$. The forward recurrence time V is defined as the time from a fixed time to the next immediate renewal. Then under model (2), it follows from Cox (1962) that its hazard function is

$$\lambda_v(t|Z) = m(t|Z)^{-1} = m_0(t)^{-1} g(\beta_0' Z)^{-1},$$

which is a proportional hazards model. When there is no censoring, the following partial score equation can be used to estimate β_0 (Prentice and Self, 1983; Cai and Schaubel, 2004),

$$\hat{E}\{h(\beta' Z)Z\} - \int_0^\tau \frac{\hat{E}[h(\beta' Z)Zg(\beta' Z)^{-1} I(V \geq t)]}{\hat{E}[g(\beta' Z)^{-1} I(V \geq t)]}\, d\hat{F}_v(t) = 0, \tag{10}$$

where $\hat{E}$ and $\hat{F}_v(t)$ are their empirical estimates of the expectation E and $F_v(t)$, respectively. Here $F_v(t)$ is the distribution function of V. However, this equality is only theoretical, since we cannot observe V. To use the sample of T's in (10), following the arguments of Maguluri and Zhang (1994) and Chen et al. (2005), we have that for any function $w(Z)$,

$$E\{w(Z)I(V \geq t)\} = m_0(0)^{-1} E\{w(Z) g(\beta_0' Z)^{-1} (T-t)^+\},$$

where $(T-t)^+$ denotes $(T-t)I(T \geq t)$. As a result,

$$dF_v(t) = \frac{E\{g(\beta_0' Z)^{-1} I(T > t)\}}{E\{g(\beta_0' Z)^{-1} T\}}\, dt.$$

Replacing the respective terms in (10), we obtain the following estimating equation for β based on T's:

$$n^{-1} \sum_{i=1}^n h(\beta' Z_i) Z_i - \int_0^\tau \frac{\sum_{i=1}^n h(\beta' Z_i) Z_i g(\beta' Z_i)^{-2} (T_i - t)^+}{\sum_{i=1}^n g(\beta' Z_i)^{-2} (T_i - t)^+} \frac{\sum_{i=1}^n g(\beta' Z_i)^{-1} I(T_i > t)}{\sum_{i=1}^n g(\beta' Z_i)^{-1} T_i}\, dt = 0. \tag{11}$$

Let $G_i(t;\gamma_0,\Lambda_0)$ be the censoring survival distribution of C_i given Z_i under model (8), that is, $G_i(t;\gamma_0,\Lambda_0) = \exp\{-\Lambda_0(t)\exp(\gamma_0' Z_i)\}$. Then for any well-defined function of v,

$$E\left\{\frac{v(X_i,Z_i,t)\delta_i}{G_i(X_i;\gamma_0,\Lambda_0)}\right\} = E\left[E\left\{\frac{v(T_i,Z_i,t)\delta_i}{G_i(T_i;\gamma_0,\Lambda_0)}\middle| Z_i\right\}\right] = E\{v(T_i,Z_i,t)\}. \tag{12}$$

In view of (11) and (12), using inverse probability of censoring weighting techniques (Robins and Rotnitzky, 1992), we propose the following class of estimating equations for β_0 when the censoring time C_i may depend on Z_i under model (8):

$$U_b(\beta) = n^{-1} \sum_{i=1}^n \frac{\delta_i}{G_i(X_i;\hat{\gamma},\hat{\Lambda}_0)} \{h(\beta' Z_i) Z_i - \overline{Z}_i(\beta,\hat{\gamma},\hat{\Lambda}_0)\} = 0, \tag{13}$$

where

$$\overline{Z}_i(\beta,\gamma,\Lambda) = \int_0^\tau h(\beta' Z_i) Z_i g(\beta' Z_i)^{-2} (X_i - t)^+ L_n(t;\beta,\gamma,\Lambda)\, dt,$$

$$L_n(t;\beta,\gamma,\Lambda) = \frac{L_{1n}(t;\beta,\gamma,\Lambda)}{L_{2n}(t;\beta,\gamma,\Lambda) L_{3n}(t;\beta,\gamma,\Lambda)},$$

and $L_{kn}(t;\beta,\gamma,\Lambda) = n^{-1} \sum_{i=1}^n V_{ki}(t;\beta) G_i(X_i;\gamma,\Lambda)^{-1}$, k=1,2,3. Here

$$V_{1i}(t;\beta) = g(\beta' Z_i)^{-1} I(X_i > t)\delta_i, \quad V_{2i}(t;\beta) = g(\beta' Z_i)^{-2} (X_i - t)^+ \delta_i,$$

and $V_{3i}(t;\beta) = g(\beta' Z_i)^{-1} X_i \delta_i$.

Let $\hat{\beta}_b$ denote the solution to $U_b(\beta) = 0$. It can be shown in Appendix A.3 that $\hat{\beta}_b$ is consistent and unique in a neighborhood of β_0. To study the asymptotic distribution of $\hat{\beta}_b$, we first show that $n^{1/2} U_b(\beta_0)$ is asymptotically normal with zero mean and covariance matrix that can be consistently estimated by $\hat{\Sigma}_b$, where

$$\hat{\Sigma}_b = n^{-1} \sum_{i=1}^n \left[\hat{\xi}_i + \int_0^\tau \frac{R_n(t)}{S^{(0)}(t;\hat{\gamma})}\, d\hat{M}_i^c(t) + B_n D_n^{-1} \int_0^\tau \{Z_i - \overline{Z}_r(t;\hat{\gamma})\}\, d\hat{M}_i^c(t)\right]^{\otimes 2},$$

$$\hat{\xi}_i = \{h(\hat{\beta}_b' Z_i) Z_i - \overline{Z}_i(\hat{\beta}_b,\hat{\gamma},\hat{\Lambda}_0)\}\delta_i G_i(X_i;\hat{\gamma},\hat{\Lambda}_0)^{-1} - \int_0^\tau Q_n(t)\left[\frac{\hat{\xi}_{1i}(t)}{\hat{L}_{2n}(t)\hat{L}_{3n}(t)} - \frac{\hat{L}_{1n}(t)\hat{\xi}_{2i}(t)}{\hat{L}_{2n}(t)^2 \hat{L}_{3n}(t)} - \frac{\hat{L}_{1n}(t)\hat{\xi}_{3i}(t)}{\hat{L}_{2n}(t)\hat{L}_{3n}(t)^2}\right] dt,$$

$$R_n(t) = n^{-1} \sum_{i=1}^n \{h(\hat{\beta}_b' Z_i) Z_i - \overline{Z}_i(\hat{\beta}_b,\hat{\gamma},\hat{\Lambda}_0)\}\exp\{\hat{\gamma}' Z_i\}\delta_i G_i(X_i;\hat{\gamma},\hat{\Lambda}_0)^{-1} Y_i(t) - \int_0^\tau Q_n(t)\left[\frac{R_{1n}(t,u)}{\hat{L}_{2n}(t)\hat{L}_{3n}(t)} - \frac{\hat{L}_{1n}(t)R_{2n}(t,u)}{\hat{L}_{2n}(t)^2 \hat{L}_{3n}(t)} - \frac{\hat{L}_{1n}(t)R_{3n}(t,u)}{\hat{L}_{2n}(t)\hat{L}_{3n}(t)^2}\right] du,$$

$$B_n = n^{-1} \sum_{i=1}^n \{h(\hat{\beta}_b' Z_i) Z_i - \overline{Z}_i(\hat{\beta}_b,\hat{\gamma},\hat{\Lambda}_0)\}\hat{\Lambda}_0(X_i)\exp\{\hat{\gamma}' Z_i\} Z_i' \delta_i G_i(X_i;\hat{\gamma},\hat{\Lambda}_0)^{-1} - n^{-1} \sum_{i=1}^n \int_0^\tau \{h(\hat{\beta}_b' Z_i) Z_i - \overline{Z}_i(\hat{\beta}_b,\hat{\gamma},\hat{\Lambda}_0)\}$$

$$\times \exp\{\hat\gamma' Z_i\}\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1} Y_i(t)\overline{Z}_r(t;\hat\gamma)\}'\,d\hat\Lambda_0(t) - \int_0^\tau Q_n(t)\left[\frac{P_{1n}(t)}{\hat L_{2n}(t)\hat L_{3n}(t)} - \frac{\hat L_{1n}(t)P_{2n}(t)}{\hat L_{2n}(t)^2\hat L_{3n}(t)} - \frac{\hat L_{1n}(t)P_{3n}(t)}{\hat L_{2n}(t)\hat L_{3n}(t)^2}\right]dt,$$

$$Q_n(t) = n^{-1}\sum_{i=1}^n h(\hat\beta_b' Z_i)Z_i g(\hat\beta_b' Z_i)^{-2}(X_i - t)^+\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1},$$

$$\hat\xi_{ki}(t) = V_{ki}(t;\hat\beta_b)G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1} - \hat L_{kn}(t), \quad k=1,2,3,$$

$$R_{kn}(t,u) = n^{-1}\sum_{i=1}^n V_{ki}(t;\hat\beta_b)G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}\exp\{\hat\gamma' Z_i\}Y_i(u),$$

$$P_{kn}(t) = n^{-1}\sum_{i=1}^n V_{ki}(t;\hat\beta_b)G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}\hat\Lambda_0(X_i)\exp\{\hat\gamma' Z_i\}Z_i' - \int_0^\tau R_{kn}(t,u)\overline{Z}_r(u;\hat\gamma)'\,d\hat\Lambda_0(u),$$

$$\hat M_i^c(t) = N_i^c(t) - \int_0^t Y_i(u)\exp\{\hat\gamma' Z_i\}\,d\hat\Lambda_0(u), \tag{14}$$

$\hat L_{kn}(t) = L_{kn}(t;\hat\beta_b,\hat\gamma,\hat\Lambda_0)$, and $D_n = -\partial U_r(\hat\gamma)/\partial\gamma'$. Then it follows that $n^{1/2}(\hat\beta_b - \beta_0)$ is asymptotically normal with zero mean and covariance matrix that can be consistently estimated by

$$\left\{\frac{\partial U_b(\hat\beta_b)}{\partial\beta'}\right\}^{-1}\hat\Sigma_b\left\{\frac{\partial U_b(\hat\beta_b)'}{\partial\beta}\right\}^{-1}.$$

To estimate the baseline mean residual life $m_0(t)$, define

$$M_i^*(t) = \frac{\delta_i I(X_i > t)}{G(X_i;\gamma_0,\Lambda_0)}[(X_i - t) - m_0(t)g(\beta_0' Z_i)], \quad i=1,\ldots,n.$$

Under models (2) and (8), $M_i^*(t)$ are zero-mean stochastic processes. Thus, for given β, a reasonable estimator for $m_0(t)$ is the solution to

$$\sum_{i=1}^n \frac{\delta_i I(X_i > t)}{G_i(X_i;\hat\gamma,\hat\Lambda_0)}[(X_i - t) - m_0(t)g(\beta' Z_i)] = 0, \quad 0 \le t \le \tau.$$

Denote this estimator by $\hat m_{b0}(t;\beta)$, which can be expressed as

$$\hat m_{b0}(t;\beta) = \frac{\sum_{i=1}^n (X_i - t)^+\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}}{\sum_{i=1}^n I(X_i > t)g(\beta' Z_i)\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}}. \tag{15}$$

Let $\hat m_{b0}(t) \equiv \hat m_{b0}(t;\hat\beta_b)$ be the corresponding estimator of the unknown baseline mean residual life $m_0(t)$ under models (2) and (8). Following the arguments of Appendices A.2 and A.3, we can check that $\hat m_{b0}(t)$ is consistent, and that $n^{1/2}\{\hat m_{b0}(t) - m_0(t)\}$ $(0 \le t \le \tau)$ converges weakly to a zero-mean Gaussian process whose covariance function at (s,t) can be estimated consistently by $\hat\Gamma_b(s,t) = n^{-1}\sum_{i=1}^n \hat\psi_i(s)\hat\psi_i(t)$, where

$$\begin{aligned}
\hat\psi_i(t) &= \hat m_{b0}(t)\overline{Z}_b(t;\hat\beta_b)'\left\{\frac{\partial U_b(\hat\beta_b)}{\partial\beta'}\right\}^{-1}\left[\hat\xi_i + B_n D_n^{-1}\int_0^\tau\{Z_i - \overline{Z}_r(u;\hat\beta_b)\}\,d\hat M_i^c(u) + \int_0^\tau \frac{R_n(u)}{S^{(0)}(u;\hat\gamma)}\,d\hat M_i^c(u)\right] \\
&\quad + \Psi_n(t;\hat\beta_b)^{-1}\left[\hat M_i^*(t) + \int_0^\tau \frac{r_n(t,u)}{S^{(0)}(u;\hat\gamma)}\,d\hat M_i^c(u) + B_n^*(t)D_n^{-1}\int_0^\tau\{Z_i - \overline{Z}_r(u;\hat\beta_b)\}\,d\hat M_i^c(u)\right], \\
\overline{Z}_b(t;\beta) &= \frac{\sum_{i=1}^n I(X_i > t)h(\beta' Z_i)Z_i g(\beta' Z_i)\delta_i G(X_i;\hat\gamma,\hat\Lambda_0)^{-1}}{n\Psi_n(t;\beta)}, \\
\Psi_n(t;\beta) &= n^{-1}\sum_{i=1}^n I(X_i > t)g(\beta' Z_i)\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}, \\
\hat M_i^*(t) &= I(X_i > t)[(X_i - t) - \hat m_{b0}(t)g(\hat\beta_b' Z_i)]\delta_i G_i(X_i;\hat\gamma,\hat\Lambda_0)^{-1}, \\
r_n(t,u) &= n^{-1}\sum_{i=1}^n \hat M_i^*(t)\exp\{\hat\gamma' Z_i\}Y_i(u), \\
B_n^*(t) &= n^{-1}\sum_{i=1}^n \hat M_i^*(t)\hat\Lambda_0(X_i)\exp\{\hat\gamma' Z_i\}Z_i' - \int_0^t r_n(t,u)\overline{Z}_r(u;\hat\gamma)'\,d\hat\Lambda_0(u),
\end{aligned}$$

and $\hat\xi_i$, B_n, $R_n(u)$ and D_n are defined in (14).

Note that the limiting process of $n^{1/2}\{\hat{m}_{b0}(t)-m_0(t)\}$ is quite complicated, and its properties are difficult to obtain analytically. As discussed in Section 2, we can show that the distribution of the process $n^{1/2}\{\hat{m}_{b0}(t)-m_0(t)\}$ can be approximated by that of the zero-mean Gaussian process $\hat{W}_b(t)$, where

$$\hat{W}_b(t)=n^{-1/2}\sum_{i=1}^{n}\hat{\psi}_i(t)\Omega_i,$$

and $(\Omega_1,\ldots,\Omega_n)$ are independent standard normal variables which are independent of the data $\{X_i,\delta_i,Z_i;i=1,\ldots,n\}$.

4. Model checking techniques

In this section, we develop testing procedures to check the adequacy of model (2) for both independent and covariate-dependent cases. Beginning with the independent case where censoring time C is independent of T and Z, let $G(t)$ be the survival function of C, and $\hat{G}(t)$ be the Kaplan–Meier estimate of $G(t)$ based on $\{X_i,1-\delta_i,i=1,\ldots,n\}$, where

$$\hat{G}(t)=\prod_{s\le t}\left\{1-\frac{\sum_{i=1}^{n}dN_i^c(s)}{\sum_{i=1}^{n}Y_i(s)}\right\}.$$

Define $H_1(t,z)=P\{X_i\le t,Z_i\le z,\delta_i=1\}$, and $H(t,z)=P\{X_i\ge t,Z_i\le z\}$, where the notation " $Z_i\le z$" means that each component of Z_i is less than or equal to the corresponding component of z. After some algebraic manipulation, model (2) leads to

$$m_0(t)=\frac{1}{H(t,z)}\int_t^{\tau}\int_0^{z}\frac{(s-t)G(t)}{g(\beta_0'w)G(s)}H_1(ds,dw), \tag{16}$$

where $\int_0^z$ stands for $\int_0^{z_1}\cdots\int_0^{z_p}$. Let us denote the right-hand side of (16) by $V(t,z)$. Note that the left-hand side is independent of the variable z. As a measure of fit for model (2), we estimate $V(t,z)$ by $V_n(t,z)$ and obtain the process

$$\theta_n(t,z)=n^{1/2}\{V_n(t,z)-V_n(t,z_u)\}, \tag{17}$$

where z_u is the vector of upper bounds for Z,

$$V_n(t,z)=\frac{1}{H_n(t,z)}\left\{\int_t^{\tau}\int_0^{z}\frac{(s-t)\hat{G}(t)}{g(\hat{\beta}_a'w)\hat{G}(s)}H_{1n}(ds,dw)\right\},$$

and H_n and H_{1n} are the empirical counterparts of H and H_1, respectively. That is, $H_n(t,z)=n^{-1}\sum_{i=1}^{n}I(X_i\ge t,Z_i\le z)$ and $H_{1n}(t,z)=n^{-1}\sum_{i=1}^{n}I(X_i\le t,Z_i\le z,\delta_i=1)$. Under model (2), the process $\theta_n(t,z)$ equals $\phi_n(t,z)-\phi_n(t,z_u)$, where $\phi_n(t,z)=n^{1/2}\{V_n(t,z)-V(t,z)\}$ is the standardized mean residual life process. Hence, based on (17), a Kolmogorov–Smirnov (KS) type test statistic $\mathcal{F}_n^{(1)}$ may be used to check the adequacy of model (2), where

$$\mathcal{F}_n^{(1)}=\sup_{t,z}|\theta_n(t,z)|.$$

Under model (2), we show in Appendix A.4 that $\theta_n(t,z)$ converges to a zero-mean Gaussian process $W(t,z)$ whose covariance function at (t,z) and (t^*,z^*) can be estimated consistently by $\hat{\sigma}(t,z;t^*,z^*)=n^{-1}\sum_{i=1}^{n}\hat{\eta}_i(t,z)\hat{\eta}_i(t^*,z^*)$, where $\hat{\eta}_i(t,z)=\hat{\rho}_i(t,z)-\hat{\rho}_i(t,z_u)$,

$$\begin{aligned}\hat{\rho}_i(t,z)=&\frac{\hat{G}(t)}{H_n(t,z)}\int_t^{\tau}\left[\int_u^{\tau}\int_0^{z}\frac{s-t}{\hat{G}(s)g(\hat{\beta}_a'w)}H_{1n}(ds,dw)\right]\frac{d\tilde{M}_i^c(u)}{\pi_n(u)}+\frac{\delta_i(X_i-t)\hat{G}(t)}{H_n(t,z)\hat{G}(X_i)g(\hat{\beta}_a'Z_i)}I(X_i\ge t,Z_i\le z)-\frac{V_n(t,z)}{H_n(t,z)}I(X_i\ge t,Z_i\le z)\\&+\frac{\hat{G}(t)}{H_n(t,z)}\int_t^{\tau}\int_0^{z}\frac{(s-t)h(\hat{\beta}_a'w)w'}{\hat{G}(s)g(\hat{\beta}_a'w)}H_{1n}(ds,dw)\hat{A}^{-1}\int_0^{\tau}\{h(\hat{\beta}_a'Z_i)Z_i-\hat{\mu}(u)\}[\hat{m}_{a0}(u)\,dN_i(u)-Y_i(u)\{g(\hat{\beta}_a'Z_i)^{-1}\,du+d\hat{m}_{a0}(u)\}],\end{aligned}$$

$$d\hat{m}_{a0}(u)=\frac{\sum_{j=1}^{n}[\hat{m}_{a0}(u)dN_j(u)-Y_j(u)g(\hat{\beta}'Z_j)^{-1}du]}{\sum_{j=1}^{n}Y_j(u)},$$

$$d\tilde{M}_i^c(u)=dN_i^c(u)-Y_i(u)\,d\Lambda_n^c(u), \tag{18}$$

and $d\Lambda_n^c(u)=n^{-1}\sum_{i=1}^{n}dN_i^c(u)/\pi_n(u)$. Consequently, $\mathcal{F}_n^{(1)}$ converges in distribution to $\mathcal{F}$, where

$$\mathcal{F}=\sup_{t,z}|W(t,z)|.$$

Obviously, the complicated structure of the covariance function (18) does not allow for an analytic treatment of the involved distributions. As discussed in Sections 2 and 3, we can show that the distribution of the process $W(t,z)$ can be approximated by that of the zero-mean Gaussian process $\tilde{W}(t,z)$, where

$$\tilde{W}(t,z)=n^{-1/2}\sum_{i=1}^{n}\hat{\eta}_i(t,z)\Omega_i,$$

and $(\Omega_1,\ldots,\Omega_n)$ are independent standard normal variables which are independent of the data $\{X_i,\delta_i,Z_i; i=1,\ldots,n\}$. Thus, the distributions of $\mathcal{F}$ can be approximated by $\tilde{\mathcal{F}}$, where

$$\tilde{\mathcal{F}}=\sup_{t,z}|\tilde{W}(t,z)|.$$

To approximate the distribution of $\mathcal{F}$, we obtain a large number, say M, of realizations from $\tilde{\mathcal{F}}$ by repeatedly generating the normal random sample $(\Omega_1,\ldots,\Omega_n)$ while fixing the data $\{X_i,\delta_i,Z_i; i=1,\ldots,n\}$ at their observed values. Then using this simulation method, we may determine an approximate critical value of the test. Specifically, the p-value of the test can be obtained as follows:

$$p=\frac{1}{M}\sum_{k=1}^{M}I(\tilde{\mathcal{F}}_k>\mathcal{F}_n),$$

where $\tilde{\mathcal{F}}_k$ (k=1,…,M) are M realizations from $\tilde{\mathcal{F}}$.

For the covariate-dependent case where C depends on Z, an analogous procedure can be developed. Let $G(t|z)$ be the censoring survival distribution of C given Z=z, and

$$\hat{G}(t|z)=\exp\{-\hat{\Lambda}_0(t)\exp(\hat{\gamma}'z)\}.$$

After some algebraic manipulation, model (2) leads to

$$m_0(t)=\frac{1}{H(t,z)}\int_t^{\tau}\int_0^{z}\frac{(s-t)G(t|w)}{g(\beta_0'w)G(s|w)}H_1(ds,dw). \tag{19}$$

Let us denote the right-hand side of (19) by $V^{*}(t,z)$. Note again that the left-hand side is independent of the variable z, and $V^{*}(t,z)$ can be estimated by $V_n^{*}(t,z)$, where

$$V_n^{*}(t,z)=\frac{1}{H_n(t,z)}\left\{\int_t^{\tau}\int_0^{z}\frac{(s-t)\hat{G}(t|w)}{g(\hat{\beta}_b'w)\hat{G}(s|w)}H_{1n}(ds,dw)\right\}.$$

Similarly, for checking the adequacy of model (2) under the covariate-dependent case, we use the Kolmogorov–Smirnov type test statistic $\mathcal{F}_n^{(2)}$, where

$$\mathcal{F}_n^{(2)}=\sup_{t,z}|\theta_n^{*}(t,z)|$$

and $\theta_n^{*}(t,z)=n^{1/2}\{V_n^{*}(t,z)-V_n^{*}(t,z_u)\}$.

Under model (2), we can also show that $\theta_n^{*}(t,z)$ converges to a zero-mean Gaussian process $W^{*}(t,z)$ whose covariance function at (t,z) and $(t^{\dagger},z^{\dagger})$ can be estimated consistently by $\hat{\sigma}^{*}(t,z;t^{\dagger},z^{\dagger})=n^{-1}\sum_{i=1}^{n}\hat{\eta}_i^{*}(t,z)\hat{\eta}_i^{*}(t^{\dagger},z^{\dagger})$, where $\hat{\eta}_i^{*}(t,z)=\hat{\rho}_i^{*}(t,z)-\hat{\rho}_i^{*}(t,z_u)$,

$$\begin{aligned}\hat{\rho}_i^{*}(t,z)=&\frac{1}{H_n(t,z)}\int_t^{\tau}\left[\int_u^{\tau}\int_0^{z}\frac{(s-t)\exp\{\hat{\gamma}'w\}\hat{G}(t|w)}{g(\hat{\beta}_b'w)\hat{G}(s|w)}H_{1n}(ds,dw)\right]\frac{d\hat{M}_i^{c}(u)}{S^{(0)}(u;\hat{\gamma})}+\frac{1}{H_n(t,z)}\int_t^{\tau}\int_0^{z}\frac{(s-t)\exp\{\hat{\gamma}'w\}\hat{G}(t|w)}{g(\hat{\beta}_b'w)\hat{G}(s|w)}\left[\int_t^{s}(w-\overline{Z}_r(v;\hat{\beta}_b)'\,d\hat{\Lambda}_0(v)\right]\\&\times H_{1n}(ds,dw)D_n^{-1}\int_0^{\tau}\{Z_i-\overline{Z}_r(u;\hat{\beta}_b)\}\,d\hat{M}_i^{c}(u)+\frac{\delta_i(X_i-t)\hat{G}(t|Z_i)}{H_n(t,z)\hat{G}(X_i|Z_i)g(\hat{\beta}_b'Z_i)}I(X_i\ge t,Z_i\le z)-\frac{V_n(t,z)}{H_n(t,z)}I(X_i\ge t,Z_i\le z)\\&+\frac{1}{H_n(t,z)}\int_t^{\tau}\int_0^{z}\frac{(s-t)h(\hat{\beta}_b'w)\hat{G}(t|w)w'}{\hat{G}(s|w)g(\hat{\beta}_b'w)}H_{1n}(ds,dw)\left\{\frac{\partial U_b(\hat{\beta}_b)}{\partial\beta'}\right\}^{-1}\left[\hat{\xi}_i+\int_0^{\tau}\frac{R_n(u)}{S^{(0)}(u;\hat{\gamma})}\,d\hat{M}_i^{c}(u)+B_nD_n^{-1}\int_0^{\tau}\{Z_i-\overline{Z}_r(u;\hat{\beta}_b)\}\,d\hat{M}_i^{c}(u)\right].\end{aligned} \tag{20}$$

Consequently, $\mathcal{F}_n^{(2)}$ converges in distribution to $\mathcal{F}^{*}=\sup_{t,z}|W^{*}(t,z)|$. As in the independent case, we can show that the distribution of the process $W^{*}(t,z)$ can be approximated by that of the zero-mean Gaussian process $\tilde{W}^{*}(t,z)=n^{-1/2}\sum_{i=1}^{n}\hat{\eta}_i^{*}(t,z)\Omega_i$ based on (20). Thus, the distributions of $\mathcal{F}^{*}$ can be approximated by $\tilde{\mathcal{F}}^{*}=\sup_{t,z}|\tilde{W}^{*}(t,z)|$, and the p-value of the test can be obtained in the same way as before.

When the model-checking procedure reveals lack of fit of model (2) for the covariate-dependent censoring case, it is possible that the assumed residual life function is correct, but that the censoring distribution has been incorrectly modeled. Therefore, it would make sense to separately evaluate the fit of the Cox model assumed for the censoring time using an existing easy-to-implement model checking procedure. Numerous graphical and analytical methods have been suggested for checking the adequacy of the Cox model. For example, Schoenfeld (1980) proposed a chi-squared goodness-of-fit statistic based on the expected and observed numbers of events in cells corresponding to a partition of the Cartesian product of the range of covariates and the time axis. Lin and Wei (1991) presented a test based on the difference of two different estimators for the inverse of the covariance matrix of the maximum partial likelihood estimators. Lin et al. (1993) used cumulative sums of martingale-based residual to produce a comprehensive and objective diagnostic methodology, and Grambsch and Therneau (1994) suggested to check the Cox model by embedding it into a larger model, where the effect of one covariate may vary over time. Other approaches can be found in Therneau and Grambsch (2000).

5. Simulation studies

We conducted simulation studies to assess the performance of the estimation procedures proposed in Sections 2 and 3 with the focus on estimating β_0 in model (2). In these studies, the survival time T was generated from model (2) with $\beta_0=0$ or 0.5, and the baseline MRLF was taken to be $m_0(t)=-0.5t+1$, which corresponds to a rescaled beta distribution (Oakes and Dasu, 1990). The covariate Z was assumed to be a Bernoulli random variable with success probability 0.5. We considered three choices for the link function $g(x)$: $g_1(x)=1+x$, $g_2(x)=e^x$ and $g_3(x)=\log(1+e^x)$. The censoring time C was generated from the exponential distribution with hazard rate $\lambda_0 e^{\gamma_0 Z}$ for $\gamma_0=0$ or 1, and λ_0 was chosen to result in two censoring percentages of approximately 10% and 30%. Note that $\gamma_0=0$ corresponds to independent censoring times, while $\gamma_0=1$ gives covariate-dependent censoring times. The results presented below are based on 1000 replications with $n=100$ and 200 for independent censoring, and $n=100$, 200 and 500 for covariate-dependent censoring.

Table 1 shows the results for independent censoring ($\gamma_0=0$). It can been seen that the bias for estimating β_0 is very small and the standard error of the estimator is very accurate for all settings. The 95% empirical coverage probability based on normal approximation are very reasonable, and the results become better when the sample size increases from 100 to 200. Table 2 shows similar results for covariate-dependent censoring ($\gamma_0=1$) except that, for a few settings with larger β_0 and heavier percentage of censoring, the coverage probability is over-estimated. We think this occurs because of the instability of the IPCW approach and the heavier tail of the distribution of T.

We also conducted simulation studies with $n=500$ to compare the two methods proposed in Sections 2 and 3 in the presence of iid censoring times (same setup as in Table 1), a setting where both methods are valid. For the method in

Table 1
Simulation results for independent censoring.

n	β_0	p (%)	$g_1(x)=1+x$				$g_2(x)=e^x$				$g_3(x)=\log(1+e^x)$			
			$\overline{\hat{\beta}_a}$	SE	SEE	CP	$\overline{\hat{\beta}_a}$	SE	SEE	CP	$\overline{\hat{\beta}_a}$	SE	SEE	CP
100	0.0	10	0.0047	0.1225	0.1220	0.945	−0.0022	0.1229	0.1215	0.943	−0.0018	0.1961	0.2061	0.968
100	0.0	30	0.0044	0.1368	0.1360	0.945	−0.0048	0.1364	0.1354	0.953	0.0010	0.2142	0.2322	0.976
100	0.5	10	0.5044	0.1505	0.1531	0.943	0.4998	0.0960	0.0962	0.948	0.5056	0.2009	0.2139	0.970
100	0.5	30	0.5053	0.1653	0.1684	0.952	0.4985	0.1072	0.1054	0.945	0.5060	0.2193	0.2393	0.974
200	0.0	10	0.0014	0.0855	0.0856	0.943	−0.0022	0.0853	0.0855	0.947	−0.0025	0.1420	0.1444	0.952
200	0.0	30	0.0031	0.0952	0.0955	0.949	−0.0014	0.0947	0.0952	0.953	−0.0009	0.1600	0.1625	0.951
200	0.5	10	0.5030	0.1080	0.1079	0.948	0.4990	0.0681	0.0680	0.957	0.5020	0.1480	0.1499	0.950
200	0.5	30	0.5030	0.1180	0.1186	0.948	0.4980	0.0742	0.0744	0.951	0.5030	0.1650	0.1675	0.957

p represents proportion of right-censoring; $\overline{\hat{\beta}_a}$ represents the mean of the point estimates of β_0; SE represents sample standard error of $\hat{\beta}_a$; SEE represents the mean of the standard error of $\hat{\beta}_a$; CP represents the empirical 95% coverage probability.

Table 2
Simulation results for covariate-dependent censoring.

n	β_0	p (%)	$g_1(x)=1+x$				$g_2(x)=e^x$				$g_3(x)=\log(1+e^x)$			
			$\overline{\hat{\beta}_b}$	SE	SEE	CP	$\overline{\hat{\beta}_b}$	SE	SEE	CP	$\overline{\hat{\beta}_b}$	SE	SEE	CP
100	0.0	10	−0.0050	0.1109	0.1149	0.949	−0.0111	0.1114	0.1155	0.958	−0.0200	0.1916	0.2012	0.954
100	0.0	30	−0.0131	0.1326	0.1433	0.949	−0.0218	0.1367	0.1454	0.961	−0.0558	0.2192	0.2531	0.964
100	0.5	10	0.5026	0.1328	0.1450	0.968	0.4987	0.0820	0.0922	0.971	0.4913	0.1902	0.2054	0.973
100	0.5	30	0.5116	0.1508	0.1837	0.980	0.5072	0.0946	0.1174	0.985	0.4775	0.2186	0.2569	0.984
200	0.0	10	−0.0043	0.0788	0.0808	0.946	−0.0074	0.0785	0.0812	0.949	−0.0138	0.1411	0.1428	0.945
200	0.0	30	−0.0119	0.0935	0.1015	0.957	−0.0163	0.0945	0.1028	0.963	−0.0500	0.1800	0.1850	0.931
200	0.5	10	0.4994	0.0954	0.1020	0.959	0.4982	0.0597	0.0650	0.970	0.4926	0.1421	0.1452	0.943
200	0.5	30	0.5046	0.1076	0.1287	0.977	0.5033	0.0665	0.0827	0.986	0.4767	0.1701	0.1825	0.957
500	0.0	10	−0.0022	0.0503	0.0511	0.953	−0.0037	0.0504	0.0513	0.953	−0.0098	0.0896	0.0906	0.954
500	0.0	30	−0.0101	0.0603	0.0641	0.955	−0.0120	0.0611	0.0648	0.966	−0.0377	0.1155	0.1190	0.932
500	0.5	10	0.4990	0.0636	0.0642	0.949	0.4997	0.0402	0.0410	0.955	0.4933	0.0919	0.0919	0.943
500	0.5	30	0.5040	0.0725	0.0812	0.968	0.5055	0.0443	0.0523	0.980	0.4802	0.1108	0.1154	0.963

p represents proportion of right-censoring; $\overline{\hat{\beta}_b}$ represents the mean of the point estimates of β_0; SE represents sample standard error of $\hat{\beta}_b$; SEE represents the mean of the standard error of $\hat{\beta}_b$; CP represents the empirical 95% coverage probability.

Table 3
Comparison of two methods in the presence of iid censoring times.

Method	β_0	p (%)	$g_1(x)=1+x$				$g_2(x)=e^x$				$g_3(x)=\log(1+e^x)$			
			$\bar{\hat{\beta}}$	*SE*	*SEE*	*CP*	$\bar{\hat{\beta}}$	*SE*	*SEE*	*CP*	$\bar{\hat{\beta}}$	*SE*	*SEE*	*CP*
Independent censoring	0.0	10	−0.0017	0.0551	0.0538	0.941	−0.0032	0.0550	0.0539	0.949	−0.0036	0.0920	0.0907	0.945
	0.0	30	−0.0005	0.0622	0.0599	0.950	−0.0024	0.0619	0.0599	0.946	−0.0021	0.1046	0.1018	0.948
	0.5	10	0.4987	0.0691	0.0679	0.941	0.4979	0.0433	0.0429	0.945	0.4988	0.0962	0.0943	0.944
	0.5	30	0.4978	0.0754	0.0745	0.949	0.4973	0.0476	0.0469	0.946	0.4993	0.1089	0.1051	0.942
Covariate-dependent	0.0	10	0.0029	0.0510	0.0525	0.955	0.0016	0.0510	0.0523	0.956	0.0042	0.0931	0.0935	0.946
	0.0	30	0.0037	0.0635	0.0696	0.967	0.0044	0.0626	0.0692	0.974	0.0106	0.1265	0.1300	0.951
	0.5	10	0.5041	0.0647	0.0665	0.959	0.5016	0.0385	0.0423	0.973	0.5054	0.0943	0.0951	0.953
	0.5	30	0.5027	0.0732	0.0880	0.980	0.5014	0.0451	0.0567	0.985	0.5109	0.1195	0.1280	0.967

p represents proportion of right-censoring; $\bar{\hat{\beta}}$ represents the mean of the point estimates of β_0; *SE* represents sample standard error of $\hat{\beta}$; *SEE* represents the mean of the standard error of $\hat{\beta}$; *CP* represents the empirical 95% coverage probability.

Section 3, we took $\gamma_0=0$ in model (8) (also $\hat{\gamma}\equiv 0$), and derived asymptotic properties of $\hat{\beta}_b$ following the same lines as the proof in Appendix A.3. The simulation results are summarized in Table 3. It can be seen from the table that, as expected, the results of the method in Section 2 are very satisfactory while the results of the method in Section 3 are overall satisfactory except for some settings with slight over-coverage. Again, this is because the method in Section 3 uses an IPCW technique, which is less stable than the method in Section 2.

To investigate the asymptotical normality of the proposed estimators of β_0 under both independent and covariate-dependent censoring, we provided some QQ-plots in Fig. 1, which suggest reasonable normal approximations to the finite-sample distributions of the proposed estimators.

6. An application

We applied the proposed estimation procedures to a data set from a clinic trial on lung cancer that has previous been analyzed by others (Lad et al., 1988; Piantadosi, 2005; Chen et al., 2005). The purpose of the trial was to assess the impact of systematic combination chemotherapy on patients' survival. Specifically, survival time of interest includes both time to death and disease-free survival time. Between November 1979 and May 1985, 172 patients were randomized to receive either postoperative radiotherapy (RT) alone or postoperative RT plus chemotherapy with Cytoxan, Adriamycin, and Platinol (RT+CAP) for 6 months and followed until death. The mean follow-up time was 1.5 years. Only 164 patients were eligible for analysis, among which 86 patients were in RT and 78 in RT+CAP group.

In our analysis, we considered examining the effect of treatment and cell type (squamous vs. nonsquamous/mixed) on patients' disease-free survival. For treatment, we let $Z_1=1$ if a patient was in RT + CAP group and 0 otherwise. For cell type, we let $Z_2=1$ if a patient had the squamous cell type and 0 otherwise. We first fit model (8) containing both covariates to the data to determine whether covariate-dependent or independent case should be considered. The logrank test shows that the overall effect of treatment and cell type on the censoring time is insignificant with a p-value of 0.876. The Kaplan–Meier estimates of survival functions of the censoring time for four subgroups were plotted in Fig. 2(a). Thus, for the illustration purpose, we then fit model (2) to the data only under the independent censoring situation (Table 4).

Table 4 shows that the estimation and test of hypothesis results for the effect of each of the covariates by using three different functions for g. The results show that both treatment and cell type had significant effect on the patients' disease-free survival after adjusting the effect of the other. More specifically, patients in RT+CAP group had significantly longer mean residual disease-free life than those in the RT group, and patients having squamous cell type have significantly longer mean residual disease-free life than those having nonsquamous/mixed cell type. Fig. 2(b) and (c) show the difference in survival functions between the treatment groups and two cell type groups, respectively. The above results are consistent with those from Chen et al. (2005) under the proportional mean residual life model and from Piantadosi (2005) under the proportional hazards model. Note that the three functions for $g(\cdot)$ yield similar results, and the result from $g(x)=e^x$ is the least conservative based on the p-values.

We also checked the adequacy of model (2) with both covariate under the three functions of $g(\cdot)$. Based on 500 realizations of $\tilde{\mathcal{F}}$, the KS-type test statistics with p-values in parentheses, are 4288.81 (0.966), 4554.43 (0.946) and 6341.43 (0.958) for $g(x)$ to be $1+x$, e^x and $\log(1+e^x)$, respectively. These results indicate that model (2) fits the data adequately.

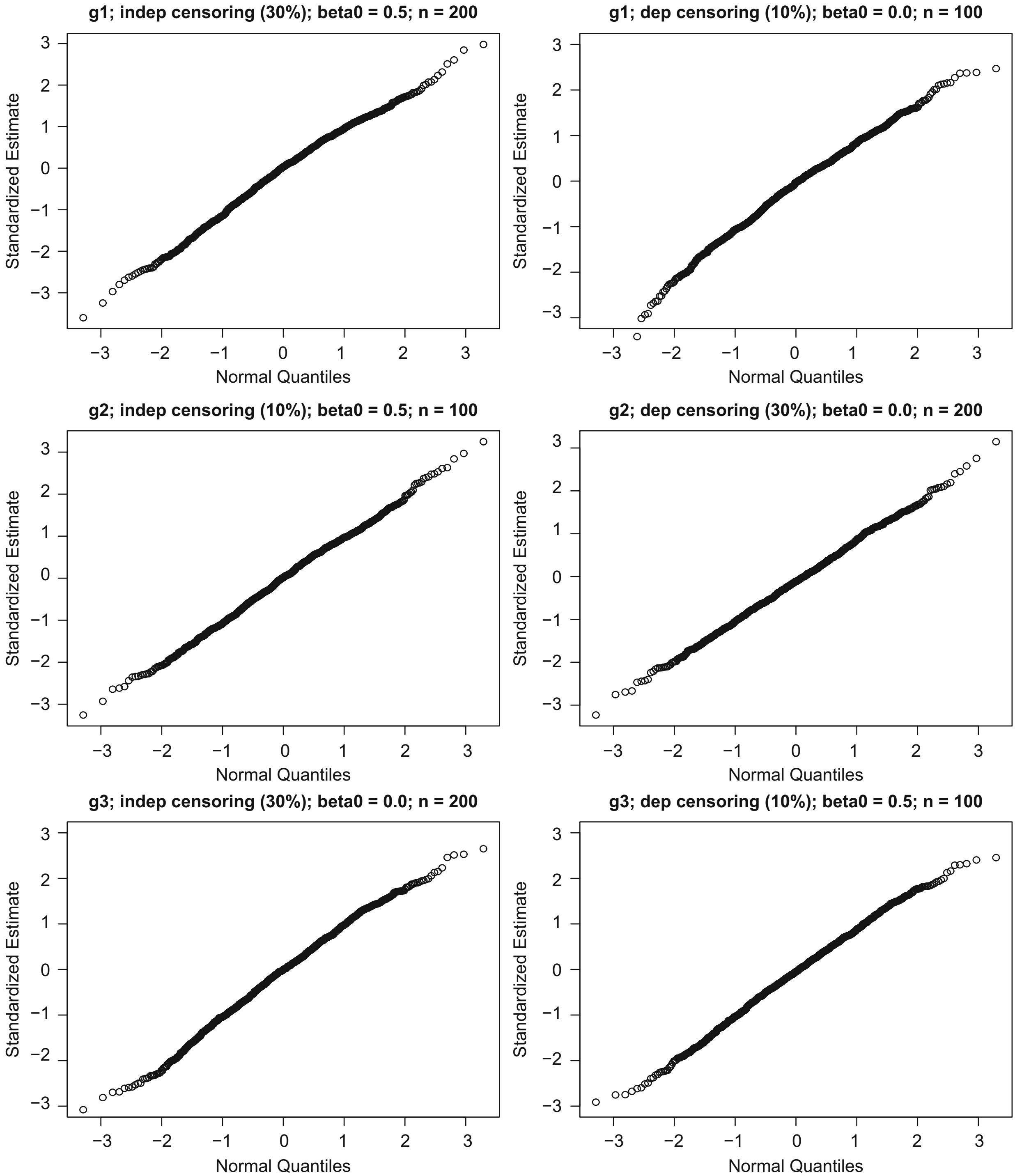

Fig. 1. Normal Q–Q Plots.

7. Concluding remarks

In this article we have studied a class of mean residual life regression models under both independent and covariate-dependent censoring. The proposed models are generalization of the proportional mean residual life model with more choices of the link function $g(\cdot)$. Estimation procedures were proposed for the model parameters, and asymptotic

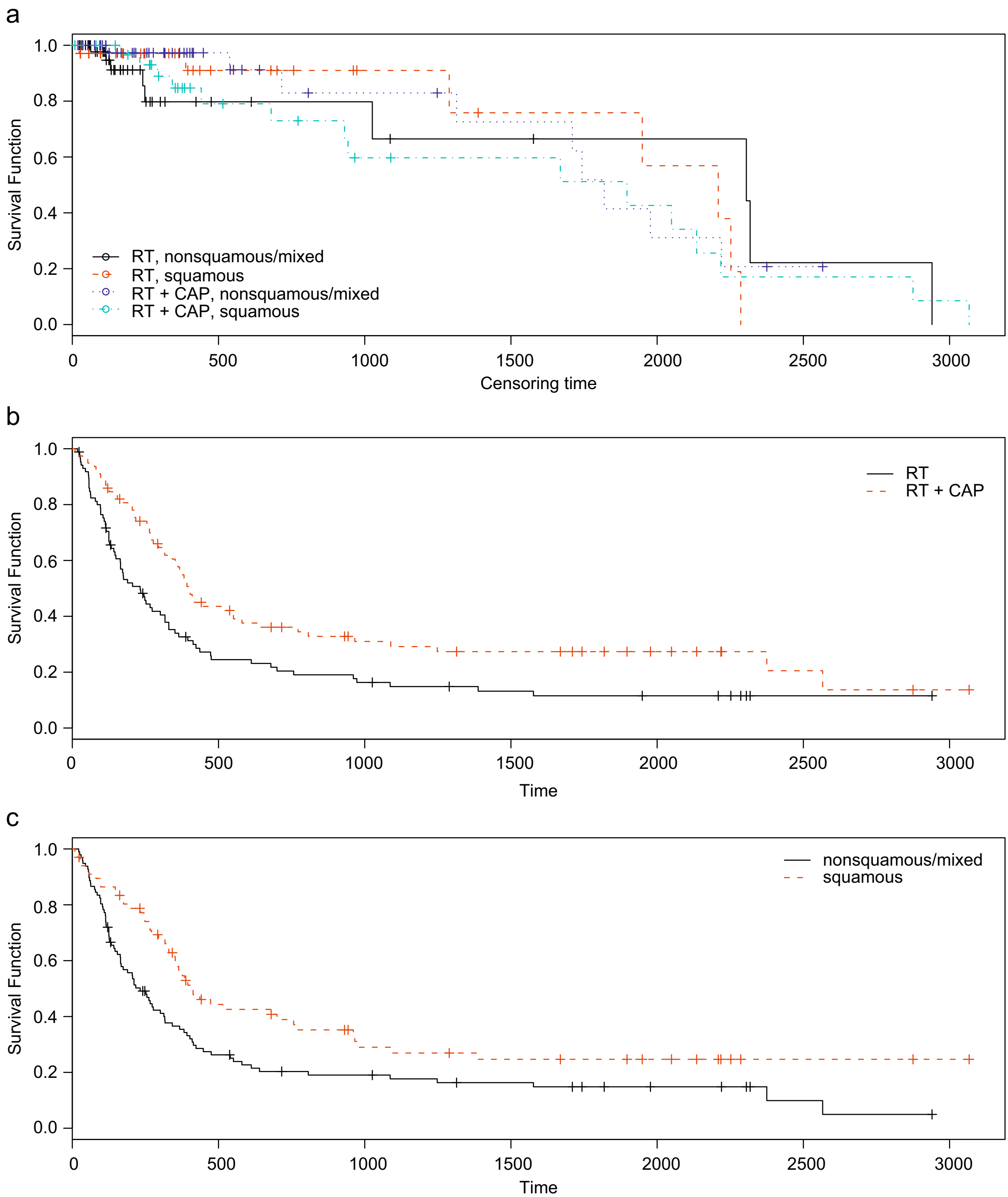

Fig. 2. Kaplan–Meier estimates of survival functions.

properties of the estimators were derived. Simulation results show that the proposed methods work well for the situations considered. The methodology was illustrated in the analysis of the cancer data from a clinic trial.

It is well-known that model checking is always a fundamental issue in regression analysis. We proposed a goodness of fit test for model (2) based on the KS type test statistics. In addition, the Cramér–von Mises type test statistics can also be

Table 4
Estimation of the effects for the lung cancer data.

Covariates	$g(x)$	Parameter estimate	SEE	p-value
Treatment (Z_1)	$1+x$	1.2249	0.6130	0.0457
	e^x	0.7334	0.2413	0.0024
	$\log(1+e^x)$	1.2487	0.5164	0.0156
Cell type (Z_2)	$1+x$	1.0596	0.6338	0.0946
	e^x	0.6528	0.2590	0.0117
	$\log(1+e^x)$	1.1035	0.5462	0.0434

Note: SEE is the standard error estimate; p-value pertains to testing no covariate effect.

used to check the adequacy of model (2):

$$\mathcal{F}_n^{(3)} = \iint \theta_n(t,z)^2 H_n^0(dt,dz),$$

which converges in distribution to

$$\mathcal{F}^{(3)} = \iint W(t,z)^2 H^0(dt,dz),$$

where H^0 and H_n^0 are the distribution function and empirical distribution function of (X_i,Z_i), respectively. Similar to the KS-type test statistics, $\tilde{\mathcal{F}}^{(3)} = \iint \tilde{W}(t,z)^2 H_n^0(dt,dz)$ can be used to approximate the distribution of $\mathcal{F}^{(3)}$.

For covariate-dependent censoring, the proportional hazards model was used as the working model for the censoring time. Of course, we can also choose some other semiparametric regression models as the working model for censoring. For example, we may use the proportional mean residual life model or the additive mean residual life model, then we can obtain the estimators of the censoring model parameters using the approach of Chen and Cheng (2005) or Chen and Cheng (2006). Thus, the estimator of the censoring survival distribution $G(t|z)$ can be obtained using the following inversion formula:

$$G(t|z) = \frac{m_G(0|z)}{m_G(t|z)} \exp\left\{ -\int_0^t m_G(u|z)^{-1}\, du \right\},$$

where $m_G(t|z) = E(C-t|z, C>t)$ is the MRLF of C at t given z. Thus, the unknown parameter in model (2) can be estimated by using the procedure in Section 3.

Since estimating functions (7) and (13) were given in a somewhat ad hoc fashion using the generalized estimating equation methods, it would be worthwhile to further investigate the efficiency of the proposed estimators. In principle, it might be possible to estimate β_0 and $m_0(\cdot)$ more efficiently by the nonparametric maximum likelihood approach, and the resulting inference procedure would be much more complicated. Another issue is that the estimates of $m_0(t)$ may be not monotonic, and there is no guarantee that the finite-sample estimator $\hat{m}_{a0}(t)+t$ or $\hat{m}_{b0}(t)+t$ would maintain the necessary monotonicity at some time point. The incorporation of the pooled-adjacent-violators algorithm may help solving the problem as mentioned in Chen and Cheng (2005).

Acknowledgements

The authors would like to thank the Editor and the reviewer for their constructive and insightful comments and suggestions, which greatly improved the paper. This research was partly supported by the National Natural Science Foundation of China Grants (Nos. 10731010, 10971015 and 10721101) and the National Basic Research Program of China (973 Program) (No. 2007CB814902) and Key Laboratory of RCSDS, CAS (No. 2008DP173182).

Appendix A. Proofs of asymptotic properties

Using the uniform strong law of large numbers (Pollard, 1990, p. 41), we have $\bar{z}_a(t) = \lim_{n\to} \overline{Z}_a(t;\beta_0)$, and $s^{(k)}(t;\gamma) = \lim_{n\to\infty} S^{(k)}(t;\gamma)$ (k=0,1) uniformly in $t \in [0,\tau]$. Let $\bar{z}_r(t) = s^{(1)}(t;\gamma_0)/s^{(0)}(t;\gamma_0)$. In addition, assume that A defined below in (A.4) is nonsingular matrix.

A.1. Asymptotic normality of $U_a(\beta_0)$ and $\hat{\beta}_a$

Note that

$$\sum_{i=1}^{n}[m_0(t)\,dN_i(t)-Y_i(t)\{g(\beta_0'Z_i)^{-1}\,dt+dm_0(t)\}]=m_0(t)\sum_{i=1}^{n}dM_i(t)$$

and

$$\sum_{i=1}^{n}[\hat{m}_{a0}(t;\beta_0)\,dN_i(t)-Y_i(t)\{g(\beta_0'Z_i)^{-1}\,dt+d\hat{m}_{a0}(t;\beta_0)\}]=0.$$

Then it follows that

$$\{\hat{m}_{a0}(t;\beta_0)-m_0(t)\}\sum_{i=1}^{n}dN_i(t)-n\pi_n d\{\hat{m}_{a0}(t;\beta_0)-m_0(t)\}=-m_0(t)\sum_{i=1}^{n}dM_i(t),$$

which is a first-order linear ordinary differential equation in $\hat{m}_{a0}(t;\beta_0)-m_0(t)$. It thus has the closed-form solution given by

$$\hat{m}_{a0}(t;\beta_0)-m_0(t)=-\Phi_n(t)^{-1}\sum_{i=1}^{n}\int_t^{\tau}\frac{\Phi_n(u)m_0(u)}{n\pi_n(u)}\,dM_i(u). \tag{A.1}$$

Write

$$U_a(\beta_0)=n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\overline{Z}_a(t;\beta_0)\}m_0(t)\,dM_i(t)+n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\overline{Z}_a(t;\beta_0)\}[\hat{m}_{a0}(t;\beta 0)-m_0(t)]\,dN_i(t).$$

Using the uniform strong law of large numbers and (A.1), the second term in the right-hand side of the above equation is equivalent to

$$-n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\mu_*(t)m_0(t)\,dM_i(t)+o_p(n^{-1/2}),$$

where

$$\mu_*(t)=\frac{S(t)}{\pi(t)}\int_0^t\frac{1}{S(u)}E[\{h(\beta'Z_i)Z_i-\overline{z}_a(u)\}\,dN_i(u)],$$

$\pi(t)=EY_1(t)$, and $S(t)$ is the marginal survival function of T. Therefore,

$$n^{1/2}U_a(\beta_0)=n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\mu(t)\}m_0(t)\,dM_i(t)+o_p(1),$$

where $\mu(t)=\overline{z}_a(t)+\mu_*(t)$. As a result, $n^{1/2}U_a(\beta_0)$ converges in distribution to zero-mean normal distribution with covariance matrix Σ_a, where

$$\Sigma_a=E\left[\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\mu(t)\}^{\otimes 2}m_0(t)^2\,dN_i(t)\right], \tag{A.2}$$

which can be consistently estimated by $\hat{\Sigma}_a$ defined in Section 2.

Since the censoring time C is independent of T and Z, and

$$\int_t^{\tau}S(u|Z)g(\beta_0'Z)^{-1}\,du=m_0(t)S(t|Z),$$

under model (2), it follows from the uniform strong law of large numbers that

$$\begin{aligned}\frac{\partial\hat{m}_{a0}(t;\beta_0)}{\partial\beta}&=-\Phi_n(t)^{-1}\int_t^{\tau}\frac{\Phi_n(u)}{\pi_n(u)}\left[n^{-1}\sum_{i=1}^{n}Y_i(u)h(\beta_0'Z_i)g(\beta_0'Z_i)^{-1}Z_i\,du\right]\\&=-\frac{1}{S(t)}E\left[h(\beta_0'Z_i)Z_i\int_t^{\tau}S(u|Z_i)g(\beta_0'Z_i)^{-1}\,du\right]+o_p(1)=-m_0(t)\overline{z}_a(t)+o_p(1).\end{aligned} \tag{A.3}$$

Let $\hat{A}=n^{-1}\partial U(\beta_0)/\partial\beta'$, and $h^{(1)}(x)=dh(x)/dx$. Then it follows from (A.3) that

$$\hat{A}=n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\left\{h^{(1)}(\beta_0'Z_i)Z_i^{\otimes 2}-\frac{\sum_{i=1}^{n}Y_i(t)h^{(1)}(\beta_0'Z_i)Z_i^{\otimes 2}}{\sum_{i=1}^{n}Y_i(t)}\right\}[\hat{m}_{a0}(t;\beta_0)\,dN_i(t)-Y_i(t)g(\beta_0'Z_i)^{-1}\,dt]$$

$$
\begin{aligned}
&+n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\overline{Z}_a(t;\beta_0)\}\left[\frac{\partial \hat{m}_{a0}(t;\beta_0)}{\partial \beta'}\,dN_i(t)+Y_i(t)h(\beta_0'Z_i)Z_i'g(\beta_0'Z_i)^{-1}\,dt\right]\\
&=n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\left\{h^{(1)}(\beta_0'Z_i)Z_i^{\otimes 2}-\frac{\sum_{i=1}^{n}Y_i(t)h^{(1)}(\beta_0'Z_i)Z_i^{\otimes 2}}{\sum_{i=1}^{n}Y_i(t)}\right\}[m_0(t)\,dM_i(t)+Y_i(t)\,dm_0(t)]\\
&-n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\overline{Z}_a(t;\beta_0)\}[z_a(t)'\{m_0(t)\,dM_i(t)+Y_i(t)g(\beta_0'Z_i)^{-1}\,dt\\
&+Y_i(t)\,dm_0(t)\}-Y_i(t)h(\beta_0'Z_i)Z_i'g(\beta_0'Z_i)^{-1}\,dt]+o_p(1)=A+o_p(1),
\end{aligned}
$$

where

$$A=E\left[\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\overline{z}_a(t)\}^{\otimes 2}Y_i(t)g(\beta_0'Z_i)^{-1}\,dt\right]. \tag{A.4}$$

Thus, the asymptotic distribution of $\hat{\beta}_a$ follows from a Taylor series expansion of $U_a(\hat{\beta}_a)$ at β_0. For future reference, we display the asymptotic approximation

$$n^{1/2}(\hat{\beta}_a-\beta_0)=-A^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\mu(t)\}m_0(t)\,dM_i(t)+o_p(1). \tag{A.5}$$

A.2. Weak convergence of $\hat{m}_{a0}(t)$

To show the weak convergence of $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$, we first note that

$$n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}=n^{1/2}\{\hat{m}_{a0}(t;\beta_0)-m_0(t)\}+n^{1/2}\{\hat{m}_{a0}(t;\hat{\beta}_a)-\hat{m}_0(t;\beta_0)\}.$$

It follows from (A.1) and the uniform strong law of large numbers that

$$n^{1/2}\{\hat{m}_{a0}(t;\beta_0)-m_0(t)\}=-S(t)^{-1}n^{-1/2}\sum_{i=1}^{n}\int_t^{\tau}\frac{S(u)m_0(u)}{\pi(u)}\,dM_i(u)+o_p(1).$$

Using the Taylor expansion of $\hat{m}_{a0}(t;\hat{\beta}_a)$ together with (A.3), we have

$$n^{1/2}\{\hat{m}_{a0}(t;\hat{\beta}_a)-\hat{m}_{a0}(t;\beta_0)\}=-m_0(t)\overline{z}_a(t)'n^{1/2}\{\hat{\beta}_a-\beta_0\}+o_p(1).$$

Thus, it follows from (A.5) that

$$n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}=n^{-/2}\sum_{i=1}^{n}\varphi_i(t)+o_p(1),$$

where

$$\varphi_i(t)=-S(t)^{-1}\int_t^{\tau}\frac{S(u)m_0(u)}{\pi(u)}\,dM_i(u)+m_0(t)\overline{z}_a(t)'A^{-1}\int_0^{\tau}\{h(\beta_0'Z_i)Z_i-\mu(u)\}m_0(u)\,dM_i(u).$$

Because φ_i ($i=1,\ldots,n$) are independent zero-mean random variables for each t, the multivariate central limit theorem implies that $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$ ($0\le t\le\tau$) converges in finite-dimensional distributions to zero-mean Gaussian process. Using the modern empirical theory as Lin et al. (2000, 2001), we can show that $n^{1/2}\{\hat{m}_{a0}(t)-m_0(t)\}$ is tight and converges weakly to zero-mean Gaussian process with covariance function $\Gamma_a(s,t)=E\{\varphi_i(s)\varphi_i(t)\}$ at (s,t), which can be estimated by $\hat{\Gamma}_a(s,t)$ given in Section 2.

A.3. Asymptotic normality of $U_b(\beta_0)$ and $\hat{\beta}_b$

It can be checked that

$$
\begin{aligned}
U_b(\beta_0)&=n^{-1}\sum_{i=1}^{n}\{h(\beta_0'Z_i)Z_i-\tilde{Z}_i\}\delta_iG(X_i;\gamma_0,\Lambda_0)^{-1}+n^{-1}\sum_{i=1}^{n}\{h(\beta_0'Z_i)Z_i-\tilde{Z}_i\}\delta_i[\hat{G}_i(X_i;\hat{\gamma},\hat{\Lambda}_0)^{-1}-G(X_i;\gamma_0,\Lambda_0)^{-1}]\\
&-\int_0^{\tau}Q(t)\{L_n(t;\beta_0,\hat{\gamma},\hat{\Lambda}_0)-L(t)\}\,dt+o_p(n^{-1/2}),
\end{aligned} \tag{A.6}
$$

where $Q(t)=\lim_{n\to\infty}Q_n(t)$, $L(t)$=$L_1(t)/(L_2(t)L_3(t))$, $L_k(t)=\lim_{n\to\infty}L_{kn}(t;\beta_0,\gamma_0,\Lambda_0)$ (k=1,2,3), and

$$\tilde{Z}_i=\int_0^{\tau}h(\beta_0'Z_i)Z_ig(\beta_0'Z_i)^{-2}(X_i-t)^{+}L(t)\,dt.$$

It is well known that (Fleming and Harrington, 1991, p. 299)

$$\hat{\Lambda}_0(t)-\Lambda_0(t)=n^{-1}\sum_{i=1}^{n}\int_0^t \frac{dM_i^c(u)}{s^{(0)}(u;\gamma_0)}-\int_0^t \bar{z}_r(u)'\,d\Lambda_0(u)(\hat{\gamma}-\gamma_0)+o_p(n^{-1/2}),$$

$$\hat{\gamma}-\gamma_0=D^{-1}n^{-1}\sum_{i=1}^{n}\int_0^\tau \{Z_i-\bar{z}_r(u)\}\,dM_i^c(u)+o_p(n^{-1/2}),$$

$$M_i^c(t)=N_i^c(t)-\int_0^t Y_i(u)\exp\{\gamma_0'Z_i\}\,d\Lambda_0(u),$$

and $D=\lim_{n\to\infty}D_n$. Thus,

$$L_{kn}(t;\beta_0,\hat{\gamma},\hat{\Lambda}_0)-L_{kn}(t;\beta_0,\gamma_0,\Lambda_0)=n^{-1}\sum_{i=1}^{n}\int_0^\tau \frac{R_k(t,u)}{s^{(0)}(u)}\,dM_i^c(u)+P_k(t)(\hat{\gamma}-\gamma_0)+o_p(n^{-1/2})$$

and

$$L_{kn}(t;\beta_0,\gamma_0,\Lambda_0)-L_k(t)=n^{-1}\sum_{i=1}^{n}\xi_{ki}(t)+o_p(n^{-1/2}),$$

where $\xi_{ki}(t)=V_{ki}(t;\beta_0)G_i(t;\gamma_0,\Lambda_0)^{-1}-L_k(t)$, and $R_k(t,u)$ and $P_k(t)$ are the limits of $R_{kn}(t,u)$ and $P_{kn}(t)$, respectively. Therefore, using the functional Delta-method, it follows from (A.6) that

$$n^{1/2}U_b(\beta_0)=n^{-1/2}\sum_{i=1}^{n}\left[\xi_i+\int_0^\tau \frac{R(t)}{S^{(0)}(t;\hat{\gamma})}\,dM_i^c(t)+BD^{-1}\int_0^\tau \{Z_i-\bar{z}_r(t)\}\,dM_i^c(t)\right]+o_p(1),$$

where

$$\xi_i=\frac{\delta_i\{h(\beta_0'Z_i)Z_i-\tilde{Z}_i\}}{G_i(X_i;\gamma_0,\Lambda_0)}-\int_0^\tau Q(t)\left[\frac{\xi_{1i}(t)}{L_2(t)L_3(t)}-\frac{L_1(t)\xi_{2i}(t)}{L_2(t)^2L_3(t)}-\frac{L_1(t)\xi_{3i}(t)}{L_2(t)L_3(t)^2}\right]dt,$$

and $R(t)$ and B are the limits of $R_n(t)$ and B_n given in (14), respectively. Utilizing the multivariate central limit theorem, $n^{1/2}U_b(\beta_0)$ is asymptotically normal with mean zero and covariance matrix Σ_b, where

$$\Sigma_b=E\left[\xi_i+\int_0^\tau \frac{R(t)}{S^{(0)}(t;\hat{\gamma})}\,dM_i^c(t)+BD^{-1}\int_0^\tau \{Z_i-\bar{z}_r(t)\}\,dM_i^c(t)\right]^{\otimes 2}.$$

An empirical covariance estimator $\hat{\Sigma}_b$ defined by (14), in which all unknown quantities are replaced with their observed counterparts, converges in probability to Σ_b.

It can be checked that $U_b(\beta)$ converges almost surely uniformly in a closed set of β to $u_b(\beta)$, and $u_b(\beta_0)=0$, where

$$u_b(\beta)=E\{h(\beta'Z)Z\}-\int_0^\tau \frac{E\{h(\beta'Z)Zg(\beta'Z)^{-2}(T-t)^+\}}{E\{g(\beta'Z)^{-2}(T-t)^+\}}\frac{E\{g(\beta'Z)^{-1}I(T>t)\}}{E\{g(\beta'Z)^{-1}T\}}\,dt.$$

For any function $w(Z)$, define

$$E_{t,\beta_0}\{w(Z)\}=\frac{E\{w(Z)g(\beta_0'Z)^{-2}(T-t)^+\}}{E\{g(\beta_0'Z)^{-2}(T-t)^+\}}$$

and

$$E_{\beta_0}\{w(Z)\}=\frac{E\{w(Z)g(\beta_0'Z)^{-1}T\}}{E\{g(\beta_0'Z)^{-1}T\}}.$$

Then

$$\frac{\partial u_b(\beta_0)}{\partial\beta'}=2\int_0^\tau \mathrm{Var}_{t,\beta_0}\{h(\beta_0'Z)Z\}E_{\beta_0}\{S(t|Z)m(0|Z)^{-1}\}-\frac{1}{m_0(0)}\int_0^\tau \mathrm{Cov}_{\beta_0}\{h(\beta_0'Z)Z,S(t|Z)\}\ \mathrm{Cov}_{t,\beta_0}\{h(\beta_0'Z)Z,g(\beta_0'Z)^{-1}\}\,dt.$$

We observe that $S(t|Z)$ is decreasing function of $g(\beta_0'Z)^{-1}$, which implies that $\{h(\beta_0'Z),S(t|Z)\}$ and $\mathrm{cov}_{t,\beta_0}\{h(\beta_0'Z)Z,g(\beta_0'Z)^{-1}\}$ must take opposite signs (Maguluri and Zhang, 1994). This gives that $\partial u_b(\beta_0)/\partial\beta'$ is positive definite. Thus, it follows that $\hat{\beta}_b$ is consistent and unique in a neighborhood of β_0. A Taylor series expansion of $U_b(\hat{\beta}_b)$ yields that $n^{1/2}(\hat{\beta}_b-\beta_0)$ is asymptotically normal with mean zero and covariance matrix given by

$$\left\{\frac{\partial u_b(\beta_0)}{\partial\beta'}\right\}^{-1}\Sigma_b\left\{\frac{\partial u_b(\beta_0)'}{\partial\beta}\right\}^{-1}.$$

A.4. Weak convergence of $\theta_n(t,z)$

It can be checked that

$$\phi_n(t,z) = \frac{n^{1/2}\{B_n(t,z)-B(t,z)\}}{H(t,z)} - \frac{B(t,z)}{H(t,z)^2} n^{1/2}\{H_n(t,z)-H(t,z)\} + o_p(1), \tag{A.7}$$

where

$$B_n(t,z) = \int_t^\tau \int_0^z \frac{(s-t)\hat{G}(t)}{g(\hat{\beta}_a' w)\hat{G}(s)} H_{1n}(ds,dw)$$

and

$$B(t,z) = \int_t^\tau \int_0^z \frac{(s-t)G(t)}{g(\beta_0' w)G(s)} H_1(ds,dw).$$

Consider the martingale representation of the Kaplan–Meier estimator (Fleming and Harrington, 1991, p. 97)

$$\frac{\hat{G}(t)-G(t)}{G(t)} = -\int_0^t \frac{\hat{G}(u-)}{G(u)} \frac{\sum_{i=1}^n M_i^c(u)}{n\pi_n(u)}, \tag{A.8}$$

where $M_i^c(t) = N_i^c(t) - \int_0^t I(X_i \geq u)\, d\Lambda^c(u)$ and $\Lambda^c(t) = -\log(G(t))$ is the cumulative hazard function of the censoring times. It is well known that $M_i^c(t)$ ($i=1,\ldots,n$) are martingales with respect to the $\sigma-$filtration

$$\sigma\{I(X_i \geq u), I(X_i \leq u, \delta_i = 0), Z_i : 0 \leq u \leq t,\ i = 1,\ldots,n\}.$$

It follows from (A.8) and a Taylor series expansion that

$$\begin{aligned} B_n(t,z) - B(t,z) = {} & n^{-1} \int_t^\tau \int_0^z \frac{(s-t)G(t)}{g(\beta_0' w)G(s)} \left(\int_t^s \frac{dM_i^c(u)}{\pi(u)} \right) H_1(ds,dw) + \int_t^\tau \int_0^z \frac{(s-t)G(t)}{g(\beta_0' w)G(s)} [H_{1n}(ds,dw) - H_1(ds,dw)] \\ & - \int_t^\tau \int_0^z \frac{(s-t)h(\beta_0' w)w' G(t)}{g(\beta_0' w)G(s)} H_1(ds,dw)(\hat{\beta}_a - \beta_0) + o_p(1). \end{aligned} \tag{A.9}$$

Thus, combining (A.5), (A.7) and (A.9), we have

$$\theta_n(t,z) = \phi_n(t,z) - \phi_n(t,z_u) = n^{-1/2} \sum_{i=1}^n \eta_i(t,z) + o_p(1),$$

where $\eta_i(t,z) = \rho_i(t,z) - \rho_i(t,z_u)$, and

$$\begin{aligned} \rho_i(t,z) = {} & \frac{G(t)}{H(t,z)} \int_t^\tau \left[\int_u^\tau \int_0^z \frac{s-t}{G(s)g(\beta_0' w)} H_1(ds,dw) \right] \frac{dM_i^c(u)}{\pi(u)} + \frac{\delta_i(X_i - t)G(t)}{H(t,z)G(X_i)g(\beta_0' Z_i)} I(X_i \geq t, Z_i \leq z) - \frac{V(t,z)}{H(t,z)} I(X_i \geq t, Z_i \leq z) \\ & + \frac{G(t)}{H(t,z)} \int_t^\tau \int_0^z \frac{(s-t)h(\beta_0' w)w'}{g(\beta_0' w)G(s)} H_1(ds,dw) A^{-1} \int_0^\tau \{h(\beta_0' Z_i)Z_i - \mu(u)\} m_0(u)\, dM_i(u). \end{aligned}$$

Thus, by the same arguments as those of Appendix A.5 in Lin et al. (2000), $\theta_n(t,z)$ converges weakly to zero-mean Gaussian process with covariance function $\sigma(t,z;t^*,z^*) = E\{\eta_i(t,z)\eta_i(t^*,z^*)\}$ at (t,z) and (t^*,z^*), which can be consistently estimated by $\hat{\sigma}(t,z;t^*,z^*)$ given in Section 4.

References

Cai, J., Schaubel, D.E., 2004. Marginal mean/rates models for multiple type recurrent event data. Lifetime Data Analysis 10, 121–138.
Chen, Y.Q., 2007. Additive regression of expectancy. Journal of the American Statistical Association 102, 153–166.
Chen, Y.Q., Jewell, N.P., Lei, X., Cheng, S.C., 2005. Semiparametric estimation of proportional mean residual life model in presence of censoring. Biometrics 61, 170–178.
Chen, Y.Q., Cheng, S.C., 2005. Semiparametric regression analysis of mean residual life with censored survival data. Biometrika 92, 19–29.
Chen, Y.Q., Cheng, S.C., 2006. Linear life expectancy regression with censored data. Biometrika 93, 303–313.
Cox, D.R., 1962. Renewal Theory. Spottiswoode Ballantyne, London.
Cox, D.R., 1972. Regression models and life-tables (with discussion). Journal of the Royal Statistical Society, Series B 34, 187–220.
Fleming, T.R., Harrington, D.P., 1991. Counting Processes and Survival Analysis. John Wiley, New York.
Grambsch, P.M., Therneau, T.M., 1994. Proportional hazards tests and diagnostics based on weighted residuals. Biometrika 81, 515–526.
Lad, T., Rubinstein, L., Sadeghi, A., 1988. The benefit of adjuvant treatment for resected locally advanced non-small-cell lung cancer. Journal of Clinical Oncology 6, 9–17.
Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.
Lin, D.Y., Wei, L.J., 1991. Goodness-of-fit tests for the general Cox regression model. Statistica Sinica 1, 1–17.
Lin, D.Y., Wei, L.J., Yang, I., Ying, Z., 2000. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society, Series B 62, 711–730.
Lin, D.Y., Wei, L.J., Ying, Z., 1993. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika 80, 557–572.

Lin, D.Y., Wei, L.J., Ying, Z., 2001. Semiparametric transformation models for point processes. Journal of the American Statistical Association 96, 620–628.
Maguluri, G., Zhang, C.-H., 1994. Estimation in the mean residual life regression model. Journal of the Royal Statistical Society, Series B 56, 477–489.
Oakes, D., Dasu, T., 1990. A note on residual life. Biometrika 77, 409–410.
Piantadosi, S., 2005. Clinical Trials: A Methodologic Perspective, second ed. John Wiley, New York.
Prentice, R.L., Self, S.G., 1983. Asymptotic distribution theory for Cox-type regression models with general relative risk form. The Annals of Statistics 11, 804–813.
Pollard, D., 1990. Empirical Processes: Theory and Applications. Institute of Mathematical Statistics, Hayward, CA.
Robins, J.M., Rotnitzky, A., 1992. Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell, N., Dietz, K., Farewell, V. (Eds.), AIDS Epidemiology-Methodologic Issues. Birkhauser, Boston, pp. 24–33.
Schoenfeld, D., 1980. Chi-squared goodness-of-fit tests for the proportional hazards regression model. Biometrika 67, 145–153.
Therneau, T.M., Grambsch, P.M., 2000. Modeling Survival Data: Extending the Cox Model. Springer, New York.
Yuen, K.C., Zhu, L.X., Tang, N.Y., 2003. On the mean residual life regression model. Journal of Statistical Planning and Inference 113, 685–698.