

Regression on Quantile Residual Life

Author(s): Sin-Ho Jung, Jong-Hyeon Jeong and Hanna Bandos

Source: *Biometrics*, Vol. 65, No. 4 (Dec., 2009), pp. 1203-1212

Published by: International Biometric Society

Stable URL: <http://www.jstor.org/stable/20640641>

Accessed: 13-03-2017 00:29 UTC

REFERENCES

Linked references are available on JSTOR for this article:

http://www.jstor.org/stable/20640641?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at

<http://about.jstor.org/terms>



International Biometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Biometrics*

Regression on Quantile Residual Life

Sin-Ho Jung,¹ Jong-Hyeon Jeong,^{2,*} and Hanna Bandos²

¹Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina 27710, U.S.A.

²Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, U.S.A.

*email: jeong@nsabp.pitt.edu

SUMMARY. A time-specific log-linear regression method on quantile residual lifetime is proposed. Under the proposed regression model, any quantile of a time-to-event distribution among survivors beyond a certain time point is associated with selected covariates under right censoring. Consistency and asymptotic normality of the regression estimator are established. An asymptotic test statistic is proposed to evaluate the covariate effects on the quantile residual lifetimes at a specific time point. Evaluation of the test statistic does not require estimation of the variance-covariance matrix of the regression estimators, which involves the probability density function of the survival distribution with censoring. Simulation studies are performed to assess finite sample properties of the regression parameter estimator and test statistic. The new regression method is applied to a breast cancer data set with long-term follow-up to estimate the patients' *median* residual lifetimes, adjusting for important prognostic factors.

KEY WORDS: Breast cancer; Martingale; Minimum dispersion statistic.

1. Introduction

In medical research, an investigator's ultimate interest may be in inferring the remaining life years of a patient, given his/her genetic and/or environmental backgrounds. The remaining life years of a patient could be prolonged by treating or preventing a disease by a medical intervention. Existing statistical methods such as an adjusted version of the Kaplan–Meier method (Kaplan and Meier, 1958) or Cox proportional hazards model (Cox, 1972, 1975) may be adopted to indirectly infer the remaining lifetimes, but they are often cumbersome and not straightforward, especially when the remaining lifetimes need to be evaluated in the middle of an observation period. Also the remaining lifetimes at a *specific* time point estimated from the semiparametric models such as proportional hazards or accelerated failure time (AFT) (Cox and Oakes, 1984) may heavily depend on model assumptions that affect the entire observation period. Another advantage of using the concept of the remaining lifetime is that it may provide investigators and patients with more straightforward interpretations regarding the efficacy results from clinical studies, so that they can communicate more clearly about potential benefits of new drugs for ongoing or future trials.

Some recent phase III breast cancer clinical trials provide practical examples of studies that could take advantage of the proposed method (Goss et al., 2003; Coombes et al., 2004; Mamounas et al., 2008). In those studies, estrogen receptor-positive patients who had been treated with tamoxifen for up to 5 years without recurrence of the original disease were re-randomized to placebo or an aromatase inhibitor, either letrozole or exemestane. A patient who would be interested in participating in this type of study would want to know how much her lifetime could be extended by accepting the new therapy after she had already survived 5 years since her initial

diagnosis and treatment. In this case, the important covariates to be included in a regression model would be the patient's age at diagnosis, seriousness of nodal involvement, status of estrogen or progesterone receptor, pathological tumor size, and type of the initial therapy.

Quantile regression models were originally introduced by Koenker and Bassett (1978). Recent developments in this area for *non-censored* data include Jung (1996), Portnoy and Koenker (1997), Wei et al. (2006), Whang (2006), and Mu and He (2007). Jung (1996) implemented the quasi-likelihood approach to obtain the parameter estimates for median regression models, which allows for the observations to be dependent. Portnoy and Koenker (1997) introduced a new computational technique that permits a faster estimation process for general l_1 -type problems over the widely used simplex method for sufficiently large sample sizes. Wei et al. (2006) applied the quantile regression approach to analyze longitudinal growth charts data. Whang (2006) proposed to use the empirical likelihood approach for parameter estimation and construction of the confidence regions. Mu and He (2007) developed the power-transformed linear quantile regression model by using the equivariance property of the quantiles under monotone transformations.

For *censored data*, although much work has been done on the quantile regression model at the *origin* of the time axis, not much attention has been given to the regression model on quantile *residual* life function. Gelfand and Kottas (2003) proposed a Bayesian inference procedure for a *median* residual life regression model that was derived by the AFT regression model. Their work relies on correct specification of the AFT model for the underlying failure time distribution. In this article, we develop a frequentist and time-specific regression method that can be used to directly infer the effect

of covariates on the quantile residual lifetimes at any specific time point, and that does not require specifying a semiparametric model for the underlying failure time. The proposed method can be viewed as a generalization of the median regression model (Ying, Jung, and Wei, 1995). The interpretation of the larger quantile residual lifetime in a treatment group at time t_0 can be viewed as a positive effect of treatment on the quantile failure time among subjects who survived beyond t_0 . In general, evaluation of the asymptotic variance of the quantiles in survival data is not straightforward because it involves estimation of the probability density function of failure time distribution under censoring. Here we propose a method that deals directly with the estimating equation to avoid the problem.

In Section 2, some notations are introduced along with the definition of the quantile residual life function. In Section 3, a time-specific quantile residual life regression model is proposed. An estimating equation is formulated to estimate the regression parameters, and consistency and asymptotic normality of the regression estimator are established. An asymptotic χ^2 -statistic is also proposed to test significance of the regression parameters in this section. In Section 4, simulation studies are performed to assess the small-sample properties of the regression method, such as bias of the estimator, and type I error and power of the test statistic. In Section 5, the proposed method is applied to a data set from one of the earliest clinical trials on breast cancer treatment performed by the National Surgical Adjuvant Breast and Bowel Project (NSABP). We conclude with discussions in Section 6.

2. Quantile Residual Life Function

Let $Y_i, i = 1, 2, \dots, n$, denote the minimum of the failure time T_i and censoring time C_i . Let Δ_i be the censoring indicator, i.e., $\Delta_i = 1$ if $Y_i = T_i$ and $\Delta_i = 0$ if $Y_i = C_i$. Suppose $\theta_{\xi|t}$ defines the ξ -quantile residual life function at time t . Then, $\theta_{\xi|t} = \xi$ -quantile $(T_i - t | T_i \geq t)$ satisfies the relation $P(T_i - t \geq \theta_{\xi|t} | T_i \geq t) = 1 - \xi$, which is equivalent to

$$P(T_i \geq t + \theta_{\xi|t}) = (1 - \xi)P(T_i \geq t). \quad (1)$$

Note that $\theta_{\xi|t}$ is the ξ -quantile of the remaining lifetimes among patients who survived beyond time t . The function $\theta_{\xi|t}$ has been extensively studied by many authors, including Schmittlein and Morrison (1981), Arnold and Brockett (1983), Csörgő and Horvath (1983), Gupta and Langford (1984), Joe and Proschan (1984), Csörgő and Csörgő (1987), Alam and Kulasekera (1993), Song and Cho (1995), among others. Gupta and Langford (1984) noted that $\theta_{\xi|t}$ does not uniquely determine $S(t) = P(T > t)$, but in practice one can model $S(t)$ first and then infer $\theta_{\xi|t}$ (Gelfand and Kottas, 2003). Jeong, Jung, and Costantino (2008) developed a nonparametric two-sample statistic to compare the median residual lifetimes at any fixed time point during the follow-up period of a study. In this article, we are particularly interested in associating the function $\theta_{\xi|t}$ at a specific time t_0 , denoted as $\theta_{\xi|t_0}$ in the sequel, with selected covariates.

Suppose that $\{(Y_i, \Delta_i), i = 1, \dots, n\}$ are random samples from a single population. Let $G(t) = P(C \geq t)$ denote the "survival function" of censoring distribution, and $I(\cdot)$ is the indicator function. Since $E\{I(Y \geq t)\} = S(t)G(t)$ under the independence of failure and censoring times, we can

consider an unbiased estimating function for $\theta_{\xi|t_0}$,

$$s_n(\theta_{\xi|t_0}) = \sum_{i=1}^n \left\{ \frac{I(Y_i \geq t_0 + \theta_{\xi|t_0})}{G(t_0 + \theta_{\xi|t_0})} - (1 - \xi) \frac{I(Y_i \geq t_0)}{G(t_0)} \right\}.$$

Here, $G(t)$ is usually unknown, so that we replace it with its Kaplan-Meier estimator (Kaplan and Meier, 1958), $\hat{G}(t)$ by using the censoring indicator $1 - \Delta_i$, to obtain a consistent estimating function under censoring

$$S_n(\theta_{\xi|t_0}) = \sum_{i=1}^n \left\{ \frac{I(Y_i \geq t_0 + \theta_{\xi|t_0})}{\hat{G}(t_0 + \theta_{\xi|t_0})} - (1 - \xi) \frac{I(Y_i \geq t_0)}{\hat{G}(t_0)} \right\}. \quad (2)$$

Consistency and asymptotic normality of the solution $\hat{\theta}_{\xi|t_0}$ to the equation can be easily developed.

In the following section, we extend the estimating equation for the quantile residual life for a single sample to a regression setting.

3. Regression Model

We consider a linear regression model for a ξ -quantile of residual lifetimes at time t_0 , on a log-scale,

$$\xi\text{-quantile}\{\log(T_i - t_0) | T_i \geq t_0, \mathbf{Z}_i\} = \beta'_{\xi|t_0} \mathbf{Z}_i, \quad (3)$$

where $\beta_{\xi|t_0} = (\beta_{\xi|t_0,0}, \beta_{\xi|t_0,1}, \dots, \beta_{\xi|t_0,p})'$ denotes a vector of the regression coefficients, and $\mathbf{Z}_i = (1, X_{1i}, \dots, X_{pi})'$ is a vector of covariates for a subject i . The model (3) specifies a linear relationship between the ξ -quantile residual lifetimes on a log-scale and the vector of covariates at a *specific* time t_0 , allowing for testing, say, a treatment effect adjusting for other covariates. Let $\beta_{\xi|t_0}^0$ denotes the true value of $\beta_{\xi|t_0}$. Without censoring the estimate of the true regression parameter can be obtained by minimizing $\sum_{i=1}^n |\log(T_i - t_0) - \beta'_{\xi|t_0} \mathbf{Z}_i| = 2 \sum_{i=1}^n \{[\log(T_i - t_0) - \beta'_{\xi|t_0} \mathbf{Z}_i][I\{\log(T_i - t_0) - \beta'_{\xi|t_0} \mathbf{Z}_i \geq 0\} - (1 - \xi)]\}$. Therefore the estimating function for noncensored case is given by $\sum_{i=1}^n \mathbf{Z}_i [I\{T_i \geq t_0 + \exp(\beta'_{\xi|t_0} \mathbf{Z}_i)\} - (1 - \xi)]$. Assuming conditional independence between T_i and C_i given \mathbf{Z}_i and independence of C_i from \mathbf{Z}_i , the following is true for the censored data,

$$\begin{aligned} E[I\{\log(Y_i - t_0) \geq \beta_{\xi|t_0}^0 \mathbf{Z}_i\} | \mathbf{Z}_i] \\ &= P\{Y_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i) | \mathbf{Z}_i\} \\ &= P\{T_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i) | \mathbf{Z}_i\} \\ &\quad \times P\{C_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i) | \mathbf{Z}_i\}, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \frac{P\{T_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i) | \mathbf{Z}_i\}}{P(T_i \geq t_0 | \mathbf{Z}_i)} \\ \times P(T_i \geq t_0 | \mathbf{Z}_i) P\{C_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i)\}. \end{aligned}$$

Under the model (3) and by definition of the ξ -quantile residual life function

$$\frac{P\{T_i \geq t_0 + \exp(\beta_{\xi|t_0}^0 \mathbf{Z}_i) | \mathbf{Z}_i\}}{P(T_i \geq t_0 | \mathbf{Z}_i)} = 1 - \xi.$$

Therefore,

$$\begin{aligned} E[I\{\log(Y_i - t_0) \geq \beta_{\xi|t_0}^{\prime} \mathbf{Z}_i\} | \mathbf{Z}_i] \\ = (1 - \xi) \frac{G\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{G(t_0)} P(Y_i \geq t_0 | \mathbf{Z}_i) \\ = E \left[(1 - \xi) \frac{G\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{G(t_0)} I(Y_i \geq t_0) \middle| \mathbf{Z}_i \right], \end{aligned}$$

and the following can be used as an estimating equation for the regression parameter $\beta_{\xi|t_0}$,

$$\begin{aligned} \mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0}) = \sum_{i=1}^n \mathbf{Z}_i \left[\frac{I\{Y_i \geq t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{\widehat{G}\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}} \right. \\ \left. - (1 - \xi) \frac{I(Y_i \geq t_0)}{\widehat{G}(t_0)} \right] \approx 0. \quad (4) \end{aligned}$$

Equation (4) mimics the least square estimating equations from the ordinary multiple linear regression model since

$$\begin{aligned} E \left[\frac{I\{Y_i \geq t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{G\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}} \middle| \mathbf{Z}_i \right] \\ = E \left\{ (1 - \xi) \frac{I(Y_i \geq t_0)}{G(t_0)} \middle| \mathbf{Z}_i \right\}. \end{aligned}$$

By the invariance property of the quantile with respect to monotone transformations, the estimating equation (4) can be evaluated on the original scale of the observed survival data, although the model (3) is based on a log-scale. This transformation invariance property of the quantile also allows for flexibility in data analysis. A solution $\widehat{\beta}_{\xi|t_0}$ to (4) may be defined as a minimizer of the function $\|\mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0})\|$, where $\|\cdot\|$ will be defined as the square root of sum of squares in our simulation study and real example. Under certain regularity conditions, it can be shown that $\widehat{\beta}_{\xi|t_0}$ is a consistent estimator of its true value $\beta_{\xi|t_0}^0$ (see Web Appendix A). Note that in practice if $\widehat{G}\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}$ and $\widehat{G}(t_0)$ are zeros in (4), $I\{Y_i \geq t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}/\widehat{G}\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}$ and $I(Y_i \geq t_0)/\widehat{G}(t_0)$ are also set to be zeros.

Suppose that we are interested in testing the null hypothesis $H_0: \beta_{\xi|t_0} = \beta_{\xi|t_0,0}$. Because the limiting variance-covariance matrix of the estimators $\widehat{\beta}_{\xi|t_0}$ involves unknown conditional probability density functions of $\log(T_i - t_0) - \beta_{\xi|t_0}^{\prime} \mathbf{Z}_i$ given \mathbf{Z}_i , we will use the estimating function $\mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0})$ directly to test H_0 . In Web Appendix B, we show that the distribution of $n^{-1/2} \mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0}^0)$ is approximately normal with mean zero and variance-covariance matrix $\mathbf{\Gamma}_{\xi|t_0} = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \boldsymbol{\tau}_{\xi|t_0,i} \boldsymbol{\tau}_{\xi|t_0,i}^{\prime}$, where

$$\begin{aligned} \boldsymbol{\tau}_{\xi|t_0,i} = \left[\frac{I\{Y_i \geq t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{G\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}} - (1 - \xi) \frac{I(Y_i \geq t_0)}{G(t_0)} \right] \mathbf{Z}_i \\ + \int_{-\infty}^{\infty} G^{-1}(s) \int_{-\infty}^s h^{-1}(v) \{dI(Y_i \leq v, \Delta_i = 0) \\ - I(Y_i \geq v) d\Lambda_G(v)\} d\mathbf{q}_1(s) \end{aligned}$$

$$\begin{aligned} - \mathbf{q}_2(t_0) \int_{-\infty}^{t_0} h^{-1}(s) \{dI(Y_i \leq s, \Delta_i = 0) \\ - I(Y_i \geq s) d\Lambda_G(s)\}, \end{aligned}$$

$\Lambda_G(\cdot)$ is the cumulative hazard function for the censoring distribution, $h(s) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n I(Y_i \geq s)$, $\mathbf{q}_1(s) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n \mathbf{Z}_i I\{t_0 + \exp(\beta_{\xi|t_0}^{\prime} \mathbf{Z}_i) \leq \min(s, Y_i)\}$, and

$$\mathbf{q}_2(t_0) = \lim_{n \rightarrow \infty} G(t_0)^{-1} \left(\frac{1 - \xi}{n} \right) \sum_{i=1}^n I(Y_i \geq t_0) \mathbf{Z}_i.$$

A consistent estimator $\widehat{\mathbf{\Gamma}}_{\xi|t_0}$ for the limiting covariance matrix of $n^{-1/2} \mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0}^0)$ can be obtained as (see Web Appendix B) $\widehat{\mathbf{\Gamma}}_{\xi|t_0} = n^{-1} \sum_{i=1}^n \widehat{\boldsymbol{\tau}}_{\xi|t_0,i} \widehat{\boldsymbol{\tau}}_{\xi|t_0,i}^{\prime}$, where

$$\begin{aligned} \widehat{\boldsymbol{\tau}}_{\xi|t_0,i} = \left[\frac{I\{Y_i \geq t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{\widehat{G}\{t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}} - (1 - \xi) \frac{I(Y_i \geq t_0)}{\widehat{G}(t_0)} \right] \mathbf{Z}_i \\ + \sum_{l=1}^n \left[\mathbf{Z}_l \frac{I\{t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_l) \leq Y_l\}}{\widehat{G}\{t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_l)\}} \right] \\ \times \left[\frac{(1 - \Delta_i) I\{Y_i \leq t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_i)\}}{\sum_{m=1}^n I(Y_m \geq Y_i)} \right. \\ \left. - \sum_{j=1}^n \frac{(1 - \Delta_j) I[Y_j \leq \min\{t_0 + \exp(\widehat{\beta}_{\xi|t_0}^{\prime} \mathbf{Z}_l), Y_i\}]}{\left\{ \sum_{m=1}^n I(Y_m \geq Y_j) \right\}^2} \right. \\ \left. - \sum_{l=1}^n \left\{ \mathbf{Z}_l \frac{(1 - \xi) I(Y_l \geq t_0)}{n \widehat{G}(t_0)} \right\} \left[\frac{(1 - \Delta_i) I(Y_i \leq t_0)}{\sum_{m=1}^n I(Y_m \geq Y_i)} \right. \right. \\ \left. \left. - \sum_{j=1}^n \frac{(1 - \Delta_j) I\{Y_j \leq \min(t_0, Y_i)\}}{\left\{ \sum_{m=1}^n I(Y_m \geq Y_j) \right\}^2} \right] \right]. \end{aligned}$$

A natural test statistic based on $\mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0,0})$ for testing H_0 would be

$$n^{-1} \mathbf{S}_{\xi|t_0,n}^{\prime}(\beta_{\xi|t_0,0}) \widehat{\mathbf{\Gamma}}_{\xi|t_0}^{-1} \mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0,0}),$$

which approximately follows a χ^2 -distribution with $p + 1$ degrees of freedom. A large observed value of this statistic suggests evidence against the null hypothesis H_0 . In Web Appendix C, we show that the distribution of $\widehat{\beta}_{\xi|t_0}$ is asymptotically normal for given \mathbf{Z}_i and fixed t_0 based on the local linearity of $\mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0})$.

Now we consider a partition of the regression coefficients, $\beta_{\xi|t_0}^{\prime} = (\beta_{\xi|t_0}^{(1)\prime}, \beta_{\xi|t_0}^{(2)\prime})$, where $\beta_{\xi|t_0}^{(1)}$ is a $r \times 1$ vector. Suppose

that $\hat{\beta}_{\xi|t_0}^{(1)}$ and $\hat{\beta}_{\xi|t_0}^{(2)}$ are the corresponding estimates, and we are only interested in testing the hypothesis $\tilde{H}_0 : \beta_{\xi|t_0}^{(1)} = \beta_{\xi|t_0,0}^{(1)}$, a specified vector, against a general alternative. For our test statistic, we will consider the minimum dispersion statistic (Basawa and Koul, 1988),

$$V(\beta_{\xi|t_0,0}^{(1)}) = \min_{\beta_{\xi|t_0}^{(2)}} \{n^{-1} \mathbf{S}_{t_0,n}'((\beta_{\xi|t_0,0}^{(1)})', (\beta_{\xi|t_0}^{(2)})')\} \\ \times \hat{\Gamma}_{\xi|t_0}^{-1} \mathbf{S}_{t_0,n}((\beta_{\xi|t_0,0}^{(1)})', (\beta_{\xi|t_0}^{(2)})'). \quad (5)$$

Note that evaluating this statistic does not require estimation of the probability density function of the survival distribution, which is needed for a Wald-type test statistic based on $\hat{\beta}_{\xi|t_0}$. By using similar arguments given in Wei, Ying and Lin (1990, Web Appendix 2) and Ying et al. (1995, Web Appendix C), it can be shown that (5) is approximately χ^2 -distributed with r degrees of freedom. We reject \tilde{H}_0 for a large value of $V(\beta_{\xi|t_0,0}^{(1)})$. By inverting this test statistic, a $100 \times (1 - \alpha)\%$ confidence region for $\beta_{\xi|t_0}^{(1)}$ can be obtained as

$$\{\beta_{\xi|t_0}^{(1)} : V(\beta_{\xi|t_0}^{(1)}) < \chi_{r,1-\alpha}^2\}, \quad (6)$$

where $\chi_{r,1-\alpha}^2$ is the $100 \times (1 - \alpha)\%$ th percentile of a χ^2 -distribution with r degrees of freedom.

By using this result we can also estimate the prediction interval. From (3), the ξ -quantile residual lifetime for a patient with a covariate vector \mathbf{z}_0 at time t_0 can be directly predicted by $\hat{\eta}(t_0 | \mathbf{z}_0) = \exp(\hat{\beta}_{\xi|t_0}' \mathbf{z}_0)$. We estimate the associated $100(1 - \alpha)\%$ confidence interval by $\{\eta : Q(\eta) < \chi_{1,1-\alpha}^2\}$, where

$$Q(\eta) = \min_{\{\beta : \beta' \mathbf{z}_0 = \log(\eta)\}} n^{-1} \mathbf{S}_{\xi|t_0,n}'(\beta) \hat{\Gamma}_{\xi|t_0}^{-1} \mathbf{S}_{\xi|t_0,n}(\beta). \quad (7)$$

4. Simulation Studies

To assess the finite sample properties of our estimator and test statistic, simulation studies were performed based on the median residual regression, i.e., $\xi = 1/2$. For an easy interpretation of the results, we considered a simple model

$$\text{med}\{\log(T_i - t_0) | T_i \geq t_0, x_{1i}\} = \beta_{t_0}^{(0)} + \beta_{t_0}^{(1)} x_{1i}, \quad (8)$$

where x_{1i} is a binary covariate, say, 0 for the control group and 1 for the treatment group. By the invariance property of the median with respect to monotone transformations, the model (8) is equivalent to

$$\text{med}(T_i - t_0 | T_i \geq t_0, x_{1i}) = \exp(\beta_{t_0}^{(0)} + \beta_{t_0}^{(1)} x_{1i}). \quad (9)$$

Here $\exp(\beta_{t_0}^{(0)})$ and $\exp(\beta_{t_0}^{(0)} + \beta_{t_0}^{(1)})$ can be interpreted as the median residual lifetimes for the control and treatment groups at time t_0 , respectively, so that the difference in median residual lifetime between the two groups is given by

$$\exp(\beta_{t_0}^{(0)}) \{\exp(\beta_{t_0}^{(1)}) - 1\}. \quad (10)$$

In this case, note that the slope parameter $\beta_{t_0}^{(1)}$ can be interpreted as the logarithm of the ratio of the two median residual lifetimes at time point t_0 .

Failure times T_i were assumed to follow a Weibull distribution with survival function $S(t) = \exp\{-(\rho t)^\kappa\}$. For this

distribution, note that under $\tilde{H}_0 : \beta_{t_0}^{(1)} = 0$ the equations (1) and (9) give

$$\theta_{t_0} \equiv \exp(\beta_{t_0}^{(0)}) = S^{-1}\{(1/2)S(t_0)\} - t_0 \\ = (1/\rho_0)\{\log(2) + (\rho_0 t_0)^\kappa\}^{1/\kappa} - t_0, \quad t_0 \geq 0. \quad (11)$$

Note that at the origin of time axis, i.e., $t_0 = 0$, $\rho_0 = \{\log(2)\}^{1/\kappa} / \exp(\beta_0^{(0)})$, where $\exp(\beta_0^{(0)})$ is the median of the failure distribution at $t_0 = 0$. By using the probability integral transformation, failure times were generated for both groups from $T_i = (1/\rho_0) \{-\log(1 - u_i)\}^{1/\kappa}$, where u_i is from a uniform random variable between 0 and 1. Censoring times C_i were generated from a uniform distribution between 0 and c , where constant c is for a certain censoring proportion. Finally the observed survival times Y_i were determined as the minimum of T_i and C_i . We assumed that $\kappa = 2$ and $\exp(\beta_0^{(0)}) = 5$ for convenience.

First, with a fixed total sample size of $n = 200$, 1000 simulations were performed for various combinations of censoring proportions and time points t_0 to evaluate the empirical distribution of the regression parameters $\beta_{t_0}^{(0)}$ and $\beta_{t_0}^{(1)}$ via the mean and standard error (SE) of the parameter estimates. The grid search method was used to find the estimates that minimize the estimating equation (4). Table 1 summarizes the results. Although the proposed regression method needs to be used at one fixed time point, we have chosen various time points only for the purpose of comprehensive investigation in our simulation studies and real example. Under $\tilde{H}_0 : \beta_{t_0}^{(1)} = 0$, equation (11) gives the true values of $\beta_{t_0}^{(0)}$ as 1.61, 1.41, 1.22, and 1.04 at $t_0 = 0, 1, 2$, and 3, respectively. We assume that the shape parameter κ does not change in the surviving population. The true value of $\beta_{t_0}^{(1)}$ must be 0 for all $t_0 \geq 0$ because an identical survival distribution was assumed for both groups. In Table 1,

Table 1

Mean (Standard Deviation) of the empirical estimates of the true regression parameters $\beta_{t_0}^{(0)} = 1.61, 1.41, 1.22, 1.04$ at $t_0 = 0, 1, 2, 3$ and $\beta_0^{(1)} = 0$ when the total number of observations (n) is 200

t_0	$c\%$	New method				MLE method			
		$\beta_{t_0}^{(0)}$	SE	$\beta_{t_0}^{(1)}$	SE	$\beta_{t_0}^{(0)}$	SE	$\beta_{t_0}^{(1)}$	SE
0	0	1.611	0.104	-0.004	0.140	1.612	0.057	0.002	0.074
	10	1.607	0.087	-0.002	0.127	1.611	0.058	-0.001	0.074
	20	1.593	0.290	0.008	0.307	1.610	0.060	0.001	0.081
	30	1.594	0.294	0.008	0.312	1.607	0.063	0.002	0.085
1	0	1.373	0.141	0.053	0.208	1.411	0.066	0.001	0.086
	10	1.401	0.093	0.001	0.142	1.408	0.068	0.002	0.090
	20	1.408	0.104	-0.001	0.159	1.413	0.069	0.001	0.095
	30	1.401	0.112	0.007	0.171	1.410	0.074	0.003	0.100
2	0	1.153	0.278	0.099	0.371	1.219	0.072	-0.003	0.100
	10	1.191	0.315	0.020	0.360	1.218	0.071	-0.002	0.101
	20	1.202	0.292	0.007	0.335	1.217	0.077	-0.003	0.110
	30	1.179	0.530	0.025	0.570	1.221	0.079	-0.006	0.113
3	0	0.865	0.768	0.219	0.898	1.038	0.078	0.001	0.109
	10	0.893	0.947	0.153	1.039	1.039	0.077	-0.001	0.115
	20	0.946	0.786	0.083	0.863	1.037	0.083	-0.001	0.116
	30	0.908	0.946	0.130	1.034	1.038	0.092	-0.003	0.130

we observe that the mean values of the estimates of $\beta_{t_0}^{(0)}$ and $\beta_{t_0}^{(1)}$ (no censoring) from the new method are very close to their true values for earlier fixed time points. Larger censoring proportions as well as later fixed time points tend to increase the SE of the parameter estimates.

In order to investigate the performance of our semiparametric estimator relative to the maximum likelihood (ML) approach, we also conducted a simulation study for the ML approach. The median residual life function for a subject i with a covariate x_i at time t_0 is determined by $\theta_{i,t_0} = (1/\rho_{i,t_0})\{\log(2) + (\rho_{i,t_0}t_0)^\kappa\}^{1/\kappa} - t_0 = \exp(\beta'_{t_0}\mathbf{Z}_i)$, where $\beta'_{t_0} = (\beta_{t_0}^{(0)}, \beta_{t_0}^{(1)})$ are the true values of the regression parameters at t_0 and $\mathbf{Z}'_i = (1, x_i)$. With a fixed value of κ , this implies

$$\rho_{i,t_0} = \{\log(2)\}^{1/\kappa} [\{\exp(\beta'_{t_0}\mathbf{Z}_i) + t_0\}^\kappa - t_0^\kappa]^{-1/\kappa}. \quad (12)$$

Denoting $\phi = (\kappa, \beta'_{t_0})$, the log-likelihood function to estimate the vector ϕ is given by $l(\phi) = \sum_{i=1}^n [\Delta_i \log\{h(y_i; \phi, \mathbf{Z}_i)\} + \log\{S(y_i; \phi, \mathbf{Z}_i)\}]$, where $\log\{h(y_i; \phi, \mathbf{Z}_i)\} = \log\{\log(2)\} + \log(\kappa) - \log[\{\exp(\beta'_{t_0}\mathbf{Z}_i) + t_0\}^\kappa - t_0^\kappa] + (\kappa - 1)\log(y_i)$ and $\log\{S(y_i; \phi, \mathbf{Z}_i)\} = -\log(2)[\{\exp(\beta'_{t_0}\mathbf{Z}_i) + t_0\}^\kappa - t_0^\kappa]^{-1/\kappa} y_i^\kappa$. We used the same simulation scenarios as before for a fair comparison. In Table 1, we observe that the bias of the regression estimators by ML method is negligible across all scenarios. Their corresponding SEs of the estimates tend to increase in censoring proportion and t_0 . The bias and SE of the MLE are small overall. The bias of our semiparametric estimators is small with small t_0 values, but slightly increases with $t_0 = 3$. As expected, the MLE is more efficient than our estimator under the correct distributional assumption.

Next, with total sample sizes of $n = 100, 200$, or 1000 , type I error probabilities for testing the null hypothesis $\tilde{H}_0: \beta_{t_0}^{(1)} = 0 (t_0 = 0, 1, 2, 3)$ with $\alpha = 0.05$ were compared between the proposed and ML methods at various time points with different average censoring proportions (Table 2). The results

indicate that the test based on the minimum dispersion statistic tends to be slightly conservative, which is consistent with the observations by Su and Wei (1993), Ying et al. (1995), and Jeong et al. (2008).

Similar settings were used for power analysis, except that the difference in median residual lifetimes between two groups were assumed to be 1, 3, and 5 at $t_0 = 0$. Note that when $\beta_0^{(0)} = 1.61$, equation (10) gives 1, 3, and 5 corresponding to $\beta_0^{(1)} = 0.18, 0.47$, and 0.69 , respectively. From (12), recall that when $t_0 = 0$ the scale parameter ρ from the Weibull distribution is given by

$$\rho_{i,0} = \{\log(2)\}^{1/\kappa} \exp(-\beta_0^{(0)} - \beta_0^{(1)}x_{1i}), \quad (13)$$

where $\beta_0^{(k)} (k = 0, 1)$ is the true values of the regression parameters at $t_0 = 0$. The covariate values of x_{1i} being generated from a Bernoulli distribution with equal probability of success, failure times were simulated from $T_i = (1/\rho_{i,0})\{-\log(1 - u_i)\}^{1/\kappa}$. Note that the median of the failure distribution at $t_0 = 0$ is determined by

$$(1/\rho_{i,0})\{\log(2)\}^{1/\kappa} = \exp(\beta_0^{(0)} + \beta_0^{(1)}x_{1i}). \quad (14)$$

Equation (14) gives 5 when $x_{1i} = 0$, and it corresponds approximately to 6, 8, and 10, respectively, when $x_{1i} = 1$ and $\beta_0^{(1)} = 0.18, 0.47$, and 0.69 . Table 3 summarizes the probabilities of rejecting the null hypothesis of $\tilde{H}_0: \beta_{t_0}^{(1)} = 0 (t_0 = 0, 1, 2, 3)$ for each true value of $\beta_0^{(1)}$. As expected, power tends to increase in n and $\beta_0^{(1)}$, and to decrease in censoring proportion.

5. A Real Example

The NSABP B-04 data set has been chosen as a real data example because it has a long-term follow-up over 30 years, preserving important survival information among breast cancer patients *without* any chemo- or hormonal therapies after

Table 2
Type I error (1–95% coverage) probabilities for testing the null hypothesis
 $H_0: \beta_{t_0}^{(1)} = 0 (t_0 = 0, 1, 2, 3)$ when the true parameter values are $\beta_0^{(1)} = 0$ and the total
sample sizes (n) are 100, 200, or 1000

t_0	$c\%$	New method			MLE method		
		$n = 100$	$n = 200$	$n = 1000$	$n = 100$	$n = 200$	$n = 1000$
0	0	0.050	0.038	0.041	0.065	0.058	0.056
	10	0.024	0.026	0.036	0.060	0.056	0.053
	20	0.027	0.032	0.035	0.062	0.054	0.055
	30	0.036	0.033	0.024	0.065	0.046	0.053
1	0	0.038	0.030	0.034	0.068	0.055	0.044
	10	0.028	0.024	0.033	0.062	0.046	0.050
	20	0.033	0.032	0.038	0.052	0.049	0.049
	30	0.038	0.037	0.033	0.043	0.050	0.047
2	0	0.041	0.033	0.040	0.068	0.051	0.058
	10	0.032	0.030	0.040	0.045	0.050	0.050
	20	0.034	0.032	0.027	0.070	0.056	0.049
	30	0.029	0.041	0.038	0.051	0.045	0.050
3	0	0.049	0.045	0.041	0.061	0.051	0.045
	10	0.028	0.036	0.032	0.069	0.054	0.040
	20	0.037	0.032	0.031	0.055	0.048	0.055
	30	0.032	0.031	0.039	0.064	0.053	0.052

Table 3

Powers for testing the null hypothesis $H_0: \beta_{t_0}^{(1)} = 0 (t_0 = 0, 1, 2, 3)$ when the true parameter values are $\beta_0^{(1)} = 0.18, 0.47, 0.69$ and the total sample sizes (n) are 100 or 200

		Censoring proportion											
		0%			10%			20%			30%		
	$\beta_0^{(1)}$	0.18	0.47	0.69	0.18	0.47	0.69	0.18	0.47	0.69	0.18	0.47	0.69
$n = 100$	$t_0 = 0$	0.18	0.87	0.99	0.13	0.76	0.98	0.14	0.69	0.96	0.09	0.61	0.91
	1	0.17	0.86	0.99	0.16	0.77	0.97	0.12	0.67	0.96	0.11	0.59	0.88
	2	0.16	0.82	0.98	0.14	0.75	0.97	0.11	0.67	0.93	0.09	0.55	0.90
	3	0.15	0.78	0.99	0.11	0.70	0.96	0.10	0.62	0.94	0.08	0.52	0.85
200	0	0.36	0.99	0.99	0.28	0.98	0.99	0.23	0.96	0.99	0.21	0.89	0.99
	1	0.36	0.99	0.99	0.27	0.97	0.99	0.24	0.94	0.99	0.22	0.90	0.99
	2	0.31	0.99	0.99	0.28	0.96	0.99	0.23	0.94	0.99	0.19	0.88	0.99
	3	0.28	0.97	0.99	0.27	0.96	0.99	0.17	0.92	0.99	0.17	0.84	0.99

surgery. This information is not available from recent studies on breast cancer because a series of standard therapies have been established for the patients being randomized to the control group. In this section, we apply the proposed method to re-analyze the B-04 data to evaluate the effects of important prognostic factors in breast cancer on the median residual lifetimes.

In the B-04 data, there were five groups being compared, three groups in node-negative patients and two groups in node-positive patients. With more than a quarter century of follow-up, the proportion of patients still alive is very low as expected, i.e., less than 30% among 1,079 eligible node-negative patients with follow-up and less than 20% among 586 eligible node-positive patients with follow-up. Fisher et al. (2002) reported a 25-year follow-up update of the B-04 data. The original goal of the study was to prove that a less aggressive total mastectomy had an equivalent effect on patients' survival to the traditional radical mastectomy. In this article, the nodal status was considered as a covariate along with age at diagnosis of breast cancer and pathological tumor size to predict median residual lifetimes of breast cancer patients at any given time point.

First we consider a simple regression model that includes only one covariate of the nodal status, i.e., $\mathbf{Z}_i = (1, X_{\text{node},i})'$, where $X_{\text{node},i}$ was coded as 0 or 1 for node-negative and node-positive, respectively, in the regression model (3). Table 4a summarizes the estimates of $\beta_{t_0}^{(\text{intercept})}$ and $\beta_{t_0}^{(\text{node})}$ ($t_0 = 0, 2, 4, 6, 7, 8, 10$) and the corresponding values of the minimum of the square root of sum of squares of two score functions, $\min \mathbf{S} = \min \|\mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0})\|$. Even though the median follow-up time was close to 30 years in this data set, the results presented here are only through 10 years because the data get sparse later on. Figure 1 shows contour plots of the surface of minus the square root of sum of squares of two score functions at year $t_0 = 0, 4$, and 8 with the intersection of the dotted horizontal and vertical lines being the regression parameter estimates found by the grid search method. Under this simple model, interpretation of $\hat{\theta}_{0,t_0} = \exp(\hat{\beta}_{t_0}^{(\text{intercept})})$ and $\hat{\theta}_{1,t_0} = \exp(\hat{\beta}_{t_0}^{(\text{intercept})} + \hat{\beta}_{t_0}^{(\text{node})})$ would be the estimated median residual lifetimes in node-negative and node-positive patients at time t_0 , respectively. Figure 2 shows a comparison of the estimated median residual lifetimes $\hat{\theta}_{0,t_0}$ and $\hat{\theta}_{1,t_0}$ from different approaches at various time points. Line type 1 repre-

sents the nonparametric estimates separately for each group by replacing the survival functions by their Kaplan-Meier estimators in (1). Line type 2 represents the estimates from the proposed regression model.

A comparison may be also considered between the proposed and proportional hazards regression models (Cox, 1972, 1975). Suppose we have a proportional hazards model $S(t; Z_i) = S_0(t)^{\eta_i}$, where $\eta_i = \exp(\beta Z_i)$ and Z_i is an indicator for the nodal status. Under this model, the ξ -quantile residual lifetime at t_0 can be evaluated by $\theta_{\xi|t_0} = S_0^{-1}\{(1 - \xi)^{1/\eta_i} S_0(t_0)\} - t_0$, by noting that $S^{-1}(y) = S_0^{-1}(y^{1/\eta_i})$ and using the fact $\theta_{\xi|t_0} = S^{-1}\{(1 - \xi)S(t_0)\} - t_0$. Line type 3 in Figure 2 represents such estimates from the NSABP B-04 data. The comparison indicates that the estimates from the proposed model seem to follow the path of the nonparametric one-sample estimates closer in each group than the Cox model does. The indication of lack of fit in the Cox model may be due to violation of the proportional hazards assumption in the B-04 data ($p = 0.0002$ by Grambsch and Therneau, 1994).

In Figure 2, we observe that the median residual lifetime of the node-positive group is shorter than that of the node-negative group during the first 10 years. The median residual lifetime in the node-negative group is almost constant for the range of t_0 values. For the node-positive group, however, the median residual lifetime slightly increases in t_0 , which may imply that the node-positive patients tend to survive longer once they pass the early high-risk period. p -values and 95% confidence intervals from the minimum dispersion statistic (5) are provided in Table 4a to compare median residual lifetimes between the two nodal groups at each fixed time point t_0 . The results indicate that the nodal status has a significant effect on the median residual lifetimes in early time points, but the effect diminishes once patients survive more than 7 years. This is similar to the findings in Jeong et al. (2008) by using a two-sample test statistic without adjusting for other prognostic factors.

Since there are no other existing methods that can be directly compared with the proposed method, first we have adopted the bootstrap resampling method (Efron, 1979) to assess our results. We have resampled 1000 times from the original sample with replacement and calculated SEs of those 1000 estimates, which were 0.07, 0.10, 0.13, 0.14, 0.12, 0.11, and 0.16 for time points $t = 0, 2, 4, 6, 7, 8$, and 10, respectively.

Table 4

Fitted simple and multiple regression models (NSABP B-04 data); (a) Regression parameter estimates from the median residual life regression model with a single covariate (simple model) and associated p -values and 95% confidence intervals from the proposed (new), bootstrap resampling (bs), and proportional hazards methods (Cox) for testing the null hypothesis $H_0 : \beta_{t_0}^{(node)} = 0 (t_0 = 0, 2, 4, 6, 7, 8, 10)$, (b) Regression parameter estimates from the median residual life regression model with multiple covariates and 95% confidence intervals (CI) for testing the null hypothesis $H_0 : \beta_t^{(cov)} = 0 (t_0 = 0, 2, 4, 6, 8, 10, \text{cov} = \text{node, age, or tsize})$ ($\min \mathbf{S} = \min \|\mathbf{S}_\xi|_{t_0, n}(\beta_{\xi|t_0})\|$)

t_0	$\min \mathbf{S}$	$\hat{\beta}_{t_0}^{(intercept)}$	$\hat{\beta}_{t_0}^{(node)}$	95% CI _{new}	p -value _{new}	95% CI _{bs}	p -value _{bs}	p -value _{Cox}
(a) Simple regression model								
0	0.51	2.54	-0.62	(-0.74, -0.47)	<0.0001	(-0.76, -0.48)	<0.0001	<0.0001
2	0.22	2.53	-0.59	(-0.77, -0.37)	<0.0001	(-0.79, -0.40)	<0.0001	<0.0001
4	0.52	2.56	-0.50	(-0.72, -0.21)	0.0003	(-0.75, -0.26)	0.0001	<0.0001
6	0.02	2.59	-0.44	(-0.71, -0.17)	0.001	(-0.71, -0.17)	0.002	<0.0001
7	0.15	2.57	-0.26	(-0.57, -0.09)	0.008	(-0.50, -0.06)	0.013	0.001
8	0.67	2.54	-0.22	(-0.42, 0.05)	0.116	(-0.44, -0.01)	0.038	0.038
10	0.97	2.46	-0.09	(-0.48, 0.11)	0.343	(-0.41, 0.19)	0.475	0.094
(b) Multiple regression model								
	$\hat{\beta}_{t_0}^{(intercept)}$	95% CI	$\hat{\beta}_{t_0}^{(node)}$	95% CI	$\hat{\beta}_{t_0}^{(age)}$	95% CI	$\hat{\beta}_{t_0}^{(tsize)}$	95% CI
0	3.29	(2.91, 4.24)	-0.51	(-0.72, -0.32)	-0.83	(-2.33, -0.21)	-1.02	(-1.85, -0.61)
2	3.94	(3.31, 4.83)	-0.44	(-0.65, -0.21)	-2.06	(-3.17, -1.10)	-0.89	(-1.18, -0.40)
4	4.06	(3.43, 4.79)	-0.35	(-0.63, -0.13)	-2.29	(-3.38, -1.63)	-0.71	(-1.59, 0.97)
6	3.92	(3.54, 4.68)	-0.14	(-0.65, 0.10)	-2.11	(-3.65, -1.59)	-0.48	(-1.40, 0.96)
8	3.88	(3.42, 4.77)	-1.60	(-0.51, 0.13)	-2.21	(-3.18, -1.61)	-0.36	(-1.05, 0.81)
10	3.87	(3.27, 4.84)	-0.10	(-0.47, 0.12)	-2.33	(-3.90, -1.60)	-0.36	(-0.91, 0.84)

At each time point, these SE estimates were used to construct 95% confidence intervals for $\beta_{t_0}^{(node)}$ as well as to obtain p -values via the Wald test. The proportional hazards model (Cox, 1972, 1975) was also chosen for a comparison of the observed significance of the slope parameter. In this case, the significance of the effect of nodal status was evaluated among surviving patients beyond each fixed time point. The results from the proportional hazards analysis are presented only through p -values for the comparison, because the quantities being evaluated are different under the proportional hazards and time-specific median residual life models. Note that the interpretation of the slope parameter is the logarithm of the hazard ratio under the proportional hazards model and the logarithm of the ratio of two median residual lifetimes at time t_0 under the proposed model, so that confidence intervals for the slope parameter from the two models cannot be compared directly. From Table 4a, we observe that the results from the three approaches are compatible in p -values, although ones from the proposed method tend to be slightly larger. We also see that the resampling and proposed methods give similar 95% confidence intervals.

Now we consider the multiple regression model, where $\mathbf{Z}_i = (1, X_{node,i}, X_{age,i}, X_{tsize,i})'$, $X_{node,i} = 0$ (node-negative) or 1 (node-positive), and both age ($X_{age,i}$) and pathological tumor size ($X_{tsize,i}$) are continuous covariates. The age and tumor size values were rescaled simply by multiplying by 0.01, so that the regularity condition holds in the estimates $\hat{G}(\cdot)$ in the estimating equation (4). Table 4b shows the estimated regression parameters and corresponding 95% confidence intervals from the minimum dispersion statistic for testing $H_0 : \beta_{t_0}^{(cov)} = 0 (t_0 = 0, 2, 4, 6, 8, 10, \text{cov} = \text{node, age, or tsize})$. The results indicate that the negative effect of age at a di-

agnosis of cancer on the median residual lifetimes tends to become stronger for later time points, whereas the effects of nodal status and pathological tumor size are significant only up to first 4 or 5 years, when adjusting for other covariates.

These results can be used to predict a patient's median residual lifetime at a given time t_0 . For example, for a patient with positive lymph nodes, age at diagnosis of 56 (median), and pathological tumor size of 30mm (median), the median residual lifetime (in years) of 6 years after the original diagnosis of cancer can be predicted as $\exp\{3.92 - 1 \times 0.28 - 2.11 \times (0.01 \times 56.0) - 0.48 \times (0.01 \times 30.0)\} = 10.1$. Similarly for node-negative patients, the predicted median residual lifetime is 13.4 years. From (7), the corresponding 95% prediction intervals are (7.3, 12.9) and (11.2, 16.4), respectively. These estimates may be used as baseline mortality information for a future clinical study to evaluate a new therapy in terms of prolonging the patients' remaining lifetimes.

6. Discussion

In this article, we proposed a regression model on quantile residual lifetimes, where the covariate effects can be directly associated with the quantile failure time among survivors beyond any specific time point, rather than inferring the quantile residual lifetime under some global regression model assumptions such as proportional hazards or AFT. An estimating equation for the regression coefficients has been formulated and justified under the proposed model. Asymptotic distributions of the proposed regression parameter estimator and test statistic to evaluate the covariate effects have been derived, and it was shown that they had reasonable finite sample properties through simulation studies. The grid search method that was used to find the roots of the estimating equation

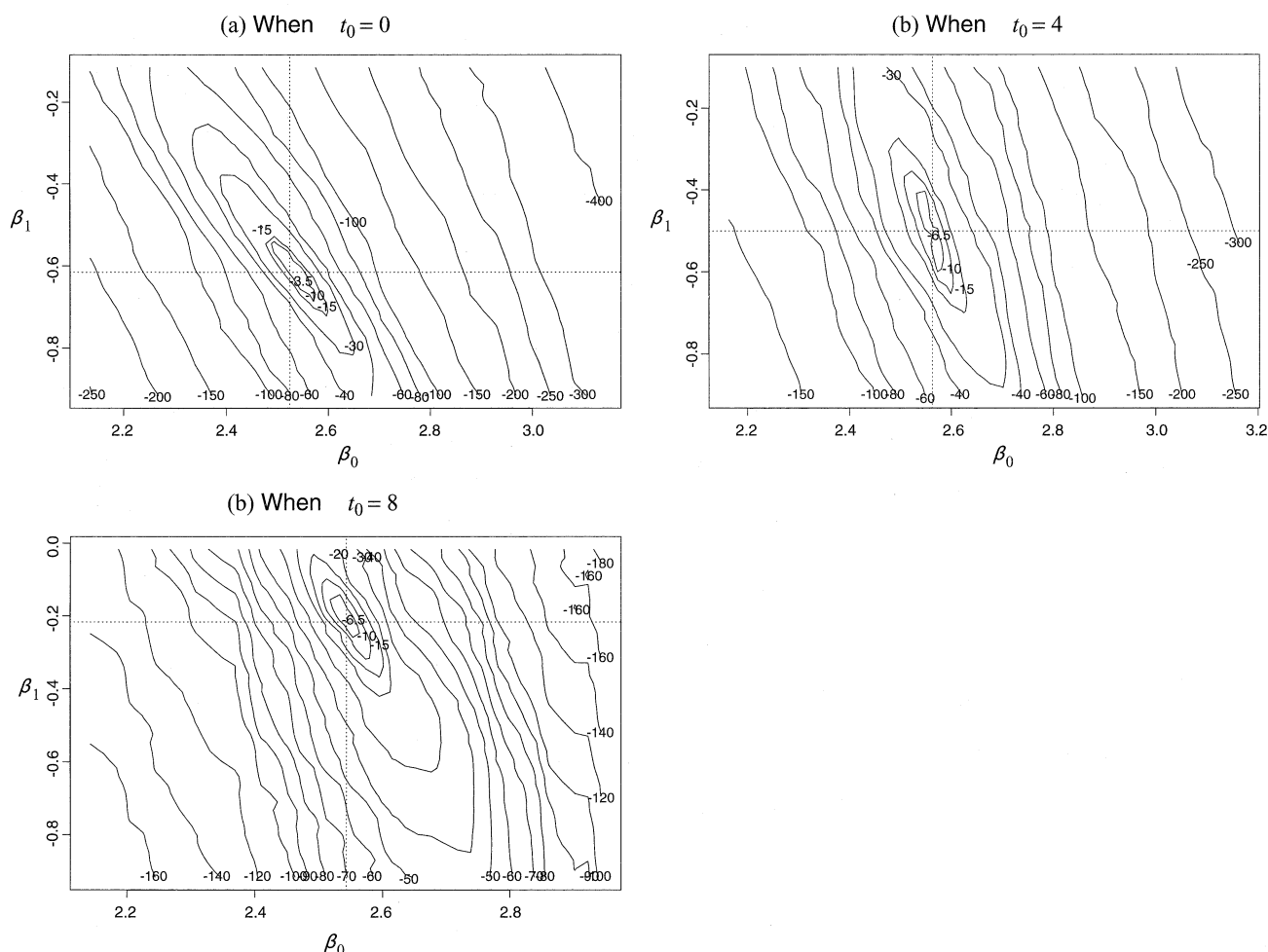


Figure 1. Contour plots of surfaces of minus the square root of sum of squares of two score equations for the median residual regression model applied to NSABP B-04 data set with the only covariate of nodal status; (a) When $t_0 = 0$, (b) When $t_0 = 4$, and (c) When $t_0 = 8$.

also provided reasonable solutions in our simulation studies and real example. The variance of the proposed estimators depends on the probability density function of the survival variables. However, the variance of the estimating function can be consistently estimated without specification of the survival distribution. Hence, we used the minimum dispersion statistic (Basawa and Koul, 1988) based on the distribution of the estimating function. The variance of the regression coefficient estimator may be estimated by using either the proposed confidence interval approach or the bootstrap resampling method.

Our theoretical results in Section 3 are based on the uniform consistency of the Kaplan–Meier estimator $\hat{S}(t)$ over $0 < t < \zeta = \sup\{t : S(t)G(t) > 0\}$, which leads to a condition that $t_0 + \exp(\beta'_{\xi|t_0} \mathbf{Z}) < \zeta$ in our case. Here ζ may be considered as the maximum follow-up time, which has the lower bound $t_0 + \exp(\beta'_{\xi|t_0} \mathbf{Z})$.

As noted in Ying et al. (1995), to check the model assumption of (3) at a fixed time t_0 , a zero-mean Gaussian process of cumulative sums of median residuals can be defined as $W(s) = n^{-1/2} \sum_{i=1}^n e_i I(\hat{\beta}'_{\xi|t_0} \mathbf{Z}_i \leq s)$, where

$$e_i = \frac{I\{Y_i \geq t_0 + \exp(\hat{\beta}'_{\xi|t_0} \mathbf{Z}_i)\}}{\hat{G}\{t_0 + \exp(\hat{\beta}'_{\xi|t_0} \mathbf{Z}_i)\}} - (1 - \xi) \frac{I(Y_i \geq t_0)}{\hat{G}(t_0)}.$$

A plot of $W(s)$ against the predicted ξ -quantile residual life-times may be used as a graphical tool to check the model assumption at time t_0 .

We have assumed that survival and censoring variables are conditionally independent given covariates. By using a censoring distribution model that is independent of covariates, however, the conditional independence assumption is reduced to an unconditional independence assumption, as a referee pointed out. The latter stronger assumption holds for most well-conducted clinical trials, including the NSABP trial discussed in this article, where the major cause of censoring is administrative. If the censoring distribution depends on a discrete covariate, we may partition the data into a number of strata defined by the covariate values so that the censoring variables have an identical distribution within each stratum. If the censoring distribution depends on a continuous covariate,

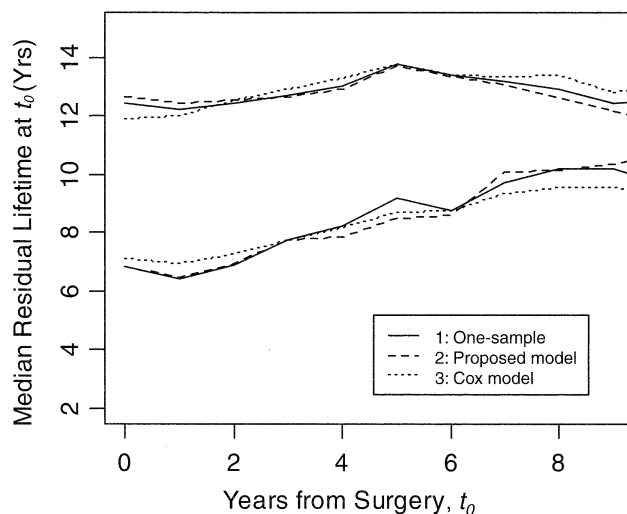


Figure 2. Comparison of the estimated median residual lifetimes by nodal status from various methods (NSABP B-04 data); upper three lines are for node-negative group and lower three lines are for node-positive group.

we may consider modeling the dependency of censoring distribution on the covariate. For example, for a continuous covariate z_i , suppose that C_i satisfies a proportional hazards model $h_i(t) = h_0(t) \exp(\gamma z_i)$, where $h_i(t)$ is the hazard function of C_i and $h_0(t)$ is the baseline hazard function. Then, $\hat{G}(t)$ in (4) will be replaced by $\hat{G}(t|z_i) = \exp\{-\hat{h}_0(t) \exp(\hat{\gamma} z_i)\}$, where $\hat{\gamma}$ is the partial MLE (Cox, 1972) for the censoring distribution and $\hat{h}_0(t)$ is the Breslow's (1974) estimator of $h_0(t)$. If the assumed censoring model is correct, the consistency and asymptotic normality of $\hat{\beta}_{\xi|t_0}$ can be derived using the properties of $\hat{\gamma}$.

An associate editor raised an issue regarding how to choose quantiles and how to interpret the results if the covariate effects differ across quantiles. In theory, any quantile ξ such that $\xi < \{\hat{S}(t_0) - \hat{S}(t_{max})\} / \hat{S}(t_0)$, where $t_{max} = \max(X_i)$ and $X_i = \min(T_i, C_i)$, can be chosen, but one closer to the middle is recommended due to less variability. In the same context, the interpretation with one toward the middle should be considered more reliable. In practice, however, determination of which quantiles to be investigated would also depend on the scientific objectives of the study. For example, if the investigator is interested in inferring a patient population with longer survival, high quantiles such as 80 or 90 percentile, rather than the median, would be more appropriate. We have re-analyzed the NSABP B-04 data across different quantiles at $t_0 = 0$ in the simple regression model. For the 10th, 25th, 50th, and 75th percentile models, the estimates of the slope parameter were -0.620 , -0.623 , -0.620 , and -0.351 , respectively, all of which were associated with the p -value smaller than 0.0001. Thus the interpretation of the effect of the nodal status on patients' survival does not change for the use of different quantiles in this example.

7. Supplementary Materials

The proofs of consistency and asymptotic normality of $\hat{\beta}_{\xi|t_0}$ and asymptotic normality of $n^{-1/2} \mathbf{S}_{\xi|t_0,n}(\beta_{\xi|t_0}^0)$ in Section 3

are available under the Paper Information link at the Biometrics website <http://www.biometrics.tibs.org>.

ACKNOWLEDGEMENTS

We are very grateful to an associate editor and a referee for their valuable comments. Dr. Jong-Hyeon Jeong's work was supported in part by the Department of Defense (DOD) Grant W81XWH-04-1-0605. Dr. Jong-Hyeon Jeong and Dr. Hanna Bandos's work was supported in part by the National Health Institute (NIH) Grants 5-U10-CA69974-09 and 5-U10-CA69651-11.

REFERENCES

- Alam, K. and Kulasekera, K. B. (1993). Estimation of the quantile function of residual life time distribution. *Journal of Statistical Planning and Inference* **37**, 327–337.
- Arnold, B. C. and Brockett, P. L. (1983). When does the β th percentile residual life function determine the distribution? *Operations Research* **31**, 391–396.
- Basawa, I. V. and Koul, H. L. (1988). Large-sample statistics based on quadratic dispersion. *International Statistical Review* **56**, 199–219.
- Breslow, N. E. (1974). Covariate analysis of censored survival data. *Biometrics* **30**, 89–99.
- Coombes, R. C., Hall, E., Gibson, L. J. et al. (2004). A randomized trial of exemestane after two to three years of tamoxifen therapy in postmenopausal women with primary breast cancer. *The New England Journal Of Medicine* **350**, 1081–1092.
- Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society Series B* **34**, 187–220.
- Cox D. R. (1975). Partial likelihood. *Biometrika* **62**, 269–276.
- Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman & Hall.
- Csörgő, M. and Csörgő, S. (1987). Estimation of percentile residual life. *Operations Research* **35**, 598–606.
- Csörgő, S. and Horvath, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **62**, 411–426.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics* **7**, 1–26.
- Fisher, B., Jeong, J., Anderson, S. et al. (2002). Twenty-five year findings from a randomized clinical trial comparing radical mastectomy with total mastectomy and with total mastectomy followed by radiation therapy. *The New England Journal Of Medicine* **347**, 567–575.
- Fleming, T. R. and Harrington, D. P. (1991). *Counting Processes and Survival Analysis*. New York: Wiley.
- Gelfand, A. E. and Kottas, A. (2003). Bayesian semiparametric regression for median residual life. *The Scandinavian Journal of Statistics* **30**, 651–665.
- Goss, P. E., Ingle, J. N., Martino, S. et al. (2003). A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early-stage breast cancer. *The New England Journal Of Medicine* **349**, 1793–1802.
- Grambsch, P. and Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81**, 515–526.
- Gupta, R. C. and Langford, E. S. (1984). On the determination of a distribution by its median residual life function: A functional equation. *Journal of Applied Probability* **21**, 120–128.
- Jeong, J.-H., Jung, S.-H., and Costantino, J. P. (2008). Nonparametric inference on median residual life function. *Biometrics* **64**, 157–163.
- Joe, H. and Proschan, F. (1984). Percentile residual life functions. *Operations Research* **32**, 668–678.

- Jung S.-H. (1996). Quasi-likelihood for median regression models. *Journal of the American Statistical Association* **91**, 251–257.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimator from incomplete observations. *Journal of the American Statistical Association* **53**, 457–481.
- Koenker, R. and Bassett, G. J. (1978). Regression quantiles. *Econometrica* **46**, 33–50.
- Mamounas, E. P., Jeong, J.-H., Wickerham, L. et al. (2008). Benefit from exemestane as extended adjuvant therapy after 5 years of adjuvant tamoxifen: Intention-to-treat analysis of the National Surgical Adjuvant Breast and Bowel Project B-33 trial. *Journal of Clinical Oncology* **26**, 1965–1971.
- Mu, Y. and He, X. (2007). Power transformation toward a linear regression quantile. *Journal of the American Statistical Association* **102**, 269–279.
- Portnoy, S. and Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Statistical Science* **12**, 279–296.
- Schmittlein, D. C. and Morrison, D. G. (1981). The median residual lifetime: A characterization theorem and an application. *Operations Research* **29**, 392–399.
- Song, J. and Cho, G. (1995). A note on percentile residual life. *Sankhya* **57**, 333–335.
- Su, J. Q. and Wei, L. J. (1993). Nonparametric estimation for the difference or ratio of median failure times. *Biometrics* **49**, 603–607.
- Ying, Z., Jung, S.-H., and Wei, L. J. (1995). Survival analysis with median regression models. *Journal of the American Statistical Association* **90**, 178–184.
- Wei, L. J., Ying, Z., and Lin, D. Y. (1990). Linear regression analysis of censored survival data based on rank tests. *Biometrika* **19**, 845–851.
- Wei, Y., Pere, A., Koenker, R., and He, X. (2006). Quantile regression methods for reference growth charts. *Statistics in Medicine* **25**, 1369–1382.
- Whang Y. J. (2006). Smoothed empirical likelihood methods for quantile regression models. *Econometric Theory* **22**, 173–205.

Received November 2007. Revised October 2008.
Accepted October 2008.