Published in final edited form as:

Stat Med. 2014 June 30; 33(14): 2467–2479. doi:10.1002/sim.6105.

Likelihood approaches for proportional likelihood ratio model with right-censored data

Hong Zhu^{*,†}

Division of Biostatistics, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX, U.S.A

Abstract

Regression methods for survival data with right censoring have been extensively studied under semiparametric transformation models [1] such as the Cox regression model [2] and the proportional odds model [3]. However, their practical application could be limited due to possible violation of model assumption or lack of ready interpretation for the regression coefficients in some cases. As an alternative, in this paper, the proportional likelihood ratio model introduced by Luo and Tsai [4] is extended to flexibly model the relationship between survival outcome and covariates. This model has a natural connection with many important semiparametric models such as generalized linear model and density ratio model, and is closely related to biased sampling problems. Compared with the semiparametric transformation model, the proportional likelihood ratio model is appealing and practical in many ways because of its model flexibility and quite direct clinical interpretation. We present two likelihood approaches for the estimation and inference on the target regression parameters under independent and dependent censoring assumptions. Based on a conditional likelihood approach using uncensored failure times, a numerically simple estimation procedure is developed by maximizing a pairwise pseudo-likelihood [5]. We also develop a full likelihood approach and the most efficient maximum likelihood estimator is obtained by a profile likelihood. Simulation studies are conducted to assess the finite-sample properties of the proposed estimators and compare the efficiency of the two likelihood approaches. An application to survival data for bone marrow transplantation patients of acute leukemia is provided to illustrate the proposed method and other approaches for handling non-proportionality. The relative merits of these methods are discussed in concluding remarks.

Keywords

conditional likelihood; pairwise pseudo-likelihood; profile likelihood; proportional likelihood ratio model; right-censored data

1. Introduction

Survival data arise frequently in medical follow-up study where failure time is observed subject to right censoring. A longstanding statistical problem is how to assess the association

*Correspondence to: Hong Zhu, Division of Biostatistics, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX 75390, U.S.A .

†hong.zhu@utsouthwestern.edu

between risk factors and survival in the population given observed survival data. Various parametric, semiparametric and nonparametric regression models have been proposed for survival data. Among others, semiparametric transformation model has been extensively investigated [1, 6], where the linear transformation function is unspecified and the error distribution is assumed to be completely known. The transformation model extends the Cox regression model with the extreme value error distribution as well as the proportional odds model with the logistic error distribution. Although the transformation model is of interest in its own right, it either relies on a strong model assumption as in the Cox regression model or suffers from a lack of ready interpretation for the regression coefficients, except in the extreme value and logistic error special cases. For complete data, Luo and Tsai [4] introduced a semiparametric proportional likelihood ratio model to specify the nonlinear relationship between outcome and covariates with an unknown baseline distribution. Huang and Rathouz [7] explicitly modeled the mean function for a proportional likelihood ratio model with discrete covariates. It has been discussed in [4] that the proportional likelihood ratio model has a strong connection with the generalized linear model [8], the density ratio model [9, 10], the selection bias model [11], the single-index model [12] and the exponential tilt regression model [13]. Particularly, Wang *et al.* [14] studied the application of the exponential tilt model to compare survival distributions in two groups, which is a special case of the proportional likelihood ratio model for right-censored data. The proportional likelihood ratio model relates the outcome Y and a $p \times 1$ covariate vector X as

$$f(y|x) = \frac{\exp(\beta^T xy) g(y)}{\int \exp(\beta^T xy) dG(y)}, \quad (1)$$

where g is the density of an unknown baseline distribution function G with respect to some dominating measure, and β is a parameter vector of interest with a meaningful interpretation. For any y_1, y_2 and x , we have

$$\log \frac{f(y_2|x)}{f(y_1|x)} = \log \frac{g(y_2)}{g(y_1)} + \beta^T x (y_2 - y_1).$$

Further, if X is one-dimensional, then

$$e^\beta = \frac{f(y+1|x+1) / f(y|x+1)}{f(y+1|x) / f(y|x)},$$

that is, β is the log likelihood ratio that the outcome increases by one unit given the covariate increases by one unit. Clearly, if the form of g is known in (1), then the proportional likelihood ratio model becomes the well-known exponential family model. However, if g is unknown, the conventional parametric likelihood approach cannot be used.

In this paper, as an alternative to the transformation model, we adapt the proportional likelihood ratio model to investigate the relationship between survival outcome and covariates in a flexible fashion. Two likelihood approaches are presented for the estimation

and inference on the target regression parameter vector β under independent and dependent censoring assumptions. Based on a conditional likelihood, we provide a computationally convenient estimation procedure by applying a pairwise pseudo-likelihood method [5], where we take advantage of the invariance property of the target regression parameter vector under the proportional likelihood ratio model for uncensored failure times. The pairwise pseudo-likelihood can eliminate the unknown baseline distribution function and thus has a nice simple form. Full likelihood approach is regarded as the most efficient means for inference in statistical analysis, and we also develop a full likelihood approach under the proportional likelihood ratio model. The nonparametric maximum likelihood estimator of regression parameter vector is obtained by extending Luo & Tsai's iterative algorithm with a profile likelihood method.

Non-proportional hazards, such as converging, diverging or crossing hazards, occur frequently in survival analysis. The proportional likelihood ratio model for right-censored data considered in the paper could be very useful in real application, where the proportionality assumption is not satisfied and commonly-used Cox regression model is not appropriate. For a binary covariate X indicating group membership, Figure 1 graphically demonstrates different situations of hazard functions of group 1 ($X = 1$) and group 2 ($X = 0$) under the proportional likelihood ratio model. Consider baseline density as the unit Exponential density, $g(y) = \exp(-y)$; $y > 0$ and proportional likelihood ratio model regression parameter $\beta = -1$, Fig. 1 (a) represents proportional hazards for groups 1 and 2. When baseline density is taken to be a Gamma density function $g(y) = y^{(a-1)} \exp(-y/\theta) / \theta^a \Gamma(a)$; $y > 0$ with shape parameter $a = 2$ and rate parameter $\theta = 1$, and the regression parameter $\beta = -1$, Fig. 1 (b) represents non-proportional (diverging) hazards for group 1 and 2. Additionally, the proportional likelihood ratio model would lead to regression effect that is easy to be interpreted to medical practitioners and patients as shown in our data analysis. Meanwhile, there are other options to appropriately handle non-proportional hazards, including (i) stratification on a covariate with non-proportional hazards, (ii) separate Cox models for different time periods, (iii) inclusion of time-dependent covariate terms, such as fractional polynomials for continuous variables, in the Cox model [15] and (iv) estimation of average hazard ratios [16]. These approaches will also be illustrated in comparative analyses of leukemia survival data and their relative merits will be summarized in the data analysis and discussion section.

We organize the remainder of the article as follows. In Section 2, we discuss the invariance property under the proportional likelihood ratio model for uncensored failure times, and present the conditional likelihood approach to estimate the target regression parameter vector under independent and dependent censoring assumptions. We develop the most efficient nonparametric maximum likelihood estimator of the regression parameter vector based on a full likelihood in Section 3. The finite-sample performances of the proposed conditional and full likelihood estimators are evaluated and compared by simulation studies in Section 4. In Section 5, we apply our methods to survival data for bone marrow transplantation patients of acute leukemia as an illustration. We provide concluding remarks and topics for further research in Section 6.

2. Conditional likelihood approach

2.1. Invariance property and estimation under independent censoring

Suppose that Y is the failure time, X is a $p \times 1$ covariate vector, C is the censoring time, and $\delta = I(Y \leq C)$ is the censoring indicator. The probability of observing a failure time is $\text{pr}(Y \leq C)$. The conditional density function of observing uncensored Y can be expressed as

$$f^c(y|x, \delta=1) = \frac{\lim_{\Delta \rightarrow 0} \text{pr}(y \leq Y < y+\Delta, C > y|x) / \Delta}{\text{pr}(Y \leq C)}.$$

Let S_C denote the survival function of C . Assume that C and (Y, X) are independent, then under the proportional likelihood ratio model in (1), the conditional density becomes

$$f^c(y|x, \delta=1) = \frac{f(y|x) S_C(y)}{\int f(y|x) S_C(y) dy} = \frac{\exp(\beta^T xy) g(y) S_C(y)}{\int \exp(\beta^T xy) g(y) S_C(y) dy} = \frac{\exp(\beta^T xy) g^*(y)}{\int \exp(\beta^T xy) dG^*(y)}, \quad (2)$$

where $g^*(y) = g(y)S_C(y)$ is an unspecified density function with distribution function $G^*(y)$, and β is the same as the one in (1), the marginal density for the population. It shows the conditional density given uncensored failure times retains a proportional likelihood ratio model framework with the same regression parameter vector β as that in the density for the population model. Therefore, under independent censoring assumption, the proportional likelihood ratio model is invariant with respect to the regression parameter vector given uncensored failure times. In this case, even though the sample is biased, uncensored failure times can be used directly for estimating β . For proportional likelihood ratio model with complete data, a nonparametric maximum likelihood method was proposed by Luo and Tsai [4] where β and G are estimated simultaneously by an iterative algorithm. Chan [17] showed the invariance property for randomly truncated data under the proportional likelihood ratio model, and suggested an alternative estimator of β by a pairwise pseudo-likelihood [5] which eliminates the functional nuisance parameter G . Similarly, under independent censoring, we rely on the invariance property for uncensored failure times, and adopt the method in Chan [17] to estimate β for right-censored data.

Let $\{\tilde{y}_i = \min(y_i, c_i), \delta_i = I(y_i \leq c_i), x_i\}$ be the observed data for the i th subject ($i = 1, \dots, n$). Given uncensored failure times $\{y_{(1)}, \dots, y_{(K)}\}$ where $0 < y_{(1)} < \dots < y_{(K)}$ and K is the number of failures among n subjects, we have the conditional pairwise pseudo-likelihood

$$L_{CP}(\beta) \propto \prod_{(i) < (k)} \frac{1}{1 + \exp\left\{-\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)})\right\}}.$$

A consistent estimate of β can be obtained by maximizing $L_{CP}(\beta)$ and is denoted by $\hat{\beta}_{CP}$. Let

$$\psi_{(i),(k)}(\beta) = - \frac{\partial \left(\log \left[1 + \exp \left\{ -\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)}) \right\} \right] \right)}{\partial \beta}.$$

Following the lines in Chan [17], it can be proved that the estimate $\hat{\beta}_{CP}$ is consistent, and as $K \rightarrow \infty$, $K^{1/2} (\hat{\beta}_{CP} - \beta_0)$ converges weakly to a multivariate normal distribution with mean zero and variance-covariance matrix $V = \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1}$, where $\Sigma_1 = E \left\{ \partial \psi_{(1),(2)}(\beta_0) / \partial \beta \right\}$ and $\Sigma_2 = 4E \left\{ \psi_{(1),(2)}(\beta_0) \psi_{(1),(3)}^T(\beta_0) \right\}$.

2.2. Estimation under dependent censoring

The invariance of proportional likelihood ratio model for the conditional density of uncensored failure times holds under independence censoring assumption but is not true in general. Next, we focus on the proportional likelihood ratio model in presence of dependent censoring. Although failure time and censoring time can be related in many applications, when much auxiliary information has been collected, it is common to assume that the covariate vector X is sufficient to explain the dependence between failure time Y and censoring time C [18]. Equivalently, Y and C are independent conditional on X . Under the conditional independent assumption of $Y \perp C | X$, the conditional density function of observing uncensored Y is

$$\begin{aligned} f^c(y|x, \delta=1) &= \frac{f(y|x) S_C(y|x)}{\int f(y|x) S_C(y|x) dy} = \frac{\exp(\beta^T xy) S_C(y|x) g(y)}{\int \exp(\beta^T xy) S_C(y|x) g(y) dy} \\ &\propto \exp(\beta^T xy) S_C(y|x) g(y). \end{aligned}$$

Accordingly, the conditional pairwise pseudo-likelihood function given uncensored failure times $\{y_{(1)}, \dots, y_{(K)}\}$ is

$$L_{CP}^*(\beta, S_C) \propto \prod_{(i)<(k)} \frac{1}{1 + b(y_{(i)}, x_{(i)}, y_{(k)}, x_{(k)}) \exp \left\{ -\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)}) \right\}},$$

where

$$b(y_{(i)}, x_{(i)}, y_{(k)}, x_{(k)}) = \frac{S_C(y_{(i)}|x_{(k)}) S_C(y_{(k)}|x_{(i)})}{S_C(y_{(i)}|x_{(i)}) S_C(y_{(k)}|x_{(k)})}.$$

If S_C is known, $L_{CP}^*(\beta, S_C)$ only involves regression parameter vector β and an estimate of β can be produced by a similar procedure to the model under independent censoring assumption in Section 2.1. The corresponding score function is

$$\tilde{U}_{CP}(\beta) = \sum_{(i) < (k)} - \frac{\partial \left(\log \left[1 + b \left(y_{(i)}, x_{(i)}, y_{(k)}, x_{(k)} \right) \exp \left\{ -\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)}) \right\} \right] \right)}{\partial \beta}.$$

When S_C is unknown, we replace S_C with a consistent estimator $\hat{S}_C$ in $\tilde{U}_{CP}(\beta)$. Further, suppose that censoring time C follows a proportional hazards model

$$h_C(t|x) = h_0(t) \exp(\gamma^T x),$$

where $h_0(t)$ is an arbitrary baseline hazard function for C and γ is a $p \times 1$ vector of regression coefficients. Let $\hat{\gamma}$ denote the maximum partial likelihood estimator of γ , and $\hat{S}_0(t)$ denote the Breslow estimator for the baseline survival function of C . The survival function S_C can be consistently estimated by $\hat{S}_C(t|x) = \hat{S}_0(t) \exp(\hat{\gamma}^T x)$. Thus, we have

$$\hat{b} \left(y_{(i)}, x_{(i)}, y_{(k)}, x_{(k)} \right) = \left\{ \frac{\hat{S}_0(y_{(i)})}{\hat{S}_0(y_{(k)})} \right\}^{\exp(\hat{\gamma}^T x_{(k)}) - \exp(\hat{\gamma}^T x_{(i)})},$$

and accordingly, the asymptotic score function for β is

$$U_{CP}(\beta) = \sum_{(i) < (k)} - \frac{\partial \left(\log \left[1 + \hat{b} \left(y_{(i)}, x_{(i)}, y_{(k)}, x_{(k)} \right) \exp \left\{ -\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)}) \right\} \right] \right)}{\partial \beta}.$$

An estimator of β , $\hat{\beta}_{CP}^*$, is provided by the solution of $U_{CP}(\beta) = 0$.

The weak convergence of $K^{-3/2} U_{CP}(\beta_0)$ can be derived by the properties of U-statistics and the asymptotic martingale representation of $n^{1/2} (\hat{S}_C - S_C)$ [19]. Since the pairwise summation in $U_{CP}(\beta)$ is over $\binom{K}{2} = O(K^2)$ items, the rate of convergence is in the order of $K^{-3/2}$. The uniform consistency of $\hat{S}_C$ entails the uniform consistency of $\hat{b}$. We see that $U_{CP}(\beta_0)$ can be decomposed to a U-statistic $\tilde{U}_{CP}(\beta_0)$ and a sum of independent and identically distributed zero-mean random variables, denoted by $s(\beta_0)$, where

$$s(\beta_0) = \sum_{(i) < (k)} \frac{(\hat{b} - b) \left(x_{(i)} - x_{(k)} \right) \left(y_{(i)} - y_{(k)} \right) \exp \left\{ -\beta_0^T \left(x_{(i)} - x_{(k)} \right) \left(y_{(i)} - y_{(k)} \right) \right\}}{\left[1 + b \exp \left\{ -\beta_0^T \left(x_{(i)} - x_{(k)} \right) \left(y_{(i)} - y_{(k)} \right) \right\} \right] \left[1 + \hat{b} \exp \left\{ -\beta_0^T \left(x_{(i)} - x_{(k)} \right) \left(y_{(i)} - y_{(k)} \right) \right\} \right]}.$$

Note that $s(\beta_0)$ can also be written in the form of a U-statistic. Given the multivariate U-statistic representation and similar to the discussion in Serfling [20], we have that $K^{-3/2} U_{CP}(\beta_0)$ converges weakly to a multivariate normal distribution with mean zero and

variance-covariance matrix Σ_2^* . By Taylor series expansion of $U_{CP}(\hat{\beta}_{CP}^*)$ around β_0 , $K^{1/2}(\hat{\beta}_{CP}^* - \beta_0)$ is asymptotically equivalent to $K^{-3/2}\Sigma_1^{*-1}U_{CP}(\beta_0)$ by Slutsky's theorem and delta method, where Σ_1^* is the Hessian matrix of the asymptotic score function $U_{CP}(\beta)$ and can be consistently estimated by

$$\hat{\Sigma}_1^* = -\frac{2}{K(K-1)} \sum_{(i)<(k)} \frac{\partial^2 \left(\log \left[1 + \hat{b} \exp \left\{ -\beta^T (x_{(i)} - x_{(k)}) (y_{(i)} - y_{(k)}) \right\} \right] \right)}{\partial \beta^2}.$$

Therefore, as $K \rightarrow \infty$, $K^{1/2}(\hat{\beta}_{CP}^* - \beta_0)$ converges weakly to a multivariate normal distribution with mean zero and variance-covariance matrix $V^* = \Sigma_1^{*-1}\Sigma_2^*\Sigma_1^{*-1}$. While the asymptotic variance-covariance Σ_2^* can be estimated by the empirical form of the U-statistic variance, the computation is rather complex. Since the asymptotic normality has been established, it is more convenient to use the bootstrap method to compute Σ_2^* and V^* .

3. Full likelihood approach

3.1. Maximum likelihood estimation

Conditional on uncensored failure times, the estimation method in Section 2 can be easily implemented, under both independent and dependent censoring assumptions. However, conditional inference is often less efficient than inference based on the full likelihood [21, 22]. In this section, we formulate the full likelihood function for survival data under the proportional likelihood ratio model by generalizing the nonparametric maximum likelihood estimation method of Luo and Tsai [4].

The density function of Y given X , $f(y|x)$, satisfies the proportional likelihood ratio model (1), and accordingly, the survival function of Y given X is

$$S(y|x) = \frac{\int_y^\infty \exp(\beta^T x u) dG(u)}{\int \exp(\beta^T x y) dG(y)}.$$

Under either independent censoring or conditional independent censoring assumption specified in Section 2, the full likelihood function of the observed data $(\tilde{y}_i, \delta_i, x_i)$ can be expressed as

$$\begin{aligned} L_F(\beta, G) &= \prod_{i=1}^n f(\tilde{y}_i|x_i)^{\delta_i} S(\tilde{y}_i|x_i)^{1-\delta_i} \\ &= \frac{\prod_{i=1}^n \left\{ \exp(\beta^T x_i \tilde{y}_i) dG(\tilde{y}_i) \right\}^{\delta_i} \left\{ \int_{\tilde{y}_i}^\infty \exp(\beta^T x_i u) dG(u) \right\}^{1-\delta_i}}{\int \exp(\beta^T x_i y) dG(y)}. \end{aligned}$$

The corresponding log full likelihood for (β, G) is

$$l_F(\beta, G) = \sum_{i=1}^n \delta_i \beta^T x_i \tilde{y}_i + \sum_{i=1}^n \delta_i \log dG(\tilde{y}_i) + \sum_{i=1}^n (1 - \delta_i) \log \left\{ \int_{\tilde{y}_i}^{\infty} \exp(\beta^T x_i u) dG(u) \right\} - \sum_{i=1}^n \log \left\{ \int \exp(\beta^T x_i y) dG(y) \right\}, \quad (3)$$

which will achieve its maximum if and only if G has jumps at the distinct uncensored and censored times, denoted by $\{y'_{(1)}, \dots, y'_{(H)}\}$. Let $p_h = dG(y'_{(h)})$ be the probability mass that the distribution function G assigns at the point $y'_{(h)}$, and $p = (p_1, \dots, p_H)$. Let $\{x'_{(1)}, \dots, x'_{(L)}\}$ be the distinct observed X values. The discrete version of the log full likelihood can be expressed as

$$l_F(\beta, p) = \beta^T \sum_{l=1}^L \sum_{h=1}^H n_{hl} x'_{(l)} y'_{(h)} + \sum_{h=1}^H n_h \log p_h + \sum_{l=1}^L \sum_{h=1}^H m_{hl} \log \left[\sum_{j=h}^H p_j \exp\{\beta^T x'_{(l)} y'_{(j)}\} \right] - \sum_{l=1}^L \lambda_l \log \left[\sum_{h=1}^H p_h \exp\{\beta^T x'_{(l)} y'_{(h)}\} \right], \quad (4)$$

where $n_{hl} = \sum_{i=1}^n I(\tilde{y}_i = y'_{(h)}, x_i = x'_{(l)}, \delta_i = 1)$, $n_h = \sum_{i=1}^n I(\tilde{y}_i = y'_{(h)}, \delta_i = 1)$, $m_{hl} = \sum_{i=1}^n I(\tilde{y}_i = y'_{(h)}, x_i = x'_{(l)}, \delta_i = 0)$, and $\lambda_l = \sum_{i=1}^n I(x_i = x'_{(l)})$. Thus, maximization of (3) is reduced to maximization of (4) with respect to p_h ($h = 1, \dots, H$) and subject to the constraints

$$\sum_{h=1}^H p_h = 1, p_h > 0 \quad (h = 1, \dots, H), \quad \|\beta\| \leq B$$

for some $B > 0$, where $\|\cdot\|$ is the Euclidean norm. We assume the true parameter vector β_0 satisfies $\|\beta_0\| \leq B$.

3.2. Profile likelihood, iterative algorithm and inference

The profile likelihood approach can be used to solve the maximization problem for (4). For any fixed β , define $p(\beta) = \arg\max_p l_F(\beta, p)$ and define the log profile likelihood $Q_F(\beta) = l_F\{\beta, p(\beta)\}$. If an estimate $\hat{\beta}_F$ maximizes the log profile likelihood $Q_F(\beta)$ in the region $\|\beta\| \leq B$,

then $\{\hat{\beta}_F, p(\hat{\beta}_F)\}$ will maximum the log full likelihood $l_F(\beta, p)$ in the region

$R_B = \{(\beta, p) : p_h \geq 0, \sum_{h=1}^H p_h = 1, \|\beta\| \leq B\}$ for some $B > 0$. Conversely, if $(\hat{\beta}_F, \hat{p})$

maximizes the log full likelihood $l_F(\beta, p)$ in the region R_B , then $\hat{\beta}_F$ will maximize the log profile likelihood $Q_F(\beta)$ in the region $\|\beta\| \leq B$. Luo and Tsai [4] suggested an iterative procedure to estimate $p(\beta)$ for complete data. We implement a similar algorithm for survival data as follows.

1. Take some initial point $p^{(0)} = (p_1^{(0)}, \dots, p_H^{(0)})$, for example, $p_h^{(0)} = n_h / \sum_{h=1}^H n_h$, to obtain

$$p_h^{(r)} = \frac{n_h}{\sum_{l=1}^L \frac{\lambda_l \exp\{\beta^T x'_{(l)} y'_{(h)}\}}{\sum_{h=1}^H p_h^{(0)} \exp\{\beta^T x'_{(l)} y'_{(h)}\}} - \sum_{l=1}^L \sum_{s=1}^h \frac{m_{sl} \exp\{\beta^T x'_{(l)} y'_{(h)}\}}{\sum_{j=s}^H p_j^{(0)} \exp\{\beta^T x'_{(l)} y'_{(j)}\}}}, h=1, \dots, H.$$

2. Normalize the point $p^{(r)} = (p_1^{(r)}, \dots, p_H^{(r)})$ to form a new point $p^{(1)} = (p_1^{(1)}, \dots, p_H^{(1)})$ where $p_h^{(1)} = p_h^{(r)} / \sum_{h=1}^H p_h^{(r)}$ for $h = 1, \dots, H$.
3. Stop if $\max_{1 \leq h \leq H} |p_h^{(1)} - p_h^{(0)}|$ is smaller than some pre-specified small number, otherwise set $p^{(0)} = p^{(1)}$ and return to Step 1.

The estimate $\hat{\beta}_F$ is then obtained by maximizing $Q_F(\beta)$ from a quasi-Newton method. Define $\hat{G}(y) = \sum_{h=1}^H \hat{p}_h I(y'_{(h)} \leq y)$. Let (β_0, G_0) be the true parameter values. Following the lines in Luo and Tsai [4], we introduce two regularity conditions to establish the asymptotic properties of the nonparametric maximum likelihood estimators $\{\hat{\beta}_F, \hat{G}(y)\}$.

Condition 1. The norm of the covariate vector $\|X\| \leq X_0$ for some positive constant X_0 , and covariance of the covariate vector, $\text{cov}(X)$, is of full rank.

Condition 2. The lower limit l_G and the upper limit u_G of the support of G satisfy $-\infty < l_G < u_G < \infty$.

In particular, if $\|\beta_0\| \leq B$, $\hat{\beta}_F$ is a consistent estimator of β_0 , and as $n \rightarrow \infty$, $n^{1/2}(\hat{\beta}_F - \beta_0)$ converges weakly to a multivariate normal distribution with mean zero and covariance matrix V_F . Generally, the covariance matrix V_F has no explicit form. Given the established weak convergence under the regularity conditions, a simple bootstrap resampling procedure is valid to estimate the asymptotic variance of $\hat{\beta}_F$. For the baseline distribution function G , $\hat{G}(y)$ is a consistent estimator, and as $n \rightarrow \infty$, $n^{1/2}\|\hat{G} - G_0\|_\infty$ is bounded in probability, where $\|\cdot\|_\infty$ denotes the supremum norm on the real line. This property implies that, for any $y_1, \dots, y_m$, the random vector $n^{1/2}\{\hat{G}(y_1) - G_0(y_1), \dots, \hat{G}(y_m) - G_0(y_m)\}$ follows approximately an m -variate normal distribution with mean zero.

For testing a null hypothesis $H_0: \beta = \beta_0$, we can construct a likelihood ratio test statistic $LRT = -2[l_F(\beta_0, \tilde{G}) - l_F(\hat{\beta}_F, p(\hat{\beta}_F))]$, where $\tilde{G}$ is the nonparametric maximum likelihood estimator of G under the null hypothesis. By the results of Murphy and van der Vaart [23, 24], the likelihood ratio test statistic has an asymptotic χ_p^2 distribution. Alternatively, we can construct a Wald test statistic, which also has a limiting χ_p^2 distribution.

4. Simulations

Simulation studies are conducted to evaluate the finite-sample performances of the estimation and inference of the conditional likelihood and the full likelihood approaches for right-censored survival data, under various scenarios. We consider an Exponential failure time as outcome.

First, we assess the performance of the conditional likelihood approach. A discrete covariate X_1 is generated from a $\text{Ber}(0.6)$, and a continuous covariate X_2 is generated from a $U(0, 1)$. The finite-dimensional parameters of interest are $(\beta_1, \beta_2) = (-1, -1)$, and baseline density $g(y)$ is taken to be the unit exponential density function. In Scenario 1, the censoring time C is generated independently from a uniform distribution corresponding to censoring percentage of around 25%. In Scenario 2, the censoring time depends on covariate, while the outcome and covariate data follow the same distributions as those in Scenario 1. The censoring time C is generated from a Weibull distribution with shape parameter $\alpha = 2$ and rate parameter μ depends on covariate X_2 by $\mu = \{2/\exp(\gamma X_2)\}^{1/2}$ where $\gamma = -1$. The censoring proportion is around 20%. For each scenario, we use sample sizes $n = 100, 200$ and 300, and repeat the simulation 1000 times. The simulation results are shown in Table 1, which summarizes the empirical bias, empirical standard error, average bootstrap standard error and 95% empirical coverage probability of the estimated parameters. The proposed estimators from the conditional likelihood perform well with small biases. Bootstrap variance estimates are in good agreement with empirical variances and coverage probabilities are close to the nominal value, which indicates the inference based on the conditional likelihood approach is reasonably good.

We then compare the performance of the full likelihood estimator to that of the conditional likelihood estimator under the proportional likelihood ratio model with dependent censoring. Covariate X_1 follows a $U(0, 1)$ with a regression coefficient $\beta_1 = -1$, and we consider two settings of baseline density $g(y)$. In setting 1, $g(y) = \exp(-y)$, $y > 0$, and the response is a positive exponential distributed random variable. In setting 2, $g(y) = (2/\pi)^{(1/2)} \Phi(0.5)^{-1} \exp\{-2(y - 1/4)^2\}$, $y > 0$, and the response is a positive continuous random variable of a half-normal distribution with an exponential tilting term. The censoring time C is generated from a Weibull distribution with shape parameter $\alpha = 2$ and rate parameter μ depends on the covariate X_1 by $\mu = \{2/\exp(\gamma X_1)\}^{1/2}$ where $\gamma = -1$, and the censoring proportion is 20% – 25%. We generate 1000 independent datasets with sample sizes 50, 100 and 200 respectively. The variance estimation is obtained by bootstrap method, where 500 bootstrap samples are used. Table 2 summarizes the simulation results for both settings, including the empirical bias, empirical standard error, average bootstrap standard error and 95% empirical coverage probability of estimators of β_1 from the conditional and full likelihood methods. The bias of the full likelihood estimator is small, and average bootstrap standard error is very close to empirical standard error. Compared with the conditional likelihood estimator, the full likelihood estimator has a smaller bias and variability. As sample size increases, the performance of the conditional likelihood estimator gets close to the full likelihood estimator in terms of bias. However, the standard error of the full likelihood method is still about 2/3 times that of the conditional method, because the conditional method does not fully utilize the information of observed data. We also compare the computation time of

each method in setting 1. For $n = 50, 100$ and 200 , it takes on average $0.2, 0.9$, and 1.5 seconds, respectively to compute the proposed estimator for the conditional likelihood approach, and $14.5, 70.3$, and 348.7 seconds for the full likelihood approach, which suggests the use of the conditional likelihood approach leads to a substantial improvement in computational efficiency.

5. Application

Bone marrow transplantation is a standard treatment for acute leukemia, and recovery following bone marrow transplant is a complex process. Prognosis for recovery may depends on risk factors known at the time of transplantation, such as the stage of initial disease, the time from diagnosis to transplantation, etc. A study was conducted to illustrate this process and characterize disease-free survival for bone marrow transplantation patients of acute leukemia. Details of the study are found in Copelan *et al.* [25]. In this multi-center clinical trial, a total of 137 patients were treated with bone marrow transplants at one of four hospitals (two in U.S. and two in Australia). They were followed for relapse or death, with the maximum follow-up of 7 years. There were 42 patients who relapsed and 41 who died while in remission. The censoring percentage was around 40%. In addition to time to relapse or death, other potential risk factors were measured at the time of transplantation. Particularly, patients were grouped into risk categories based on their initial disease status at the time of transplantation as follows: 38 acute lymphoblastic leukemia (ALL) patients, 54 acute myelotic leukemia (AML) low-risk patients, and 45 acute myelotic leukemia (AML) high-risk patients. Patients at the two hospitals in Australia were given a graft-versus-host prophylactic combining methotrexate (MTX) with cyclosporine and possibly methylprednisolone. Patients at the two hospitals in U.S. were not given methotrexate but rather a combination of cyclosporine and methylprednisolone.

The interest is in studying the effects of initial disease status at the time of transplantation and MTX status on disease-free survival. In our analysis, Y denotes time to relapse or death, X_1 denotes the indicator of AML low-risk, X_2 denotes the indicator of AML high-risk, and X_3 denotes the indicator of MTX status. A patient has ALL if $X_1 = 0$ and $X_2 = 0$. We first consider the Cox regression model and check the proportional hazards assumption for the two factors using scaled Schoenfeld residual plot. It suggests the validity of the proportional hazards assumption for indicators of initial disease status (X_1, X_2) as shown in Fig. 2, but non-proportionality for MTX status X_3 as shown in Fig. 3. Additionally, one way to formally test the proportional hazards assumption for MTX is to create a time-dependent covariate $X_4(t) = X_3 \times \ln(t)$ and consider a Cox model with X_3 and $X_4(t)$. Testing for time-variation for MTX reduces to verify the null hypothesis that the Cox regression coefficient for $X_4(t)$ equals 0. The p -value for such test is 0.04, which implies the violation of proportional hazards assumption for MTX and a simple Cox regression model with (X_1, X_2, X_3) as covariates is inappropriate. A Cox regression model stratified by MTX status could be used to estimate the effect of initial disease status, if the primary interest is to compare the three groups of ALL, AML low-risk and AML high-risk patients. Based on such a stratified Cox regression model, adjusting for MTX status, the estimated hazard ratio is 0.61 comparing AML low-risk patients to ALL patients with 95% confidence interval (0.34, 1.08) and p -value 0.09, and the estimated hazard ratio is 1.55 comparing AML high-risk

patients to ALL patients with 95% confidence interval (0.91, 2.63) and p -value 0.10. Moreover, from the Kaplan-Meier estimation, the estimated median survival time for patients being given MTX is 192 days, and that for patients without being given MTX is 625 days. The log-rank test comparing survival by MTX status has a p -value of 0.09. When this differential effect of MTX status on survival is also of interest, the proposed proportional likelihood ratio model has the advantage of not requiring the proportional hazards assumption nor modeling the non-proportional hazard for MTX status.

Therefore, we investigate the effects of initial disease status and MTX status under the proportional likelihood ratio model. We assume independent censoring and estimate regression coefficients ($\beta_1, \beta_2, \beta_3$) in the proportional likelihood ratio model from the proposed conditional and full likelihood approaches. The standard error is obtained from 500 bootstrap samples. Table 3 summarizes the estimation results. Both approaches yield similar parameter estimates, but the full likelihood estimators are more efficient with smaller standard errors than the conditional likelihood estimators. Based on the full likelihood estimates, the likelihood ratio that a patient's time to relapse or death is prolonged by sixty days will increase by a factor $\exp(0.00144 \times 60) = 1.09$, if the patient has AML low-risk instead of ALL. The likelihood ratio that a patient's survival time is prolonged by sixty days will decrease by a factor $\exp(-0.00291 \times 60) = 0.84$, if the patient has in AML high-risk instead of ALL, and such likelihood ratio will decrease by a factor $\exp(-0.00527 \times 60) = 0.73$, should the patient was given MTX. Compared with ALL, AML low-risk is associated with a lower risk of relapse or death (p -value=0.20) and AML high-risk is associated with a higher risk (p -value=0.06), but the relationship is not statistically significant. MTX is significantly associated with a higher risk of relapse or death. The results on survival pattern are consistent with those by the Cox regression model and log-rank test. Overall, we provide precise information about the magnitudes of the effects of initial disease status and MTX status, and demonstrate that MTX status is a significant factor on disease-free survival (p -value<0.01).

In addition to stratification and the use of proportional likelihood ratio model, there are other options to handle non-proportionality for MTX within the Cox model framework. A common and direct alternative is to fit piecewise constant Cox model in pre-specified disjunctive time intervals. By studying scatterplot smoother of the scaled Schoenfeld residuals versus time for MTX in Fig. 3, we choose a "change point" to be 600 days. Based on this model, it shows that, up to 600 days, the estimated hazard ratio is 1.61 comparing patients given MTX to those without given MTX (p -value=0.06), whereas after 600 days, the estimated hazard ratio is 0.36 (p -value=0.32). We also check for the proportionality assumption within the two time intervals using formal test by fitting Cox models with two time-dependent covariates, and the corresponding p -values are 0.11 and 0.23 respectively. Thus, there is no need to further divide the subintervals of time.

However, the main problem of this method is that it requires a pre-specification of the time-variation by a functional form or a partitioning of time-axis, and in some cases the improvement of fit by including time-variation may not justify the increase in model complexity. For continuous covariates, a flexible method for dynamic Cox modeling based on fractional polynomials was proposed by Berger *et al.* in [15] to detect and simultaneously

analyze time-varying effect structures, and is straightforward to apply. In presence of non-proportional hazards, weighted estimation in Cox model is a parsimonious alternative without including additional parameters for time-dependent effects. The suitability of the weighted approach for estimating average hazard ratios was first demonstrated by Schemper in [16], and this method is applied to analyze the effects of the prognostic factors for the leukemia survival data. Based on unadjusted and adjusted analysis results summarized in Table 4, we find that for the non-proportional MTX effect, it makes much difference in terms of the hazard ratio and the associated p -value whether unweighted or weighted approach is employed. But for the proportional factors, there is no much difference. Although it reduces the complexity of an analysis, the main goal of this method is to provide a summary of the average effect and we can not expect substantial advantages of better modeling time dependence. Further, to provide a better visualization for comparing the various methods under consideration, we present the estimated survival curves for the MTX=0 and MTX=1 groups in Figure 4, using the aforementioned Cox model, stratified Cox model by MTX, piecewise constant Cox model and proportional likelihood ratio model (PLRM). Among these models, for both groups survival curves estimated from the proportional likelihood ratio model agree well with survival curves estimated from the Cox model stratified by MTX status, which may indicate a good fit of the proportional likelihood ratio model for the leukemia survival data. The straightforward Cox model and piecewise constant Cox model seem to fit the data relatively poorly.

6. Discussion

Regression analyses for survival data based on semiparametric transformation model have been extensively studied in literature. These methods either rely on a strong assumption of proportionality as in the Cox regression model or generally suffer from a lack of ready interpretation for the regression parameters. As an alternative, the proportional likelihood ratio model considered in this paper assumes an explicit probability model for survival data up to an unspecified infinite-dimensional distribution parameter, and flexibly models the relationship between survival outcome and covariates, resulting in relatively easy-to-interpret regression coefficients. The development of such model for survival data could be considered as practical, especially when many or all covariates exhibit non-proportional hazards. Several other approaches to cope with non-proportional hazards are discussed. Stratification is not a very attractive option because it does not allow estimation and test of effect for a covariate with non-proportional hazards. If a detailed exploration of time-dependence of covariate effect is of scientific interest and the amount of data information is sufficient, inclusion of time-dependent covariate terms by functional polynomials or separate Cox models for different time periods are appropriate, but may result in more complex models, an increased risk of over-fitting or decreased power. The estimation of average hazard ratios is preferable if only a single summary measure of the non-proportional covariate effect is desirable.

We have proposed two likelihood approaches for analyzing survival data under the proportional likelihood ratio model. By taking advantage of the invariance property, based on a conditional likelihood, the target regression parameter vector is estimated using a pairwise pseudo-likelihood method. We also generalize Luo & Tsai's iterative algorithm

with the profile likelihood method, and jointly estimate the target regression parameter vector and the unknown baseline distribution function parameter from a full likelihood approach. Inference based on the full likelihood is more efficient than that based on the conditional likelihood because the censored data are fully utilized in the iterative algorithm. In addition, the full likelihood approach is more robust to various censoring patterns, because it does not require to estimate the survival function of censoring time under dependent censoring. Although methods for modeling dependent censoring process are well established and will typically be straightforward, the consistency of the estimator of regression parameter vector relies on the consistency of the survival function estimator of censoring time. Therefore, the misspecification of the model for censoring may result in bias for the target regression parameter vector. We consider proportional hazards model for censoring time, while a class of semiparametric or parametric model can be used when it is appropriate.

Further, the full likelihood can simultaneously estimate the regression parameter vector and the baseline distribution function, but the conditional likelihood cannot estimate the entire distribution. In many applications, the estimation of the conditional mean or other quantities of outcome is of interest. The conditional likelihood approach is not applicable under such situations. The full likelihood estimators are more informative, and have estimation efficiency and robustness advantages over the conditional likelihood estimators. However, the conditional approach is appealing due to its straightforward computation under the proportional likelihood ratio model. Our simulations indicate that the condition likelihood method performs quite favorably compared with the full likelihood method for estimating the regression parameter vector, and has a substantially lower computation burden.

As suggested by one reviewer, it is possible to improve the computational efficiency of the full likelihood method by a wise choice of initial values. In the original iterative algorithm, we start with arbitrary initial estimates $p_h^{(0)} = n_h / \sum_{h=1}^H n_h$, $h = 1, \dots, H$ for the baseline distribution function $G(y)$. Alternatively, we can use $\hat{\beta}_{CP}^*$ the maximum conditional likelihood estimator of β , as an initial estimate of β and

$$\sum_{i=1}^n \exp(-\tilde{y}_i \hat{\beta}_{CP}^{*T} x_i) I(\tilde{y}_i \leq y) / \left\{ \sum_{i=1}^n \exp(-\tilde{y}_i \hat{\beta}_{CP}^{*T} x_i) \right\}$$

as an initial estimate of $G(y)$. Our empirical experience suggests that this algorithm is computationally more efficient than the original algorithm. The average computation time of the full likelihood approach for simulation setting 1 in Table 2 reduces to 3.8, 16.5 and 107.2 seconds for $n = 50, 100$ and 200, respectively. Another option for initial values is to perform an initial analysis with a flexible parametric model for $G(y)$, take the resulting estimate of as an initial estimate in the full semiparametric likelihood fitting, and construct an initial estimate of $G(y)$ by redistributing the mass of the parametric $G(y)$ as point masses on the sample points.

We have illustrated the proposed methods under the proportional likelihood ratio model by applying them to a leukemia survival data set where the proportionality assumption is violated to some extent. It would be interesting to develop a formal goodness-of-fit test to assess the adequacy of the proportional likelihood ratio model. The investigation of such a

procedure and other model diagnostics is beyond the scope of this paper but will be explored in future research.

Acknowledgements

This work was supported in part by the Cancer Center Support Grant from the National Cancer Institute awarded to the Harold C. Simmons Cancer Center at the University of Texas Southwestern Medical Center. The author thank the editor, an associate editor and two reviewers for their constructive comments that have greatly improved the initial version of this paper.

References

1. Cheng SC, Wei LJ, Ying Z. Analysis of transformation models with censored data. *Biometrika*. 1995; 82:835–845.
2. Cox DR. Regression models and life tables (with discussion. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*. 1972; 34:184–220.
3. Bennett S. Analysis of survival data by the proportional odds model. *Statistics in Medicine*. 1983; 2:273–277. [PubMed: 6648142]
4. Luo X, Tsai WY. A proportional likelihood ratio model. *Biometrika*. 2012; 99:211–222.
5. Liang KY, Qin J. Regression analysis under non-standard situations: a pairwise pseudolikelihood approach. *Journal of the Royal Statistical Society. Series B (Methodological)*. 2000; 62:773–786.
6. Kalbfleisch, JD.; Prentice, RL. *The Statistical Analysis of Failure Time Data*. 2nd ed. John Wiley and Sons; New York: 2002.
7. Huang A, Rathouz PJ. Proportional likelihood ratio models for mean regression. *Biometrika*. 2012; 99:223–229. [PubMed: 24421412]
8. Nelder JA, Wedderburn RWM. Generalized linear models. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*. 1972; 135:370–384.
9. Qin J, Zhang B. A goodness-of-fit test for logistic regression models based on case-control data. *Biometrika*. 1997; 84:609–618.
10. Fokianos K, Kaimi I. On the effect of misspecifying the density ratio model. *Annals of the Institute of Statistical Mathematics*. 2006; 58:475–497.
11. Gilbert PB, Lele SR, Vardi Y. Maximum likelihood estimation in semiparametric selection bias models with application to AIDS vaccine trials. *Biometrika*. 1999; 86:27–43.
12. Ichimura H. Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics*. 1993; 58:71–120.
13. Rathouz PJ, Gao LP. Generalized linear models with unspecified reference distribution. *Biostatistics*. 2009; 10:205–218. [PubMed: 18824517]
14. Wang C, Tan Z, Louis TA. Exponential tilt models for two-group comparison with censored data. *Journal of Statistical Planning and Inference*. 2011; 141:1102–1117.
15. Berger U, Schafer J, Ulm K. Dynamic Cox modelling based on fractional polynomials: time-variations in gastric cancer prognosis. *Statistics in Medicine*. 2003; 22:1163–1180. [PubMed: 12652560]
16. Schemper M. Cox analysis of survival data with non-proportional hazard functions. *The Statistician*. 1992; 41:455–465.
17. Chan KC. Nuisance parameter elimination for proportional likelihood ratio models with nonignorable missingness and random truncation. *Biometrika*. 2013; 100(1):269–276.
18. Zeng D. Estimating marginal survival function by adjusting for dependent censoring using many covariates. *The Annals of Statistics*. 2004; 32:1533–1555.
19. Van der Vaart, AW. *Asymptotic Statistics*. Cambridge University Press; Cambridge: 1998.
20. Serfling, RJ. *Approximation Theorems of Mathematical Statistics*. Wiley; New York: 1980.
21. Wang MC. Nonparametric estimation from cross-sectional survival data. *Journal of the American Statistical Association*. 1991; 86:130–143.

22. Zeng D, Lin DY. Maximum likelihood estimation in semiparametric regression models with censored data. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*. 2007; 69:507–564.
23. Murphy SA, van der Vaart AW. Semiparametric likelihood ratio inference. *Annals of Statistics*. 1997; 25:1471–1509.
24. Murphy SA, van der Vaart AW. On the profile likelihood. *Journal of the American Statistical Association*. 2000; 95:449–465.
25. Copelan EA, Biggs JC, Thompson JM, Crilley P, Szer J, Klein JP, Kapoor N, Avalos BR, Cunningham I, Atkinson K, Downs K, Harmon GS, Daly MB, Brodsky I, Bulova SI, Tutschka PJ. Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with Bu/Cy. *Blood*. 1991; 78:838–843. [PubMed: 1859895]

Figure 1.

Hazard functions of group 1 and group 2 for two situations of group effects. (a) proportional hazards: $\text{Exp}(1/2)$ for group 1 and $\text{Exp}(1)$ for group 2. (b) diverging hazards: $\text{Gamma}(a = 2, \theta = 1/2)$ for group 1 and $\text{Gamma}(a = 2, \theta = 1)$ for group 2.

Figure 2.
Scaled Schoenfeld residual plot of indicators of initial disease status.

Figure 3.
Scaled Schoenfeld residual plot of MTX status.

Figure 4.

Estimated survival curves for MTX=0 and MTX=1 groups using Cox model (solid), stratified Cox model (dash), piecewise constant Cox model (dash-dot), and proportional likelihood ratio model (PLRM, dot).

Table 1

Simulation results based on conditional likelihood approach

n	Censoring	Bias $\times 10^3$		SSE $\times 10^3$		SEB $\times 10^3$		95%CP	
		β_1	β_2	β_1	β_2	β_1	β_2	β_1	β_2
100	Independent	-27	-94	880	1466	885	1471	94	95
	Dependent	-26	-150	726	1193	732	1198	96	94
200	Independent	-51	-42	591	968	595	973	93	94
	Dependent	-29	-54	492	738	498	745	94	95
300	Independent	-24	-39	482	799	488	805	94	96
	Dependent	-18	-34	366	595	371	602	95	96

Bias is the empirical bias; SSE is the empirical standard error; SEB is the average bootstrap standard error; CP is the empirical coverage probability.

Table 2

Simulation results comparing conditional and full likelihood approaches.

<i>n</i>	Approach	Bias $\times 10^3$	SSE $\times 10^3$	SEB $\times 10^3$	95%CP
Setting 1					
50	Con. Lik.	-145	1588	1592	94
	Full Lik.	-53	1263	1269	93
100	Con. Lik.	-116	1096	1102	94
	Full Lik.	-34	756	761	94
200	Con. Lik.	-22	669	673	95
	Full Lik.	-25	443	448	94
Setting 2					
50	Con. Lik.	-157	1953	1958	93
	Full Lik.	-59	1297	1303	93
100	Con. Lik.	-130	1331	1338	95
	Full Lik.	-41	883	887	94
200	Con. Lik.	-26	845	850	95
	Full Lik.	-23	571	574	95

Table 3

Estimation results under the proportional likelihood ratio model for the leukemia survival data.

Parameter (Factor)	Approach	Estimate $\times 10^3$	SEB $\times 10^3$	95%CI $\times 10^3$	<i>p</i> -value
β_1 (AML low)	Con. Lik.	1.42	1.25	(-1.03, 3.87)	0.26
	Full Lik.	1.44	1.13	(-0.77, 3.65)	0.20
β_2 (AML high)	Con. Lik.	-2.89	1.62	(-6.07, 0.29)	0.07
	Full Lik.	-2.91	1.53	(-5.91, 0.09)	0.06
β_3 (MTX)	Con. Lik.	-5.24	2.03	(-9.22, -1.26)	0.01
	Full Lik.	-5.27	1.84	(-8.88, -1.66)	< 0.01

SEB is the bootstrap standard error; CI is the confidence interval.

Table 4

Hazard ratios (p -values) of the prognostic factors for the leukemia survival data by unadjusted/adjusted Cox models using unweighted/weighted estimation.

Factor	Unadjusted Unweighted	Weighted	Adjusted Unweighted	Weighted
AML low	0.56 (0.04)	0.59 (0.06)	0.60 (0.08)	0.65 (0.14)
AML high	1.47 (0.15)	1.51 (0.14)	1.53 (0.11)	1.62 (0.08)
MTX	1.49 (0.09)	1.72 (0.03)	1.41 (0.15)	1.64 (0.06)