The Canadian Journal of Statistics
Vol. 34, No. 4, 2010, Pages 665–679
La revue canadienne de statistique

Semiparametric median residual life model and inference

Yanyuan MA[1] and Guosheng YIN[2,*]

[1]*Department of Statistics, Texas A&M University, College Station, TX 77843, USA*
[2]*Department of Statistics and Actuarial Science, The University of Hong Kong, Pokfulam Road, Hong Kong*

Key words and phrases: Residual lifetime; median regression; survival function; identifiability; semiparametric model.

MSC 2000: Primary 62N01; secondary 62N02.

Abstract: For randomly censored data, the authors propose a general class of semiparametric median residual life models. They incorporate covariates in a generalized linear form while leaving the baseline median residual life function completely unspecified. Despite the non-identifiability of the survival function for a given median residual life function, a simple and natural procedure is proposed to estimate the regression parameters and the baseline median residual life function. The authors derive the asymptotic properties for the estimators, and demonstrate the numerical performance of the proposed method through simulation studies. The median residual life model can be easily generalized to model other quantiles, and the estimation method can also be applied to the mean residual life model. *The Canadian Journal of Statistics* 34: 665–679; 2010 © 2010 Statistical Society of Canada

Résumé: Les auteurs proposent une classe générale de modèles semi-paramétriques pour la durée de vie résiduelle médiane de données censurées aléatoirement. Les covariables sont incorporées dans la partie linéaire généralisée tandis que la durée de vie résiduelle médiane de référence est entièrement non spécifiée. Malgré que la fonction de survie soit non identifiable pour une fonction de durée de vie résiduelle médiane donnée, une procédure simple et naturelle est proposée pour estimer les paramètres de régression ainsi que la fonction de durée de vie résiduelle médiane de référence. Les auteurs obtiennent les propriétés asymptotiques des estimateurs et ils illustrent la performance numérique de la méthode proposée à l'aide de simulation. Le modèle de durée de vie résiduelle médiane peut facilement être généralisé pour modéliser les autres quantiles et la méthode d'estimation peut aussi être appliquée au modèle de durée de vie résiduelle moyenne. *La revue canadienne de statistique* 34: 665–679; 2010 © 2010 Société statistique du Canada

1. INTRODUCTION

The residual lifetime characterizes the remaining survival time of a subject, given that the subject has already survived up to time t. Correspondingly, the mean residual lifetime is the remaining life expectancy conditional on survival at time t, which is defined as

$$m(t) = E(T - t|T > t) \quad \text{for } t \geq 0,$$

where T is the failure time. If we denote X as a p-dimensional vector of covariates, Oakes & Dasu (1990) proposed the proportional mean residual life model

$$m(t|X) = m(t)\exp(X^{\mathrm{T}}\beta),$$

* *Author to whom correspondence may be addressed.*
E-mail: gyin@hku.hk

where $m(t|X) = E(T - t|T > t, X)$ and $m(t)$ is the unknown and unspecified baseline mean residual lifetime. When there is no censoring, Maguluri & Zhang (1994) proposed the Cox-type estimating equation based on the proportional hazards structure (Cox, 1972), and studied the forward recurrence times in the renewal process. Under general censorship, Oakes & Dasu (2003) developed the inferential procedures for one-sample and two-sample cases. Recently, Chen & Cheng (2005) proposed a semiparametric regression model and a quasi-partial-score approach based on counting process theories.

In contrast to the proportional mean residual lifetime model, Hall & Wellner (1984) studied a class of survival distributions characterized by linear mean residual lifetimes. To handle covariates in a general regression setting, Chen & Cheng (2006) proposed a semiparametric linear life expectancy model

$$m(t|X) = m(t) + X^{\mathrm{T}}\beta,$$

where the baseline function $m(t)$ is left unspecified. Both the proportional mean residual life model and the linear life expectancy model can be viewed as special cases of a more general formulation given by

$$m(t|X) = g\{m(t), X; \beta\},$$

where $g(\cdot)$ is a known link function. The mean residual lifetime has an explicit one-to-one correspondence to the conditional survival function

$$S(t|X) = \frac{m(0|X)}{m(t|X)} \exp\left\{-\int_0^t \frac{1}{m(u|X)}\,\mathrm{d}u\right\}. \tag{1}$$

This nice feature greatly facilitates the semiparametric estimation procedure, because it uniquely determines a closed form of the conditional survival function $S(t|X)$ for a given mean residual lifetime $m(t|X)$. Also, see Chen, Jewell & Cheng (2005) and Chen (2007).

However, it may be undesirable to use the mean residual lifetime when the underlying distribution is highly skewed or heavy tailed. In such cases, a single long-term survivor would impose a notable influence on the mean because the mean of survival is sensitive to outliers. Furthermore, we observe that the mean residual lifetime is given by

$$E(T|T > t) = t + \frac{1}{S(t)}\int_t^\infty S(u)\,\mathrm{d}u,$$

in which the integration does not converge for an improper survival function $S(t)$, or for any $S(t)$ that decreases at a rate of $1/t$ or slower. Thus, a mean residual lifetime may not even exist. As a natural alternative, instead of modelling the mean residual lifetime, one could consider the median or any other quantile of the residual lifetime. The median residual life model can be viewed as a more flexible and robust version than the mean residual life model in two different aspects. First, it is well known that the median is less sensitive to outliers in contrast to the mean, which is also true for the residual lifetimes. Second, there are an infinite number of survival functions that correspond to the same median residual life function, which is a multiple-to-one relationship. Thus, the median residual life model is less restrictive than the model based on the mean residual lifetime which is one-to-one correspondent to the survival function as in (1). As a result, inference based on the median residual life model would be more robust to model misspecification. Moreover, by modelling a broad range of quantiles of the residual lifetime, we

can offer a more complete evaluation of the distribution of the residual lifetime. Based on these considerations, the median and other quantiles of the residual lifetime are often more relevant in practice.

Quantile regression models provide a natural way for robust and complete assessment of covariate effects, which have been extensively studied in various contexts (Koenker & Bassett, 1978; Yu, Lu & Stander, 2003; and Koenker, 2005). In addition, quantile regression has been extended to modelling censored data, including both fixed and random censoring cases (Powell, 1984; Ying, Jung & Wei, 1995; Yang, 1999; Bang & Tsiatis, 2002; Portnoy, 2003; and Peng & Huang, 2008). However, limited research has been carried out with respect to the median residual lifetime. Gelfand & Kottas (2003) proposed a Bayesian semiparametric approach to the median residual life model. For the two-sample comparison, Jeong, Jung & Costantino (2008) studied a nonparametric median residual life function that can be constructed through the Kaplan–Meier survival estimator. In an extension, Jung, Jeong & Bandos (2009) proposed a general regression model based on the median residual lifetime, which is built upon the work by Ying, Jung & Wei (1995). These methods typically begin by modelling a survival function and then make inference on the median residual life function. Such backward inferential procedures are not very appealing or intuitive, and in certain cases, the survival function may lead to a complicated or even intractable median residual life function. The conventional mean residual life model approach entails reconstructing the survival function, then proceeding with the typical nonparametric maximum likelihood estimation in a semiparametric setting. However, the non-identifiability of the survival function from the median residual life specification invalidates such an approach. This dominates the major difference in the treatment between the median and mean residual life models.

Our goal is to model the median residual lifetime by incorporating covariates, and to provide a computationally simple estimation and inference procedure. If we focus on the τth quantile of the residual lifetime, then a general semiparametric quantile residual life model can be expressed as

$$q_\tau(T - t|T > t, X) = g\{m(t), X; \beta\}, \tag{2}$$

where $q_\tau(\cdot|T > t, X)$ is the conditional quantile function given $T > t$ and X. In practice, the covariates are often modeled through a generalized linear form, that is, $g\{m(t), X; \beta\} = g\{m(t), X^{\mathrm{T}}\beta\}$ and the baseline residual life function $m(t)$ is unspecified. The function $g(\cdot)$ can be chosen as linear or of any other convenient form, which greatly enhances the model flexibility. For the median residual life model corresponding to $\tau = 0.5$, a survival function cannot be uniquely identified from the median model specification. This is true for any other quantile value $0 < \tau < 1$ as well. Thus, the standard approaches used in the mean residual life model are not applicable to the median residual life model. This includes the direct use of the mean residual life function to construct the corresponding survival function. In lieu of that approach, we introduce a simple and computationally effective procedure to estimate β and $m(t)$ directly in the median residual life model. The proposed estimation procedure is very general and intuitive, and can also be applied to the mean residual life model. In addition, the proposed procedure estimates the parameters in the median residual life model without the need to recover the survival function.

The rest of the paper is organized as follows. In Section 2, we introduce the notation, present the estimation procedure, and derive the asymptotic properties of the parameter estimates for β and $m(t)$. In Section 3, we conduct simulation studies to examine the finite sample properties of the proposed estimators. We illustrate the proposed model with a real data set from a leukemia study in Section 4, and give some concluding remarks in Section 5.

2. MEDIAN RESIDUAL LIFE MODEL

2.1. Two-Stage Estimation Procedure

Under random censorship, let T_i be the failure time and C_i be the censoring time for subject i. We observe $Y_i = \min(T_i, C_i)$, and the censoring indicator $\Delta_i = I(T_i < C_i)$, where $I(\cdot)$ is the indicator function. We assume that C_i is independent of T_i and covariates X_i. In addition, the observations (Y_i, Δ_i, X_i) are assumed to be independent and identically distributed for $i = 1, \ldots, n$.

The τth quantile residual life model in (2) can be equivalently expressed as

$$S[t + g\{m(t), X; \beta\}|X] = (1 - \tau)S(t|X), \tag{3}$$

where $S(t|X)$ is the conditional survival function given covariates X. The functional equation in (3) can be recognized as Schröder's equation, which does not yield a unique solution for $S(t|X)$, but a very rich family of survival functions (Gupta & Langford, 1984). In general, the solutions of Schröder's equation cannot be written out explicitly. In fact, any solution $S(t|X)$ would involve the inverse of a function $f(t, X)$, where f itself is defined through a recursive formula. Specifically, a basic solution of (3) has the form of

$$S(t|X) = (1 - \tau)^{f^{-1}(t,X)+1},$$

where f^{-1} represents the inverse function of f. The function f is recursively defined as

$$f\{h(u + 1, X), X\} = (1 - \tau)f\{h(u, X), X\},$$

where h is a function that is strictly increasing in $[0, 1]$ and also satisfies $h(-1, X) = 0,\ h(1, X) = g\{m(0), X; \beta\}$, and $h(u, X) = g[m\{h(u - 1, X)\}, X; \beta]$. It worths noting that there are two recursive relationships: one is in terms of the definition of f, and the other is in the definition of h. The second recursive relationship involving h is more difficult to handle, since to define $h(u, X)$, it requires evaluation of the unspecified baseline median or τth quantile residual life function $m(t)$ at the times resulting from the previous definition $h(u - 1, X)$. Such a recursively defined function f is formidable to invert in practice, hence a direct estimation procedure based on even one fixed special $S(t|X)$ is difficult to implement.

However, we observe that model (2) holds for all $t \geq 0$, thus at $t = 0$, the equality in fact reduces to $q_\tau(T|X) = g\{m(0), X; \beta\}$, or equivalently,

$$T = g\{m(0), X; \beta\} + \epsilon \quad \text{with } q_\tau(\epsilon|X) = 0.$$

Viewing $m(0)$ and β as the regression parameters and ϵ as the error, this model at $t = 0$ simplifies to a standard censored quantile regression problem. If we had fully observed all of the data $\{(T_i, X_i), i = 1, \ldots, n\}$, the standard quantile regression method could be directly applied (Koenker, 2005). Under random censoring, we can modify the estimation procedure by implementing the inverse probability weighting technique. Let $\hat{K}(y)$ be the Kaplan–Meier estimator of the survival function for the censoring time, based on the observations $\{(Y_i, 1 - \Delta_i), i = 1, \ldots, n\}$. Then, we obtain the estimators $\{\hat{m}(0), \hat{\beta}\}$ by minimizing

$$\sum_{i=1}^{n} \frac{\Delta_i \rho_\tau[Y_i - g\{m(0), X_i; \beta\}]}{\hat{K}(Y_i)}, \tag{4}$$

where $\rho_\tau(\cdot)$ is the usual "check function", defined as $\rho_\tau(u) = u\{I(u \geq 0) - (1-\tau)\}$. Here, to ensure the validity of the method, we make the standard assumptions that the survival times are continuous and bounded in $[0, L]$ for a fixed value L, and the censoring times are continuous and satisfy $\Pr(C_i \geq L) > 0$ for $i = 1, \ldots, n$. This permits our technical derivations, and also ensures that there can be patients observed throughout the study period.

Not only our method provides an estimator for the parameter of interest, β, but also it estimates the baseline median residual life function $m(t)$. Following this route, we can in fact estimate $m(t)$ at an arbitrarily chosen time t. In the situation with no censoring, we can simply collect the observations from $\{(T_i, X_i), i = 1, \ldots, n\}$ such that $T_i > t$, and estimate $\{m(t), \beta\}$ by minimizing $\sum_{i=1}^n I(T_i > t)\rho_\tau[T_i - t - g\{m(t), X_i; \beta\}]$. The selection of the subsample that satisfies $T_i > t$ would guarantee the conditional requirement specified in model (2). Under random censoring, by using the fact that $I(T_i > t)\Delta_i = I(Y_i > t)\Delta_i$, we would need to modify the estimator $\{\hat{m}(t), \hat{\beta}\}$ to the minimizer of

$$\sum_{i=1}^n \frac{I(Y_i > t)\Delta_i \rho_\tau[Y_i - t - g\{m(t), X_i; \beta\}]}{\hat{K}_t(Y_i)},$$

where $\hat{K}_t(y)$ is the Kaplan–Meier estimator of $K_t(y)$, the survival function for the censoring time in the subpopulation where the survival time of each individual is beyond t. However, a direct estimation of $K_t(y)$ is not available because for subjects censored before t, it is impossible to know if they would survive beyond t or not, hence it is impossible to form a random sample of the subpopulation. Fortunately, because of the independence between the event and the censoring process, $K_t(y) = \Pr(C \geq y | T > t) = \Pr(C \geq y)$, we can directly use the Kaplan–Meier estimator of the entire sample $\hat{K}(Y_i)$ to replace $\hat{K}_t(Y_i)$, and no new estimation for $K_t(Y_i)$ is needed.

The parameter β stays the same for the different values of t, and we can thus use the largest amount of information from the data to estimate β when $t = 0$. We propose a two-stage estimation procedure for $m(t)$: in the first stage, we estimate $m(0)$ and β at $t = 0$; and in the second stage, for other values of t, we fix the parameter β at the pre-estimated value $\hat{\beta}$ from stage 1, and estimate the corresponding $m(t)$. That is, we estimate $m(t)$ for each $t > 0$ by minimizing

$$\sum_{i=1}^n \frac{I(Y_i > t)\Delta_i \rho_\tau[Y_i - t - g\{m(t), X_i; \hat{\beta}\}]}{\hat{K}(Y_i)}$$

with respect to $m(t)$, while fixing $\hat{\beta}$ at the estimator obtained from (4) in stage 1.

The estimators of $\{m(0), \beta\}$ from (4) correspond to a standard M-estimation procedure based on the inverse probability weighting, thus it has the usual square root-n consistency and asymptotic normality. The same argument applies to the estimator of $m(t)$ for each specified value of $t > 0$. However, as t increases, fewer observations would contribute to the estimator, which is equivalent to reducing the effective sample size. Hence, the performance of $\hat{m}(t)$ deteriorates as t increases.

2.2. Asymptotic Properties

For a fixed time t, the proposed estimation procedure leads to a median regression model with randomly censored responses. Using similar arguments as those in Bang & Tsiatis (2000; 2002), we can obtain the asymptotic properties of the estimators.

For ease of exposition, at a fixed t, we define $\theta = \{m(t), \beta^{\mathrm{T}}\}^{\mathrm{T}}$, $g(X_i; \theta) = g\{m(t), X_i; \beta\}$,

$$s(X_i, Y_i, \hat{\theta}) = [I\{Y_i < t + g(X_i; \hat{\theta})\} - \tau]\frac{\partial g(X_i; \theta)}{\partial \theta}\Big|_{\theta=\hat{\theta}},$$

and assume that the first n_t survival times $(T_1, \ldots, T_{n_t})$ are larger than t. Then, the estimating equation takes the form of

$$\begin{aligned}\sum_{i=1}^{n} \frac{I(Y_i>t)\Delta_i}{\hat{K}(Y_i)} s(X_i, Y_i, \hat{\theta}) &= \sum_{i=1}^{n} \frac{I(T_i>t)\Delta_i}{\hat{K}(Y_i)} s(X_i, T_i, \hat{\theta}) \\ &= \sum_{i=1}^{n_t} \frac{\Delta_i}{\hat{K}(Y_i)} s(X_i, T_i, \hat{\theta}),\end{aligned}$$

which is $O_p(1)$ instead of 0 due to the non-differentiability of the check function. To further separate the effect of the Kaplan–Meier estimator of the censoring distribution, we write

$$\begin{aligned}O_p(1) &= \sum_{i=1}^{n_t} \frac{\Delta_i}{\hat{K}(Y_i)} s(X_i, T_i, \hat{\theta}) \\ &= \sum_{i=1}^{n_t} \frac{\Delta_i}{\hat{K}(Y_i)} s(X_i, T_i, \theta) + \sum_{i=1}^{n_t} \frac{\partial}{\partial \theta^{\mathrm{T}}} E\left\{ \frac{\Delta_i s(X_i, T_i, \theta^*)}{\hat{K}(Y_i)} \middle| T_i > t \right\} (\hat{\theta} - \theta) \\ &= \sum_{i=1}^{n_t} s(X_i, T_i, \theta) + \sum_{i=1}^{n_t} \left\{ \frac{\Delta_i}{K(Y_i)} - 1 \right\} s(X_i, T_i, \theta) + \sum_{i=1}^{n_t} \frac{\Delta_i \{K(Y_i) - \hat{K}(Y_i)\}}{K(Y_i)\hat{K}(Y_i)} s(X_i, T_i, \theta) \\ &\quad + n_t \left[\frac{\partial E\{s(X_i, T_i, \theta) | T_i > t\}}{\partial \theta^{\mathrm{T}}} + o_p(1) \right] (\hat{\theta} - \theta)\end{aligned} \tag{5}$$

where θ^* lies on the line segment between $\hat{\theta}$ and θ. To circumvent the non-differentiability of $s(\cdot)$, we take the derivative of its expectation, which is a typical technique used in quantile regression (see, e.g., Ying, Jung & Wei, 1995). In addition, because the sample contributed to the estimation is a random sample from the subpopulation with survival times larger than t, all the expectations are in fact conditional on $T_i > t$. The derivation in (5) is based on the understanding that $\hat{\theta}$ is a consistent estimator. Because the estimating equation is consistent under the true parameter values (Tsiatis, 2006), the consistency of $\hat{\theta}$ should hold intuitively, as inherently assumed in Bang & Tsiatis (2002).

Let L be the upper limit of all the event and censoring times, and define the filtration $\mathcal{F}_t(u)$ for u to be the set of σ-algebras generated by

$$\sigma\{I(C_i \le v), v \le u; I(t < T_i \le y), X_i, s(X_i, T_i, \theta), 0 \le y < L, i = 1, \ldots, n\}.$$

Let $N_i^c(u) = I(Y_i \le u, \Delta_i = 0)$ be the counting process for the censoring time, $R_i(u) = I(Y_i \ge u)$ be the risk process, and let $\lambda^c(u)$ be the hazard function for the censoring distribution. Furthermore, we consider the martingale $\mathcal{M}_i^c(u) = N_i^c(u) - \int_0^u \lambda^c(v) R_i(v)\, \mathrm{d}v$ and denote $\mathcal{M}^c(u) = \sum_{i=1}^n \mathcal{M}_i^c(u)$, $N^c(u) = \sum_{i=1}^n N_i^c(u)$ and $R(u) = \sum_{i=1}^n R_i(u)$. Note that $R(u) = n\hat{K}(u^-)\hat{S}(u^-)$, where $\hat{K}(u^-)$ is the left continuous version of the Kaplan–Meier estimator for the survival function of the censoring times $K(u)$, and $\hat{S}(u)$ is the Kaplan–Meier estimator for the survival function $S(u) = \Pr(T \ge u)$. Then, using a martingale integral representation (Gill, 1980, p. 37), we have that

$$\begin{aligned}\frac{K(v) - \hat{K}(v)}{K(v)} &= \int_0^v \frac{\hat{K}(u^-)}{K(u)} \frac{\mathrm{d}\mathcal{M}^c(u)}{R(u)} \\ &= \int_0^L \frac{I(v \ge u)\hat{K}(u^-)}{K(u)} \frac{\mathrm{d}\mathcal{M}^c(u)}{R(u)} \\ &= \int_0^L \frac{I(v \ge u)}{nK(u)\hat{S}(u^-)}\, \mathrm{d}\mathcal{M}^c(u).\end{aligned}$$

Define

$$\hat{F}_t(s,u) = \frac{1}{n_t \hat{S}(u^-)} \sum_{i=1}^{n_t} \frac{\Delta_i s(X_i,T_i,\hat{\theta}) I(Y_i \geq u)}{\hat{K}(Y_i)} = \frac{1}{n_t \hat{S}(u^-)} \sum_{i=1}^{n_t} \frac{\Delta_i s(X_i,T_i,\hat{\theta}) I(T_i \geq u)}{\hat{K}(Y_i)},$$

and denote $\tilde{F}_t(s,u)$ to be the same as $\hat{F}_t(s,u)$ except that $\tilde{F}_t(s,u)$ is evaluated at θ. Thus, we obtain

$$\begin{aligned} n_t{}^{-1/2} \sum_{i=1}^{n_t} \frac{\Delta_i\{K(Y_i)-\hat{K}(Y_i)\}}{K(Y_i)\hat{K}(Y_i)} s(X_i,T_i,\theta) &= n_t{}^{-1/2} \sum_{i=1}^{n_t} \frac{\Delta_i s(X_i,T_i,\theta)}{\hat{K}(Y_i)} \int_0^L \frac{I(Y_i \geq u)}{nK(u)\hat{S}(u^-)} \, \mathrm{d}\mathcal{M}^c(u) \\ &= n_t{}^{-1/2} \int_0^L \frac{n_t \tilde{F}_t(s,u)}{nK(u)} \, \mathrm{d}\mathcal{M}^c(u) \\ &= n_t{}^{-1/2} \sum_{i=1}^{n} \int_0^L \frac{n_t F_t(s,u)}{nK(u)} \, \mathrm{d}\mathcal{M}_i{}^c(u) + o_p(1), \end{aligned}$$

where $F_t(s,u) = E\{s(X_i,T_i,\theta) I(T_i \geq u) | T_i > t\} / S(u)$. Using the property similar to Robins and Rotnitzky (1992, p. 313) that

$$\frac{\Delta_i}{K(Y_i)} = 1 - \int_0^L \frac{\mathrm{d}\mathcal{M}_i^c(u)}{K(u)},$$

we obtain

$$n_t{}^{-1/2} \sum_{i=1}^{n_t} \left\{ \frac{\Delta_i}{K(Y_i)} - 1 \right\} s(X_i,T_i,\theta) = -n_t{}^{-1/2} \sum_{i=1}^{n_t} \int_0^L \frac{s(X_i,T_i,\theta)}{K(u)} \, \mathrm{d}\mathcal{M}_i^c(u).$$

Inserting these relationships into (5), we obtain

$$\begin{aligned} &- \left[\frac{\partial E\{s(X_i,T_i,\theta)|T_i>t\}}{\partial \theta^{\mathrm{T}}} + o_p(1) \right] \left\{ n_t{}^{1/2}(\hat{\theta}-\theta) \right\} \\ &= n_t{}^{-1/2} \sum_{i=1}^{n_t} s(X_i,T_i,\theta) - n_t{}^{-1/2} \sum_{i=1}^{n_t} \int_0^L \frac{s(X_i,T_i,\theta)}{K(u)} \, \mathrm{d}\mathcal{M}_i^c(u) \\ &+ \frac{n_t^{1/2}}{n} \sum_{i=1}^{n} \int_0^L \frac{F_t(s,u)}{K(u)} \, \mathrm{d}\mathcal{M}_i^c(u) + o_p(1). \end{aligned} \tag{6}$$

Because $s(X_i,T_i,\theta)$ is $\mathcal{F}_t(0)$ measurable, the first term and the remaining two terms on the right-hand side of equation (6) are uncorrelated. Thus, we have that $n_t{}^{1/2}(\hat{\theta}-\theta) \sim N(0, A^{-1}B(A^{-1})^{\mathrm{T}})$, where

$$\begin{aligned} A &= -\frac{\partial E\{s(X_i,T_i,\theta)|T_i>t\}}{\partial \theta^{\mathrm{T}}} \\ &= E\left[\frac{\partial g(X_i;\theta)}{\partial \theta} \left\{ \frac{\partial g(X_i;\theta)}{\partial \theta} \right\}^{\mathrm{T}} f_\epsilon(0|X_i,T_i>t) \right] \\ B &= E\{s(X_i,T_i,\theta)^{\otimes 2}|T_i>t\} + E\left\{ \int_0^L \frac{s(X_i,T_i,\theta)^{\otimes 2}}{K(u)^2} \lambda^c(u) R_i(u) \, \mathrm{d}u \middle| T_i > t \right\} \\ &+ \frac{n_t}{n} E\left\{ \int_0^L \frac{F_t(s,u)^{\otimes 2}}{K(u)^2} \lambda^c(u) R_i(u) \, \mathrm{d}u \right\} \\ &- \frac{2n_t}{n} E\left\{ \int_0^L \frac{s(X_i,T_i,\theta)F_t(s,u)^{\mathrm{T}}}{K(u)^2} \lambda^c(u) R_i(u) \, \mathrm{d}u \middle| T_i > t \right\}, \end{aligned}$$

where $f_\epsilon(\cdot|X, T > t)$ is the probability density function of $T - t - g(X;\theta)$ conditional on $T > t$, and $a^{\otimes 2} = aa^{\mathrm{T}}$ for an arbitrary vector a.

The last three terms in B are obtained through the martingale central limit theorem (Fleming & Harrington, 1991). When $n_t = o(n)$, B simplifies to

$$E\{s(X_i, T_i, \theta)^{\otimes 2}|T_i > t\} + E\left\{\int_0^L \frac{s(X_i, T_i, \theta)^{\otimes 2}}{K(u)^2}\lambda^c(u)R_i(u)\,\mathrm{d}u|T_i > t\right\}$$

to the first order, which indicates that the estimation of $K(u)$ does not have any effect on the asymptotic efficiency compared to knowing the true $K(u)$. It is generally known that in the inverse probability weighting procedure, using an estimated weight outperforms using a known weight in terms of variance estimation (Masayuki & Shinto, 2004). Our discovery finds that such improvement can occur only if the weights are estimated at the same rate as the parameters of interest. This agrees with the intuition that since $\hat{K}(u)$ has a root-n convergence rate, its estimation variance and its contribution to the final estimation of θ is ignorable compared to the root-n_t rate for $\hat{\theta}$ as a consequence of using a subsample of size n_t. Moreover, the asymptotic result is under the condition $n_t \to \infty$. In practice, this implies that t should be reasonably small to ensure that a sufficient amount of observations belong to the subsample and participate in the estimation procedure. Note that here, n_t is the number of subjects who would survive beyond time t, $n_t \geq \sum_{i=1}^n I(Y_i > t)$. Thus, as long as we choose time t such that $\sum_{i=1}^n I(Y_i > t)$ is sufficiently large, n_t would also be sufficiently large. Based on a subsample, the estimators $\hat{\beta}$ and $\hat{m}(t)$ are both root-n_t consistent and asymptotically normal. When estimating the parameter of interest β, we choose $t = 0$ to carry out the estimation so as to obtain the standard root-n rate. We do not average the $\hat{\beta}$'s obtained at different values of t, because these $\hat{\beta}$'s are estimated from a sequence of nested subsamples and thus they are highly correlated. At $t = 0$, we estimate β based on the maximum amount of data information, and as t increases, the available data information shrinks. Therefore, averaging the $\hat{\beta}$'s at different values of t would cause a deterioration in the estimation efficiency. Whereas, the convergence rate of $\hat{m}(t)$ is root-n_t depending on t because $m(t)$ is estimated using the data truncated by t. When letting $t = 0$, we can obtain the classical root-n asymptotic property of θ with $m(0)$.

The matrix A can be estimated by using its corresponding sample version of the expectation and plugging in the estimated parameter values,

$$\hat{A} = -n_t^{-1}\frac{\partial}{\partial\theta^{\mathrm{T}}}\left\{\sum_{i=1}^n \frac{I(Y_i > t)\Delta_i s(X_i, Y_i, \hat{\theta})}{\hat{K}(Y_i)}\right\},$$

where the partial derivative with respect to θ can be replaced with a numerical derivative if the closed form solution is difficult to obtain. The matrix B has four terms, sequentially denoted as $B_j,\ j = 1, \ldots, 4$, which can be approximated by using their respective empirical counterparts,

$$\hat{B}_1 = n_t^{-1}\sum_{i=1}^n \frac{s(X_i, Y_i, \hat{\theta})^{\otimes 2} I(Y_i > t)\Delta_i}{\hat{K}(Y_i)}$$

$$\hat{B}_2 = n^{-1}\sum_{i=1}^n \frac{\hat{F}_t(s^{\otimes 2}, Y_i)(1-\Delta_i)}{\hat{K}(Y_i)^2}$$

$$\hat{B}_3 = \frac{n_t}{n^2}\sum_{i=1}^n \frac{\hat{F}_t(s, Y_i)^{\otimes 2}(1-\Delta_i)}{\hat{K}(Y_i)^2}$$

$$\hat{B}_4 = -\frac{2n_t}{n^2}\sum_{i=1}^n \frac{\hat{F}_t\{s\hat{F}_t^{\mathrm{T}}(s,\cdot), Y_i\}(1-\Delta_i)}{\hat{K}(Y_i)^2}.$$

As an illustration, we write out the details for the estimation of B_2, and other terms can be derived similarly,

$$\begin{aligned} B_2 &= E\left\{\int_0^L \frac{s(X_i,T_i,\theta)^{\otimes 2}}{K(u)^2}\lambda^c(u)R_i(u)\,\mathrm{d}u \,\middle|\, T_i > t\right\} \\ &= E\left[\int_0^L \frac{E\left\{s(X_i,T_i,\theta)^{\otimes 2} I(T_i>u)|T_i>t\right\} I(C_i>u)}{K(u)^2}\lambda^c(u)\,\mathrm{d}u\right] \\ &= E\left\{\int_0^L \frac{F_t(s^{\otimes 2},u)S(u)I(C_i>u)}{K(u)^2}\lambda^c(u)\,\mathrm{d}u\right\} \\ &= E\left\{\int_0^L \frac{F_t(s^{\otimes 2},u)}{K(u)^2}R_i(u)\lambda^c(u)\,\mathrm{d}u\right\}. \end{aligned}$$

When the sample size is small, the approximation may be unstable, therefore bootstrap is typically implemented as in the usual quantile regression (Koenker, 2005).

3. SIMULATION STUDY

We examined four different median residual life models in the simulation study. We first considered a simple linear additive median residual life model

$$g\{m(t), X; \beta\} = m(t) + \beta_1 X_1 + \beta_2 X_2, \tag{7}$$

where the baseline residual life function was a constant $m(t) = 1$ and the $g(\cdot)$ function had a linear additive form. We generated covariate X_1 as a binary variable taking a value of 1 or 0 with probability 1/2, and X_2 from a uniform distribution on [0, 1]. Although there were many survival functions corresponding to model (7), we chose to generate the failure times from the survival function

$$S(t|X) = \exp\left\{\frac{\log(1-\tau)t}{m(0) + \beta_1 X_1 + \beta_2 X_2}\right\},$$

with the true parameter values $\beta_1 = 0.7$, $\beta_2 = -1$ and $m(0) = 1$. We simulated the censoring times from uniform distributions, so that the censoring rate was approximately 10% or 20%.

To examine the scenario in which the baseline median residual life function $m(t)$ increases as t increases, in the second simulation study, we set the true function $m(t) = c + \alpha t$ in model (7). Corresponding to a median residual life function of this form, we chose to generate the failure times from the survival function

$$S(t|X) = \exp\left\{\log\left(1 + \frac{\alpha t}{c + \beta_1 X_1 + \beta_2 X_2}\right)\frac{\log(1-\tau)}{\log(1+\alpha)}\right\},$$

with $c = 1$, $\alpha = 0.1$, $\beta_1 = 0.7$, $\beta_2 = -1$, and $\tau = 0.5$.

In the third numerical example, we investigated a proportional residual life structure. More specifically, the median residual life function had the form

$$g\{m(t), X; \beta\} = m(t)\exp(-\beta_1 X_1 - \beta_2 X_2),$$

and we set $m(t) = -\log(1-\tau)/\lambda$, which equals $\log(2)$ when $\lambda = 1$ and the median ($\tau = 0.5$) is considered. One survival function that would yield such a median residual life function was the

exponential model,

$$S(t|X) = \exp(-\lambda t e^{\beta_1 X_1 + \beta_2 X_2}),$$

which we used to simulate the failure times. We took the true parameter values $\lambda = 1$, $\beta_1 = 1$ and $\beta_2 = -1$. The covariates X_1, X_2 and the censoring times were simulated in the same way as those in model (7).

Finally, we studied a more complicated model given by

$$g\{m(t), X; \beta\} = (1 + \beta_1 X_1 + \beta_2 X_2)\log(1-\tau)\{(1 + \beta_1 X_1 + \beta_2 X_2)\log(1-\tau) - 2m(t)\},$$

with $m(t) = \sqrt{t}$. It can be verified that the data generated from the survival function

$$S(t|X) = \exp\left(\frac{-\sqrt{t}}{1 + \beta_1 X_1 + \beta_2 X_2}\right)$$

yield such a median residual structure. We set the true parameters values to be $\beta_1 = -0.5$, $\beta_2 = 1$, and generated the covariates in the same way as before. We simulated the censoring times from a mixture of an exponential distribution and a point mass at infinity to achieve the target censoring rates.

We replicated 1,000 data sets with a sample size of 200. For each data realization, we computed the parameter estimates for β_1, β_2 and $m(o)$, and estimated the corresponding variances based on the usual bootstrap procedure to avoid the nonparametric functional estimation of the error density function.

In Table 1, we present the average of the parameter estimates over 1,000 simulations, the empirical variances, the median variance estimates using the bootstrap procedure, and the coverage probabilities of the 95% confidence intervals. For each replicated data set, we took 500 bootstrap samples for variance estimation.

The simulation results show that the parameter estimates are close to the true values, and the biases increase as the censoring rates increase. The estimated variances provide a good approximation of the variation in the parameter estimates, and the 95% coverage probabilities are close to the nominal level. Thus, the proposed estimation procedure performed reasonably well for the finite sample size.

In the second simulation study, if we had known that the baseline function $m(t)$ was of a linear form, it would be interesting to also estimate the slope α. A straightforward approach is to obtain $\hat{m}(t)$ at different values of t, and then carry out a simple linear regression. Using this approach, our estimate for the slope was 0.0902, with an empirical 95% confidence interval $[-0.1216, 0.4437]$ for the case with 10% censoring. Further discussions on estimation with a known parametric or smooth function $m(t)$ are given in Section 5.

4. EXAMPLE

As an illustration, we applied the proposed median residual life model to a data set from a leukemia study conducted at M.D. Anderson Cancer Center (Tsimberidou et al., 2006). In that study, 130 patients were diagnosed with Richter's syndrome (RS) via biopsy or fine-needle aspiration. RS is a rare and aggressive type of acute adult leukemia that often results from a transformation of chronic lymphocytic leukemia into diffuse large cell lymphoma. RS is usually fatal within a short period of time. The patients in the study were treated by either chemoimmunotherapy with rituximab or chemotherapy alone. Figure 1 shows the Kaplan–Meier survival curves stratified by treatment groups. We can see that there is a sharp change point in the survival curve around 2 years of

TABLE 1: Simulation studies under four different median residual life models with a sample size of 200.

Median residual life model	Estimate	10% censoring			20% censoring		
		β_1	β_2	$m(0)$	β_1	β_2	$m(0)$
Additive ($m(t) = 1$)	True value	0.7	−1	1	0.7	−1	1
	Est	0.6900	−0.9841	0.9923	0.6529	−0.9445	0.9626
	var	0.0287	0.0476	0.0330	0.0304	0.0526	0.0355
	$\widehat{\text{var}}$	0.0312	0.0547	0.0358	0.0338	0.0613	0.0386
	CP (%)	94.8	95.5	94.0	94.4	95.2	93.9
Additive ($m(t) = 1 + 0.1t$)	True value	0.7	−1	1	0.7	−1	1
	Est	0.6861	−0.9772	0.9879	0.6320	−0.9210	0.9455
	var	0.0319	0.0525	0.0364	0.0342	0.0588	0.0396
	$\widehat{\text{var}}$	0.0339	0.0605	0.0397	0.0374	0.0674	0.0427
	CP (%)	94.6	95.7	94.0	92.8	95.1	93.8
Proportional ($m(t) = \log(2)$)	True value	1	−1	0.6931	1	−1	0.6931
	Est	1.0004	−0.9759	0.7102	0.9418	−0.9082	0.7055
	var	0.0562	0.2577	0.0437	0.0647	0.2691	0.0494
	$\widehat{\text{var}}$	0.0613	0.2298	0.0399	0.0716	0.2648	0.0449
	CP (%)	95.4	95.3	94.7	94.9	94.7	92.6
Nonlinear ($m(t) = \sqrt{t}$)	True value	−0.5	1	0	−0.5	1	0
	Est	−0.5047	1.0454	0.0145	−0.5145	1.0457	0.0216
	var	0.1063	0.2850	0.0446	0.1294	0.3439	0.0520
	$\widehat{\text{var}}$	0.1100	0.3079	0.0483	0.1362	0.3847	0.0590
	CP (%)	96.2	95.6	93.6	95.6	96.3	95.2

Est is the averaged parameter estimate over 1,000 simulations; var represents the sample empirical variance; $\widehat{\text{var}}$ is the median variance estimate using the bootstrap procedure; and CP (%) is the coverage probability of the 95% bootstrap confidence interval.

follow-up, which would typically cause a violation of the usual proportional hazards assumption. The censoring rate of the data was approximately 12%. Three covariates were included in our analysis: treatment (1 if chemotherapy alone, and 0 if chemoimmunotherapy with rituximab), patient age (ranging from 29 to 77 years with a median of 60 years), and patient sex (1 if male, and 0 if female).

We fitted two of the proposed median residual life models including the linear additive and the proportional model structures, as both were reasonable candidate models based on an initial inspection of the scatter plots. In addition to the median, we also examined the quantile levels at $\tau = 25\%$ and 75% for each model. As shown in Table 2, under the additive quantile residual life model, we found no difference in treatment at any of the three quantiles, with P-values of 0.437, 0.889 and 0.267, for $\tau = 25\%, 50\%$ and 75%, respectively. Under the proportional quantile residual life model, the treatment effect was not significant at the $\tau = 25\%$ and 50% quantile level, with P-values of 0.381 and 0.972, respectively. However, at $\tau = 75\%$, the treatment exhibited a significant difference with a P-value of 0.025, indicating that patients treated with chemoimmunotherapy plus rituximab had a longer residual lifetime than those treated with chemotherapy alone. This

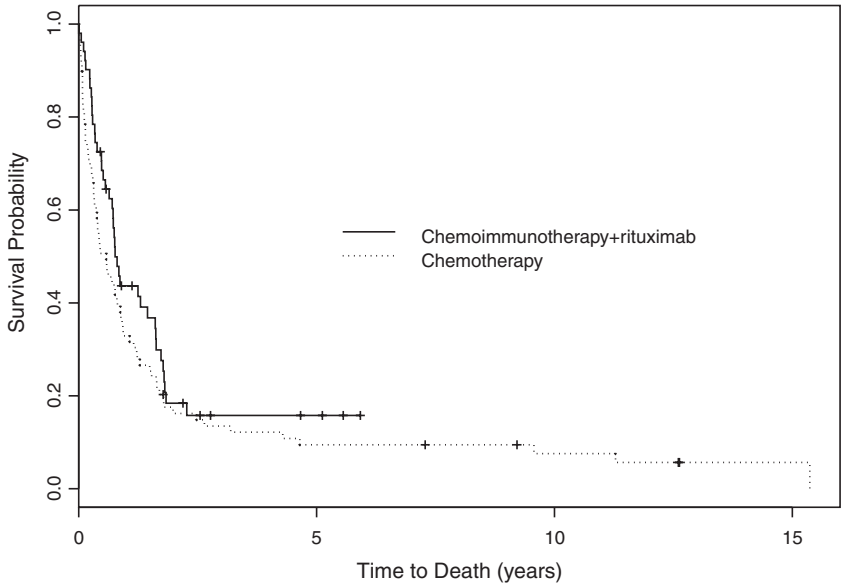

FIGURE 1: Estimated Kaplan–Meier survival curves stratified by treatment for the Richter's syndrome data.

finding can also be empirically observed in Figure 1, which shows a spread in the survival curves between 1 and 2 years of follow-up. In both models at all of the three quantile levels, we did not find patients' age or sex to have a significant effect on the residual lifetime. To determine which model structure was more suitable for the RS data, we carried out an ad-hoc model selection procedure based on the optimization value of the objective function given in (4). We found that the additive and proportional models yielded very similar minima for $\tau = 25\%$ and 50%. When $\tau = 25\%$, the objective function in (4) took a value of 0.5347 under the additive model, and 0.5341 under the proportional model; and when $\tau = 50\%$, it took values of 0.9882 and 0.9883, respectively. Hence, there did not seem to be any significant favouring of one model over the other. However, for $\tau = 75\%$, the objective function of the proportional model took a value of 1.0032, while that of the additive model was 1.2397. Therefore, the proportional model appeared to outperform the additive model in terms of model fitting.

For the residual lifetime, one could also consider the mean instead of the median. Thus, we fitted the mean residual life models under the additive and proportional structures, respectively. In Table 2, we can see that the mean residual life model yielded quite different results from those based on quantile residual life models. In particular, under the additive model, the mean residual life model produced a statistically significant treatment difference, while none of the quantile residual life models demonstrated this treatment difference. We expect that the mean residual life model might not fit the data well because it focuses on the central covariate effects, while Figure 1 shows that the survival curves are very irregular and several change or crossing points exist in the survival curves. This may violate the assumptions for any model that tries to characterize the central effects of covariates for the entire survival curve. In contrast, quantile residual life models are more focused on local covariate effects centering around certain conditional quantiles, which can particularly characterize the local features of the survival curve without imposing any requirement for the entire curve. In additional to robustness, the quantile residual life models based on different values of τ provide a comprehensive evaluation of the covariate effects.

TABLE 2: Analysis of the Richter's syndrome data using the quantile and mean residual life models.

	Additive model				Proportional model			
	Treat	Sex	Age	$m(0)$	Treat	Sex	Age	$m(0)$
Quantile residual life model, $\tau = 0.25$								
Est	−0.0861	−0.0212	−0.0476	0.3127	0.4391	0.1549	0.2066	0.3398
SE	0.1108	0.1162	0.0581	0.1245	0.5008	0.5371	0.3702	0.1552
P-value	0.4372	0.8552	0.4128	0.0120	0.3806	0.7730	0.5767	0.0286
Quantile residual life model, $\tau = 0.5$								
Est	0.0329	−0.1897	0.0315	0.8136	−0.0351	0.2604	−0.0396	0.8138
SE	0.2358	0.3621	0.1213	0.3127	0.9849	1.1687	0.3258	0.2904
P-value	0.8890	0.6002	0.7948	0.0093	0.9716	0.8237	0.9034	0.0051
Quantile residual life model, $\tau = 0.75$								
Est	0.9675	−1.3293	0.1000	2.0354	−2.5895	2.2155	0.3198	1.1793
SE	0.8715	4.8752	0.3018	4.7387	1.1575	1.1776	0.2897	0.4031
P-value	0.2669	0.7851	0.7402	0.6675	0.0253	0.0599	0.2696	0.0034
Mean residual life model								
Est	2.3082	−2.7257	−0.2931	2.3497	−1.6999	1.3340	0.1271	1.0207
SE	0.9738	1.5796	0.3309	0.9715	1.3564	2.3524	0.5530	0.4523
P-value	0.0178	0.0844	0.3758	0.0156	0.2101	0.5707	0.8183	0.0240

5. DISCUSSION

We have proposed a semiparametric median residual life model and a simple estimation procedure. The estimation method is quite general and can be applied to other residual life models, for example, the mean residual life model. The estimation procedure is also applicable to the varying-coefficient model

$$q_\tau(T - t|T > t, X) = g\{m(t), X; \beta(t)\},$$

where at each time t, we estimate the fixed unknown parameters $\{m(t), \beta(t)\}$. This makes the varying-coefficient residual life model attractive, as it is more flexible yet does not require extra computational complexity. Model selection in quantile regression has received much attention (Koenker & Machado, 1999; and He & Zhu, 2003). In selecting the best model among a set of candidate models, we can simply use the model selection criterion similar to the AIC or BIC. It would be interesting to further develop model selection procedures in quantile residual life models.

If we assume additionally that $m(t)$ is smooth or has certain parametric form as in the second simulation example, we could use the estimated $m(t)$ at different values of t as pseudo observations and then smooth them to obtain a nonparametric estimation of $m(t)$ or directly fit the parametric model. Although it is conceptually straightforward, the smoothing here is vastly different from the traditional nonparametric curve fitting due to the finite sample bias, the inherent correlation among these pseudo observations and their heteroscedasticity. In addition, by taking into account the smoothness of $m(t)$, the estimation of $m(t)$ at each fixed time t could be carried out differently, which may warrant further investigation.

The non-uniqueness of the survival function for a given median residual life function makes the estimation and inference very different from those of other regression models. Multiple survival functions correspond to the same median residual life function, which, however, further enhances the robustness of the proposed method.

ACKNOWLEDGEMENTS

Ma's research was supported by a US NSF grant. The authors are grateful to the Associate Editor, referee and Paul Gustafson for their critical and insightful comments, which led to great improvements in the revised manuscript.

BIBLIOGRAPHY

H. Bang & A. A. Tsiatis (2000). Estimating medical costs with censored data. *Biometrika*, 87, 329–343.

H. Bang & A. A. Tsiatis (2002). Median regression with censored cost data. *Biometrics*, 58, 643–649.

Y. Q. Chen (2007). Additive expectancy regression. *Journal of the American Statistical Association*, 102, 153–166.

Y. Q. Chen & S. C. Cheng (2005). Semiparametric regression analysis of mean residual life with censored survival data. *Biometrika*, 92, 19–29.

Y. Q. Chen & S. C. Cheng (2006). Linear life expectancy regression with censored data. *Biometrika*, 93, 303–313.

Y. Q. Chen, N. P. Jewell & S. C. Cheng (2005). Semiparametric estimation of proportional mean residual life model in presence of censoring. *Biometrics*, 61, 170–178.

D. R. Cox (1972). Regression models & life tables (with discussion). *Journal of the Royal Statistical Society Series B*, 34, 187–220.

T. R. Fleming & D. Harrington (1991). "*Counting Processes and Survival Analysis*." Wiley, New York.

A. E. Gelfand & A. Kottas (2003). Bayesian semiparametric regression for median residual life. *Scandinavian Journal of Statistics*, 30, 651–665.

R. D. Gill (1980). *Censoring and Stochastic Integrals.* Mathematical Centre Tracts 124. Amsterdam: Mathematisch Centrum.

R. C. Gupta & E. S. Langford (1984). On the determination of a distribution by its median residual life function: a functional equation. *Journal of Applied Probability*, 21, 120–128.

W. J. Hall & J. A. Wellner (1984). Mean residual life. In *Proceedings of the International Symposium on Statistics and Related Topics*, Eds: M. Csörgő, D. A. Dawson, J. N. K. Rao & A. K. Md. E. Saleh. North Holland, Amsterdam, pp. 169–184.

X. He & L.-X. Zhu (2003). A Lack-of-Fit Test for Quantile Regression. *Journal of the American Statistical Association*, 98, 1013–1022.

J. H. Jeong, S. H. Jung & J. P. Costantino (2008). Nonparametric inference on median residual life function. *Biometrics*, 64, 157–163.

S. H. Jung, J. H. Jeong & H. Bandos (2009). Regression on quantile residual life. *Biometrics*, 65, 1203–1212.

R. Koenker (2005). "*Quantile Regression*." Cambridge University Press, New York.

R. Koenker & G. J. Bassett (1978). Regression quantiles. *Econometrica*, 46, 33–50.

R. Koenker & J. A. F. Machado (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94, 1296–1310.

G. Maguluri & C.-H. Zhang (1994). Estimation in the mean residual life regression model. *Journal of the Royal Statistical Society Series B*, 56, 477–489.

H. Masayuki & E. Shinto (2004). A paradox concerning nuisance parameters and projected estimating functions. *Biometrika*, 91, 929–941.

D. Oakes & T. Dasu (1990). A note on residual life. *Biometrika*, 77, 409–410.

D. Oakes & T. Dasu (2003). Inference for the proportional mean residual life model. In *Crossing Boundaries: Statistical Essays in Honor of Jack Hall*, Institute of Mathematical Statistics Lecture Notes Monograph Series, 43, Eds: J. E. Kolassa & D. Oakes. Institute of Mathematical Statistics, Hayward, CA, 105–116.

L. Peng & Y. Huang (2008). Survival analysis with quantile regression models. *Journal of the American Statistical Association*, 103, 637–649.

S. Portnoy (2003). Censored regression quantiles. *Journal of the American Statistical Association*, 98, 1001–1012.

J. L. Powell (1984). Least absolute deviations estimation for the censored regression. *Journal of Econometrics*, 25, 303–325.

J. M. Robins & A. Rotnitzky (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. *AIDS Epidemiology - Methodological Issues.* Eds: Jewell N., Dietz K., Farewell V. Birkhäser, Boston, MA, pp. 297–331.

A.A. Tsiatis (2006). *Semiparametric Theory and Missing Data*. Springer, New York.

A. M. Tsimberidou, S. O'Brien, I. Khouri, F. G. Giles, H. M. Kantarjian, R. Champlin, S. Wen, K.-A. Do, S. C. Smith, S. Lerner, E. J. Freireich & M. J. Keating (2006). Clinical outcomes and prognostic factors in patients with Richter's syndrome treated with chemotherapy or chemoimmunotherapy with or without stem-cell transplantation. *Journal of Clinical Oncology*, 24, 2343–2351.

S. Yang (1999). Censored median regression using weighted empirical survival and hazard function. *Journal of the American Statistical Association*, 94, 137–145.

Z. Ying, S. H. Jung & L.J. Wei (1995). Survival Analysis with median regression models. *Journal of the American Statistical Association*, 90, 178–184.

K. Yu, Z. Lu & J. Stander (2003). Quantile regression: Applications and current research areas. *Journal of the Royal Statistical Society Series D*, 52, 331–350.

Received 1 October 2009
Accepted 7 May 2010