東海大學

統計研究所

碩士論文

半參數轉換模型下兩存活中位數差異之區間構建

Confidence interval for the difference between two median survival times with semiparametric transformation models

指導教授：張玉媚　博士

學　　生：湯語心

中華民國一〇六年七月

東海大學碩士班研究生

論文口試委員審定書

統計學系碩士班湯語心君所提之論文

Confidence interval for the difference between two median survival times with semiparametric transformation models

---

經本委員會審議，認為符合碩士資格標準。

論文口試委員召集人 陳春樹 （簽章）

委員 張文昭

江弘文

中華民國 106 年 07 月 13 日

致謝詞

在論文完成之際，也代表著研究生生涯告一段落。此篇論文的完成，要感謝許多在碩士生涯中幫助過我的人，其中最要感謝的是我的指導教授張玉媚博士，感謝老師在這幾年悉心指導，每當我遇到困難或疑惑時，老師總會不厭其煩地協助我、幫忙我，給予我良好的意見，讓我在這段學習及撰寫的過程中收穫良多，不僅是知識上的吸收，更讓我成長不少。再來我也十分感謝沈葆聖博士，在這些年來給我的指導與協助，使我的思路更為清晰明確。

論文審查及口試期間內，承蒙沈葆聖博士與陳春樹博士，提供許多寶貴的意見與建議，使得本篇論文更加完整與充實，在此致上最深的謝意。此外我亦感謝東海與統計系上所有的老師與助教們，有你們的教導、協助與照顧，才有學生今日之成就與成長。

接著要感謝研究室的同學們，大家一起研究、一起努力，並且互相扶持，共同度過種種的難關，很幸運能遇到你們，讓我在求學期間有著充滿快樂及充實的美好時光。

最後，感謝默默支持與陪伴我的家人，在這論文的撰寫時刻，無怨無悔地包容與鼓勵，給予我學習機會與發揮空間，使我能專注在論文上，並順利完成學業。僅將本論文獻給一路上幫助與關心我的人。

Abstract

In medical studies, we usually are interested in comparing the treatment effects of the drug according to the difference of two median survival times. In this paper, based on the estimating equation estimator (Chen et al., 2002) and the maximum likelihood estimator (Zeng and Lin, 2006, Chen, 2009), we propose two methods for constructing conditional confidence interval for the difference of two median survival times given the covariates under the semiparametric transformation models. The effectiveness of the associated coverage probability and the average length of the interval are investigated via a simulation study. An application of the proposed method is illustrated with a real data set.

Key words: Confidence interval; Median survival time; Estimating equation; Maximum likelihood; Weighted Breslow; Semiparametric transformation mode

Contents

1. Introduction

In medical studies, Cox proportional hazards model (Cox, 1972) is the most commonly used method to analyze the survivor function of patients when the right-censored survival data are accompanied with covariates which are associated with patients' physiological conditions. Although the estimated survival function can be used to identify covariates that affect survival times when expressing the results of clinical trials, it would be better to consider the hazard ratio together with a measure of time, such as median (or quantile) survival time, which provides the additional information regarding the point at which half the subjects have experienced the event. Thus, when one is assessing the treatment effects of a drug, how to construct the confidence intervals for the median survival time is essential. Under the Cox proportional hazards model (Cox, 1972), many methods have been proposed to construct confidence intervals of median/mean survival time. Dabrowska and Doksum (1989) considered a confidence-interval estimation of a conditional median survival time given a value of covariate vector. Burr and Doss (1993) constructed confidence bands for the median survival time as a function of the covariates. For comparing two survival functions, Zucker (1998) proposed the statistical tests for comparing two treatments based on the difference between two restricted means with covariates adjusted based on the stratified Cox model, under which the ratio of hazards or cumulative hazards for subjects with the same covariates in two treatment groups would be independent of the covariates. Kim (2001) suggested two confidence intervals for the difference of median survival times, one is constructed only for the patients with baseline covariates and the other is developed with covariates adjusted. Chen and Chang (2007) extended the work in Su and Wei (1993) and suggested covariates-dependent confidence intervals for the difference or ratio of two median

lifetimes under both stratified and treatment-specific Cox models. Chen et al. (2015) further developed confidence bands for the difference between two median survival times as a function of the associated covariates under the stratified and treatment-specific Cox models, respectively.

The proportionality assumption of the Cox model can be violated in practice. The class of semiparametric transformation model, which includes Cox and proportional odds model (Bennett, 1983) as special cases and allows various nonproportional hazard structures, have received tremendous recent attention (e.g. Cheng et al., 1995, 1997; Murphy et al., 1997; Scharfstein et al., 1998; Cai et al., 2002; Chen et al., 2002; Tsodikov, 2003; Kosorok et al., 2004; Lu and Ying, 2004; Lu and Tsiatis, 2006; Zeng and Lin, 2006, 2007). Cheng et al. (1995) proposed and justified a general estimation method for linear transformation models with right-censored data. The method was further developed in Cheng et al. (1997), Fine et al. (1998) and Cai et al. (2000). A key step in their approach is the estimation of the survival function for the censoring variable by the Kaplan–Meier estimator, which relies on the assumption that the censoring variable is independent of the covariates. In practice, however, the independent assumption is often too restrictive, even for randomized clinical trials. Unlike Cox's partial likelihood approach, this estimation method fails when the independence assumption is violated. Chen et al.'s (2002) estimator is valid under the assumption that failure time is conditionally independent of censoring time given covariates and it is the same as the Cox partial likelihood estimator in the case of proportional hazards model. Chen et al.'s (2002) estimator is obtained through the estimating equations (EE) and its asymptotic variance has a closed-form expression even though it is not efficient. Using the nonparametric maximum likelihood methods, Zeng and Lin (2006) proposed an efficient estimation of a broad class of

transformation models which can accommodate time-varying covariates and recurrent events. Chen (2009) further showed that intruding weights into the Breslow type estimator can lead to nonparametric maximum likelihood estimation (MLE) of Zeng and Lin (2006) and the weighted Breslow-type estimator suggests an iterative reweighting algorithm for nonparametric maximum likelihood estimation. Chen et al. (2012) introduced time-dependent residuals for semiparametric transformation models and used the cumulative sums of the residuals to construct graphical and numerical procedures for model checking.

In this paper, we consider the problem of constructing the conditional confidence interval given the associated covariates for the difference of two median survival times under the semiparametric transformation models. Based on the conditional confidence intervals, we can identify the possible range of covariates over which the two groups would provide different median survival times.

In Section 2, based on the EE estimator (Chen et al., 2002) and the MLE (Zeng and Lin, 2006, Chen, 2009), we consider construction of the conditional confidence interval for the difference of two median survival times given the covariates under the semiparametric transformation model. In Section 3, a simulation study is conducted to investigate the coverage probability and the expected length of the conditional intervals. In Section 4, the proposed method is illustrated using a data set in a two-arm lung cancer study (Ying et al., 1995). Conclusions and discussions on the application of the confidence intervals are finally given in Section 5.

2. The proposed confidence intervals

2.1 Semiparametric transformation model

Let $\{(T_i, C_i),\ i = 1, \dots, n\}$ be failure time and censoring time of the n patients. Denote the associated $q \times 1$ covariates vector $Z_i = [z_{1i}, z_{2i}, \dots, z_{qi}]^T$. Without loss of generality, let $z_{1i} = 0, 1$, denote two treatment groups. Assume that failure time and censoring time are conditionally independent given the covariate Z_i. The observed right-censored lifetime data are $(\tilde{T}_i, \delta_i, Z_i)$, where $\tilde{T}_i = \min(T_i, C_i)$, and $\delta_i = I(T_i \le C_i),\ i = 1, \dots, n$.

Semiparametric transformation models specify the corresponding cumulative intensity of T_i, given the covariate Z_i, as

$$\Lambda(t|Z_i) = G\{\Lambda(t)\exp(\beta^T Z_i)\}, \tag{1}$$

where $\Lambda(\cdot)$ is an unspecified increasing function, β is a $q \times 1$ vector of regression parameters and G is a specified transformation function that is continuously differentiable and strictly increasing with $G(0) = 0$, $G'(0) > 0$ and $G(\infty) = \infty$. The Box-Cox transformation model is the most commonly used model:

$$G(x) = \{(1 + x)^\rho - 1\}/\rho \quad (\rho \ge 0)$$

with $\rho = 0$ corresponding to $G(x) = \log(1 + x)$. The other model is the class of logarithmic transformations:

$$G(x) = \log(1 + rx)/r \quad (r \ge 0)$$

with $r = 0$ corresponding to $G(x) = x$. The choice of $\rho = 1$ or $r = 0$ yields the proportional hazards model; while the choice of $\rho = 0$ or $r = 1$ yields the proportional odds model.

Under model (1), the conditional survival function can be written as

$$S(t|Z_i) = \exp(-\Lambda(t|Z_i)).$$

Let $S(t|z_{1i} = k, x_i),\ k = 0, 1$, denote the survival function of the treatment group k,

given covariate $x_i = (z_{2i}, z_{3i}, \dots, z_{qi})$. The $100p$th percentile of the treatment k is given by

$$\xi_p^k(x_i) = \sup\{t: S(t|z_{1i} = k, x_i) \geq 1 - p\},\ k = 0, 1.$$

When $p = 0.5$, $\xi_{0.5}^k(x_i)$ is the median survival time for treatment k. For conveniences, we use $\xi_k(x_i)$ to represent the median survival time for treatment k in the whole article. Therefore, given the covariate x_i, the $100(1-\alpha)\%$ conditional confidence intervals for $\Delta(x_i) = \xi_0(x_i) - \xi_1(x_i)$ can be constructed.

2.2 The method based on the estimator of Chen et al. (2002)

First, we consider constructing confidence intervals based on EE estimator of Chen et al.'s (2002). Note that model (1) is equivalent to

$$H(t) = -\beta^T Z_i + \varepsilon_i,\ i = 1, \dots, n,$$

where $H(t) = \log\Lambda(t)$, $\varepsilon_i = \log G^{-1}(-\log U)$ is independent of Z_i, U has a uniform distribution on (0, 1). Let $Y_i(t) = I(\tilde{T}_i \geq t)$ denote the at-risk process and $N_i(t) = I(\tilde{T}_i \leq t, \delta_i = 1)$ be the observed counting process. For a random sample of n subjects, the data consist of $\{N_i(t), Y_i(t), Z_i: i = 1, \dots, n, 0 \leq t \leq \tau\}$, where τ denotes the end point of the study. Let $\lambda_\varepsilon(\cdot)$ and $\Lambda_\varepsilon(\cdot)$ be the known hazard and cumulative hazard function of ε, respectively. Following the usual counting process notation, let

$$M_i(t) = N_i(t) - \int_0^t Y_i(u) d\Lambda_\varepsilon\{\beta^T Z_i + H(u)\},$$

where $\Lambda_\varepsilon(t) = G(e^t)$. The β and H are evaluated at their true values when the assumed model holds. Under model (1), $M_i(t)$ is a martingale process. The EE proposed by Chen et al. (2002) are

$$U(\beta, H) \equiv \sum_{i=1}^n \int_0^\infty Z_i[dN_i(t) - Y_i(t) d\Lambda_\varepsilon\{\beta^T Z_i + H(t)\}] = 0, \qquad (2)$$

and

$$\sum_{i=1}^{n}[dN_i(t) - Y_i(t)d\Lambda_\varepsilon\{\beta^T Z_i + H(t)\}] = 0 \quad (t \geq 0), \tag{3}$$

where H is a nondecreasing function satisfying $H(0) = -\infty$. This requirement ensures that $\Lambda\{a + H(0)\} = 0$ for any finite a. Let $\hat{\beta}$ and $\widehat{H}(t, \hat{\beta})$ denote the solutions of (2) and (3). Note that $\widehat{H}(t, \hat{\beta})$ is a step functions in t that jumps only at the observed failure times t_*.

For the special case of the Cox model, $\lambda_\varepsilon(t) = \exp(t)$ and it then follows from (2) and (3) that the estimate of $\hat{\beta}$ satisfies the following equation:

$$\sum_{i=1}^{n}\int_0^\infty \left\{Z_i - \frac{\sum_{j=1}^{n} Z_j Y_j(t)\exp(\beta^T Z_j)}{\sum_{j=1}^{n} Y_j(t)\exp(\beta^T Z_j)}\right\} dN_i(t) = 0,$$

which is precisely the Cox partial likelihood score equation for right-censored data. Equations (2) and (3) suggest the following iterative algorithms for computing $\hat{\beta}$ and $\widehat{H}(t, \hat{\beta})$:

Step 0: Choose an initial value of $\hat{\beta}$ denoted by $\hat{\beta}^{(0)}$.

Step 1: Let $t_1 < t_2 < \cdots < t_D$ denote the order failure time. We obtain $\widehat{H}^{(0)}(t_1, \hat{\beta}^{(0)})$ by solving

$$\sum_{i=1}^{n} Y_i(t_1)\Lambda_\varepsilon\{\beta^T Z_i + H(t_1)\} = 1,$$

with $\beta = \hat{\beta}^{(0)}$. Then, obtain $\widehat{H}^{(0)}(t_k)$, for $k = 2, \ldots, D$, one-by-one by solving the equation

$$H(t_k) = H(t_k -) + \frac{1}{\sum_{i=1}^{n} Y_i(t_k)\lambda_\varepsilon\{\beta^T Z_i + H(t_k -)\}},$$

with $\beta = \hat{\beta}^{(0)}$.

Step 2: Obtain a new estimate of β by solving (2) with $H(t_k) = \widehat{H}^{(0)}(t_1, \hat{\beta}^{(0)})$.

Step 3: Set $\hat{\beta}^{(0)}$ to be the estimate obtained in Step 2 and repeat Step 1 and 2 until prescribed coverage criteria are met.

Let $\hat{\beta}$ and $\widehat{H}(t, \hat{\beta})$ be the estimators of β and $H(t, \beta)$. Also let

$$\hat{\xi}_{1k}\left(x_i, \hat{\beta}, \widehat{H}\right) = \hat{\xi}_{1k}(x_i) = \sup\left\{t: \hat{S}_1(t|z_{1i} = k, x_i) \geq 1 - p\right\},\ k = 0, 1,$$

where $x_i = \left(z_{2i}, z_{3i}, \ldots, z_{qi}\right)$, and

$$\hat{S}_1(t|Z_i) = \exp(-\widehat{\Lambda}_1(t|Z_i))$$

with $\widehat{\Lambda}_1(t|Z_i) = \Lambda_\varepsilon\left\{\hat{\beta}^T Z_i + \widehat{H}(t)\right\}$. Therefore, a natural estimator of $\Delta(x_i)$ is then obtained as $\widehat{\Delta}_1(x_i) = \hat{\xi}_{10}(x_i) - \hat{\xi}_{11}(x_i)$.

Since it is difficult to derive the asymptotic variance of the estimated median $\hat{\xi}_{1k}(x_i),\ k = 0, 1$, we consider using the nonparametric bootstrap method. The validity of bootstrap method for estimating variance of the median was justified by Reid (1981). The bootstrap procedure is as follows

Let $U_i = \left\{\left(\tilde{T}_i, \delta_i, Z_i\right),\ i = 1, \ldots, n\right\}$

1. Choose n sample points, U_i^* by sampling with replacement from U_i.

2. Calculate the $\hat{\beta}^*$ and $\widehat{H}^*(t)$ using equations (2) and (3), and derive the estimated medians $\hat{\xi}_{1k}^*(x_i),\ k = 0, 1$, and then obtain $\widehat{\Delta}_1^*(x_i) = \hat{\xi}_{10}^*(x_i) - \hat{\xi}_{11}^*(x_i)$.

3. Repeat step 1 and step 2 B times, we then have $\widehat{\Delta}_{1(1)}^*(x_i), \ldots, \widehat{\Delta}_{1(B)}^*(x_i)$. Let $d_1(\alpha)$ be the $100(1-\alpha)\%$ of the $\left|\widehat{\Delta}_1(x_i) - \Delta(x_i)\right|$, and the corresponding estimate is $\hat{d}_1(\alpha)$, which is the $100(1-\alpha)\%$ of $\{\left|\widehat{\Delta}_{1(1)}^*(x_i) - \hat{\Delta}(x_i)\right|, \ldots, \left|\widehat{\Delta}_{1(B)}^*(x_i) - \hat{\Delta}(x_i)\right|\}$. Therefore, the $100(1-\alpha)\%$ confidence interval for $\Delta(x_i)$ is given by

$$CI_1(x_i): \widehat{\Delta}_1(x_i) \pm \hat{d}_1(\alpha). \tag{4}$$

2.3 The methods based on the MLE

Next, we consider constructing confidence intervals using the MLE (Zeng and Lin, 2006, Chen, 2009). Based on the model (1), the unspecified parts of the model include the vector of regression coefficients β, and the increasing function Λ, both are to be estimated based on the observed data. Denote by $d\Lambda(t_*)$ the jump size of Λ at some observed event time t_*, and $\{d\Lambda\}$ the set of the jump sizes of Λ at the observed

event time. Let $g_i(t;\,\beta,\Lambda) = g\{\xi_i(t;\,\beta,\Lambda)\}$ and $\psi_i(t;\,\beta,\Lambda) = \psi\{\xi_i(t;\,\beta,\Lambda)\}$, with $g(t) = G'(t) \equiv dG(t)/dt,\ \psi(t) = g'(t)/g(t) = G"(t)/G'(t)$, and

$$\xi_i(t;\,\beta,\Lambda) = \Lambda(t)\exp(\beta^T Z_i).$$

The log-likelihood (Zeng and Lin, 2006, Chen, 2009) concerning the parameters β and $\Lambda(\cdot)$ is given by

$$l(\beta,\{d\Lambda\}) = \sum_{i=1}^{n}\left[\int_0^\tau \{\beta^T Z_i + \log g_i(t-;\,\beta,\Lambda) + \log d\Lambda(t)\}dN_i(t) - G_i(\beta,\Lambda)\right], \quad (5)$$

where

$$G_i(\beta,\Lambda) = G\{\xi_i(\tau;\,\beta,\Lambda)\} = \Lambda(t)\exp(\beta^T Z_i)g_i(t-;\,\beta,\Lambda).$$

To estimate $\{d\Lambda\}$ for fixed β, we differentiate the log-likelihood function in Equation (5) with respect to $d\Lambda_* \equiv d\Lambda(t_*)$, and then arrive at the following score function for $d\Lambda_*$:

$$U_{d\Lambda_*} \equiv \sum_{i=1}^{n}\left\{\frac{dN_i(t_*)}{d\Lambda(t_*)} - w_i(t_*;\,\beta,\Lambda)Y_i(t_*)e^{\beta^T Z_i}g_i(t_*-;\,\beta,\Lambda)\right\}, \quad (6)$$

where

$$w_i(t_*;\,\beta,\Lambda) = 1 - \frac{\kappa_i(t_*;\,\beta,\Lambda)}{g_i(t_*-;\,\beta,\Lambda)} \quad (7)$$

$$\kappa_i(t_*;\,\beta,\Lambda) = \int_{t_*+}^{\tau} \psi_i(u-;\,\beta,\Lambda)\,dM_i(u),$$

with $dM_i(t) = dN_i(t) - Y_i(t)e^{\beta^T Z_i}g_i(t-;\,\beta,\Lambda)d\Lambda(t)$. Hence, we obtain the weighted Breslow-type estimator for the jump size of Λ at t_* from Equation (6), which satisfies

$$d\Lambda(t_*) = \frac{\sum_{i=1}^{n} dN_i(t_*)}{\sum_{i=1}^{n} w_i(t_*;\,\beta,\Lambda)Y_i(t_*)e^{\beta^T Z_i}g_i(t-;\,\beta,\Lambda)}. \quad (8)$$

Note that, the weight $w_i(t_*;\,\beta,\Lambda)$ in Equation (7) depends on the weighted martingale residual $\kappa_i(t_*;\,\beta,\Lambda)$, and is determined by the function $\psi = g'/g$. For the proportional hazards model, $\psi \equiv 0$ and hence $w \equiv 1$, so that the Equation (8) reduces to the standard Breslow estimator (Breslow, 1974).

The score functions for β obtained by differentiating the log-likelihood function

in Equation (5) with respect to β is of the form

$$U_\beta = \sum_{i=1}^{n} \int_0^\tau Z_i\{dN_i(t) - w_i(t;\ \beta, \Lambda)Y_i(t)e^{\beta^T Z_i} g_i(t-;\ \beta, \Lambda)d\Lambda(t)\}.$$

Replacing $d\Lambda$ by the weighted Breslow-type estimator $d\widehat{\Lambda}$ and Λ by $\widehat{\Lambda} = \int d\widehat{\Lambda}(t)$ in U_β, we obtain the profile likelihood score equation for β:

$$0 = \sum_{i=1}^{n} \int_0^\tau \left\{ Z_i - \frac{\sum_{j=1}^{n} w_j(t;\beta,\widehat{\Lambda})Y_j(t)Z_j e^{\beta^T Z_j} g_j(t-;\beta,\widehat{\Lambda})}{\sum_{j=1}^{n} w_j(t;\beta,\widehat{\Lambda})Y_j(t)e^{\beta^T Z_j} g_j(t-;\beta,\widehat{\Lambda})} \right\} dN_i(t). \qquad (9)$$

Equation (9) suggests the following iterative algorithms for nonparametric maximum likelihood estimation in semiparametric transformation models. Starting with initial weights $w_i{}^{(0)} \equiv 1$ and the initial values for β and $\{d\Lambda\}$, e.g. $\beta^{(0)} = 0$ and $d\Lambda^{(0)} = 1/n$, for $k = 0, 1, \dots,$

Step 1: the estimating equations (9) and the weighted Breslow estimator (8) are solved to obtain $\left(\hat{\beta}^{(k+1)}, d\widehat{\Lambda}^{(k+1)}\right)$ with the weights fixed at $w_i{}^{(k)}$;

Step 2: the updated weighted $w_i{}^{(k+1)}$ are obtained from (7) with the newly solved $\left(\hat{\beta}^{(k+1)}, \widehat{\Lambda}^{(k+1)}\right)$.

Let $\hat{\beta}$ and $\widehat{\Lambda}$ be the estimators of β and Λ. Also let

$$\hat{\xi}_{2k}\left(x_i, \hat{\beta}, \widehat{\Lambda}\right) = \hat{\xi}_{2k}(x_i) = \sup\{t: \hat{S}_2(t|z_{1i} = k, x_i) \geq 1 - p\},\ k = 0, 1,$$

where $x_i = (z_{2i}, z_{3i}, \dots, z_{qi})$, and

$$\hat{S}_2(t|Z_i) = \exp(-\widehat{\Lambda}_2(t|Z_i)).$$

with $\widehat{\Lambda}_2(t|Z_i) = G\{\widehat{\Lambda}(t)\exp(\hat{\beta}^T Z_i)\}$. Therefore, a natural estimator of $\Delta(x_i)$ is then obtained as $\widehat{\Delta}_2(x_i) = \hat{\xi}_{20}(x_i) - \hat{\xi}_{21}(x_i)$.

Since it is difficult to derive the asymptotic variance of the estimated median $\hat{\xi}_{2k}(x_i),\ k = 0, 1$, we consider again the nonparametric bootstrap method. The description of procedure can be found in section 2.2. Let $d_2(\alpha)$ be the $100(1-\alpha)\%$ of the $\left|\widehat{\Delta}_2(x_i) - \Delta(x_i)\right|$, and the corresponding estimate is $\hat{d}_2(\alpha)$, which is the

100 $(1-\alpha)\%$ of $\{|\widehat{\Delta}^*_{2(1)}(x_i) - \hat{\Delta}(x_i)|, \ldots, |\widehat{\Delta}^*_{2(B)}(x_i) - \hat{\Delta}(x_i)|\}$. Therefore, the $100(1-\alpha)\%$ confidence interval for $\Delta(x_i)$ is given by

$$CI_2(x_i): \widehat{\Delta}_2(x_i) \pm \hat{d}_2(\alpha). \tag{10}$$

2.4 Model Selection

The class of semiparametric transformation models as shown in (1) requires specification of the transformation function $G(x)$. Misspecifying any of these components can result in erroneous inference and inaccurate prediction. For right-censored data, Chen et al. (2012) introduced time-dependent martingale residuals for semiparametric transformation models and used the cumulative sums of the residuals for model assessment. The estimated martingale $M_i(t;\, \hat{\beta}, \widehat{\Lambda})$ is as follow:

$$M_i(t;\, \hat{\beta}, \widehat{\Lambda}) = N_i(t) - G\left\{\int_0^t Y_i(\mathrm{s})\exp(\hat{\beta}^T Z_i) d\widehat{\Lambda}(s)\right\},$$

and the cumulative sums of residuals over the argument of the transformation function is:

$$R_{tr}(\omega, t) = n^{-1/2} \sum_{i=1}^{n} \int_0^t I\left(\int_0^u Y_i(\mathrm{s})\exp(\hat{\beta}^T Z_i) d\widehat{\Lambda}(s) \le \omega\right) dM_i(u;\, \hat{\beta}, \widehat{\Lambda}), \tag{11}$$

where ω is a constant.

Chen et al. (2012) showed that the residual process converges weakly to a zero-mean Gaussian process and suggested conducting goodness-of-fit test based on the *p*-value of a supremum test, which can be obtained by using Monte Carlo procedure. Here, in real data analysis, we select the best-fit model according to the criteria $\sup_{\omega,t}|R_{tr}(\omega, t)|$ and $\sup_{\omega}|R_{tr}(\omega, \infty)|$.

3. Simulation studies

In this section, a simulation study was conducted to investigate the coverage probability and expected length of the proposed $100(1-\alpha)\%$ intervals $CI_1(x)$ (4) and $CI_2(x)$ (10), for the difference between two median survival times. We simulated the survival time using the logarithmic transformation with cumulative intensity

$$\Lambda(t|z_1, z_2) = \log\left[1 + r\left\{\Lambda(t)e^{\beta_1 z_1 + \beta_2 z_2}\right\}\right]/r,$$

where $\Lambda(t) = t$ and $r = 0$, 0.5, 1 or 2. Note that the proportional hazards and proportional odds models correspond to $r = 0$ and $r = 1$. Two covariates z_1 and z_2 are generated from the Bernoulli distribution with success probability 0.5, and the uniform distribution over (0, 1), respectively. The censoring time is independent of the covariates and follows the uniform distribution over (c_1, c_2). The values of c_1 and c_2 are chosen such that the censoring proportion is equal to 20%. We consider the sample size $n = 100$ and 200, the number of bootstrap repetitions $B = 200$, and all simulations are based on 1000 replications.

The values of parameters are set as $(\beta_1, \beta_2) =$ (0.5, 0.5), (0, 0.5), (0.5, -0.5) and (-0.5, 0.5). For each simulated dataset, we first obtain the estimators of β and Λ. Given $z_2 = 0.5$ and $z_2 = 0.8$, we then calculate the 95% and 90% confidence intervals for the difference of two median survival times given the covariates. The estimated coverage probabilities are obtained by calculating the proportion of the 1000 confidence intervals which cover the true difference of median survival times. Tables 1 and 2 show the results for $n = 100$ and 200, respectively. We also obtain the expected length of the confidence intervals by calculating the average length of the 1000 confidence intervals. The results are reports in Tables 3 and 4.

The results in Table 1 show that when the sample size are small, all the

confidence intervals $CI_1(0.5)$, $CI_1(0.8)$, $CI_2(0.5)$ and $CI_2(0.8)$, are slightly higher than their nominal levels. When the sample size increases, the results in Table 2 show that for the case of $r = 0$, most of the confidence intervals hold their nominal level. Note that, the coverage probabilities of both CI_1 and CI_2 increase as r increases.

Table 3 and 4 indicate that the expected lengths of the interval $CI_1(0.5)$ and $CI_1(0.8)$, are almost the same as that of $CI_2(0.5)$ and $CI_2(0.8)$, for $r = 0$. When r is large, the expected lengths of the interval $CI_2(0.5)$ and $CI_2(0.8)$, are slightly shorter than that of $CI_1(0.5)$ and $CI_1(0.8)$. Moreover, the expected lengths of CI_1 and CI_2 decrease when the sample size increases from 100 to 200.

To sum up, when the sample sizes are small, most of the confidence intervals are unable to reach the specified nominal level. However, when sample size increases to $n = 200$, the confidence intervals CI_2 reach the specified level for the case of $r = 0$.

In terms of coverages, the confidence intervals CI_2 perform better than confidence intervals CI_1 and have a shorter expected lengths compared to CI_1.

Table 1. The estimated coverage probability of $100(1-\alpha)\%$ confidence interval for the difference between two median survival times for $n = 100$.

Model	β_1	β_2	$1-\alpha$	$n = 100$			
				$CI_1(x)$		$CI_2(x)$	
				$x = 0.5$	$x = 0.8$	$x = 0.5$	$x = 0.8$
$r = 0$	0.5	0.5	0.95	0.966	0.969	0.966	0.971
			0.90	0.910	0.928	0.920	0.934
	0	0.5	0.95	0.971	0.984	0.979	0.981
			0.90	0.940	0.947	0.939	0.937
	0.5	-0.5	0.95	0.968	0.973	0.958	0.969
			0.90	0.924	0.930	0.919	0.931
	-0.5	0.5	0.95	0.967	0.963	0.970	0.964
			0.90	0.914	0.918	0.915	0.923
$r = 0.5$	0.5	0.5	0.95	0.973	0.978	0.977	0.973
			0.90	0.937	0.932	0.943	0.936
	0	0.5	0.95	0.977	0.979	0.974	0.980
			0.90	0.936	0.937	0.940	0.928
	0.5	-0.5	0.95	0.965	0.978	0.971	0.969
			0.90	0.916	0.930	0.920	0.927
	-0.5	0.5	0.95	0.971	0.970	0.971	0.971
			0.90	0.919	0.939	0.926	0.930
$r = 1$	0.5	0.5	0.95	0.975	0.983	0.976	0.980
			0.90	0.931	0.951	0.941	0.947
	0	0.5	0.95	0.987	0.985	0.983	0.988
			0.90	0.938	0.952	0.940	0.954
	0.5	-0.5	0.95	0.977	0.982	0.977	0.982
			0.90	0.934	0.938	0.929	0.932
	-0.5	0.5	0.95	0.979	0.983	0.979	0.982
			0.90	0.940	0.940	0.945	0.948
$r = 2$	0.5	0.5	0.95	0.984	0.989	0.979	0.985
			0.90	0.949	0.948	0.947	0.949
	0	0.5	0.95	0.986	0.993	0.987	0.991
			0.90	0.957	0.959	0.954	0.961
	0.5	-0.5	0.95	0.983	0.983	0.982	0.984
			0.90	0.945	0.950	0.952	0.957
	-0.5	0.5	0.95	0.983	0.982	0.983	0.987
			0.90	0.944	0.955	0.946	0.953

Table 2. The estimated coverage probability of $100(1-\alpha)\%$ confidence interval for the difference between two median survival times for $n = 200$.

Model	β_1	β_2	$1-\alpha$	$n = 200$			
				$CI_1(x)$		$CI_2(x)$	
				$x = 0.5$	$x = 0.8$	$x = 0.5$	$x = 0.8$
$r = 0$	0.5	0.5	0.95	0.960	0.952	0.959	0.956
			0.90	0.916	0.917	0.905	0.904
	0	0.5	0.95	0.958	0.967	0.955	0.965
			0.90	0.909	0.918	0.909	0.924
	0.5	-0.5	0.95	0.949	0.948	0.950	0.946
			0.90	0.894	0.898	0.896	0.902
	-0.5	0.5	0.95	0.953	0.958	0.956	0.955
			0.90	0.891	0.903	0.895	0.899
$r = 0.5$	0.5	0.5	0.95	0.962	0.964	0.961	0.961
			0.90	0.913	0.925	0.910	0.914
	0	0.5	0.95	0.956	0.970	0.953	0.969
			0.90	0.905	0.918	0.919	0.924
	0.5	-0.5	0.95	0.958	0.963	0.957	0.963
			0.90	0.913	0.909	0.908	0.916
	-0.5	0.5	0.95	0.954	0.963	0.960	0.965
			0.90	0.905	0.911	0.908	0.913
$r = 1$	0.5	0.5	0.95	0.967	0.972	0.960	0.972
			0.90	0.914	0.922	0.919	0.921
	0	0.5	0.95	0.969	0.981	0.968	0.977
			0.90	0.928	0.941	0.931	0.938
	0.5	-0.5	0.95	0.959	0.960	0.962	0.961
			0.90	0.911	0.922	0.915	0.912
	-0.5	0.5	0.95	0.958	0.968	0.955	0.957
			0.90	0.902	0.922	0.920	0.920
$r = 2$	0.5	0.5	0.95	0.975	0.984	0.971	0.961
			0.90	0.932	0.935	0.925	0.929
	0	0.5	0.95	0.979	0.984	0.984	0.986
			0.90	0.932	0.946	0.945	0.957
	0.5	-0.5	0.95	0.964	0.966	0.973	0.975
			0.90	0.919	0.922	0.926	0.932
	-0.5	0.5	0.95	0.971	0.974	0.971	0.974
			0.90	0.937	0.929	0.937	0.944

Table 3. The estimated average length of $100(1-\alpha)\%$ confidence interval for the difference between two median survival times for $n = 100$.

Model	β_1	β_2	$1-\alpha$	$n = 100$			
				$CI_1(x)$		$CI_2(x)$	
				$x = 0.5$	$x = 0.8$	$x = 0.5$	$x = 0.8$
$r = 0$	0.5	0.5	0.95	0.526	0.479	0.527	0.483
			0.90	0.427	0.386	0.430	0.388
	0	0.5	0.95	0.633	0.562	0.633	0.563
			0.90	0.511	0.452	0.514	0.452
	0.5	-0.5	0.95	0.904	1.133	0.901	1.136
			0.90	0.732	0.906	0.732	0.908
	-0.5	0.5	0.95	0.901	0.818	0.902	0.819
			0.90	0.733	0.665	0.736	0.667
$r = 0.5$	0.5	0.5	0.95	0.841	0.776	0.827	0.777
			0.90	0.673	0.616	0.671	0.613
	0	0.5	0.95	0.999	0.909	0.982	0.901
			0.90	0.805	0.724	0.799	0.718
	0.5	-0.5	0.95	1.383	1.768	1.366	1.754
			0.90	1.107	1.395	1.092	1.388
	-0.5	0.5	0.95	1.364	1.265	1.343	1.242
			0.90	1.101	1.005	1.082	0.989
$r = 1$	0.5	0.5	0.95	1.290	1.228	1.231	1.171
			0.90	1.024	0.956	0.981	0.900
	0	0.5	0.95	1.576	1.473	1.507	1.384
			0.90	1.258	1.143	1.203	1.082
	0.5	-0.5	0.95	2.193	2.889	2.126	2.801
			0.90	1.728	2.232	1.674	2.144
	-0.5	0.5	0.95	2.195	2.107	2.140	2.036
			0.90	1.734	1.636	1.691	1.574
$r = 2$	0.5	0.5	0.95	3.178	3.252	2.893	2.852
			0.90	2.430	2.402	2.202	2.088
	0	0.5	0.95	3.819	3.876	3.434	3.344
			0.90	2.952	2.890	2.639	2.514
	0.5	-0.5	0.95	5.416	7.342	5.063	6.687
			0.90	4.139	5.483	3.850	5.046
	-0.5	0.5	0.95	5.298	5.413	4.914	4.874
			0.90	4.041	4.059	3.701	3.604

Table 4. The estimated average length of $100(1-\alpha)\%$ confidence interval for the difference between two median survival times for $n = 200$.

Model	β_1	β_2	$1-\alpha$	$n = 200$			
				$CI_1(x)$		$CI_2(x)$	
				$x = 0.5$	$x = 0.8$	$x = 0.5$	$x = 0.8$
$r = 0$	0.5	0.5	0.95	0.352	0.317	0.352	0.314
			0.90	0.292	0.261	0.292	0.260
	0	0.5	0.95	0.410	0.362	0.411	0.364
			0.90	0.337	0.295	0.338	0.297
	0.5	-0.5	0.95	0.601	0.732	0.598	0.733
			0.90	0.496	0.605	0.494	0.602
	-0.5	0.5	0.95	0.589	0.528	0.589	0.529
			0.90	0.488	0.436	0.488	0.435
$r = 0.5$	0.5	0.5	0.95	0.541	0.489	0.535	0.490
			0.90	0.445	0.399	0.440	0.400
	0	0.5	0.95	0.641	0.567	0.622	0.549
			0.90	0.526	0.464	0.510	0.449
	0.5	-0.5	0.95	0.891	1.089	0.878	1.089
			0.90	0.732	0.891	0.723	0.885
	-0.5	0.5	0.95	0.875	0.796	0.865	0.782
			0.90	0.725	0.653	0.716	0.641
$r = 1$	0.5	0.5	0.95	0.810	0.755	0.785	0.725
			0.90	0.663	0.609	0.641	0.583
	0	0.5	0.95	0.976	0.890	0.916	0.833
			0.90	0.798	0.718	0.749	0.671
	0.5	-0.5	0.95	1.353	1.728	1.293	1.641
			0.90	1.104	1.377	1.052	1.314
	-0.5	0.5	0.95	1.349	1.243	1.278	1.162
			0.90	1.101	1.004	1.047	0.942
$r = 2$	0.5	0.5	0.95	1.819	1.727	1.571	1.466
			0.90	1.464	1.357	1.259	1.159
	0	0.5	0.95	2.185	2.078	1.870	1.738
			0.90	1.757	1.639	1.506	1.371
	0.5	-0.5	0.95	3.017	4.036	2.732	3.618
			0.90	2.423	3.137	2.188	2.817
	-0.5	0.5	0.95	3.018	2.916	2.727	2.578
			0.90	2.409	2.282	2.172	2.025

4. Data analysis

In a two-arm lung cancer study, as described in Ying et al. (1995), the standard therapy is to use a combination of etoposide (E) and cisplatin (P); however, the optimal sequencing and administration schedule have not been established. Thus, 121 patients with small cell lung cancer were randomly assigned to two treatment groups, Arm A: P followed by E and Arm B: E followed by P. In addition to the survival time, patient's entry age was recorded. Note that there are only 8 patients aged below 50. To avoid a possible misleading interpretation of the effect of age, we illustrate the proposed methods by analyzing the data set involving 113 patients aged 50 or more. Among the 58 patients in Arm A, 15 had their survival times censored, while among the 55 patients in Arm B, there were 7 censored survival times. In this section, we explore the difference in median survival time between the two arms as a function of the entry age when survival times are subject to random right-censorship.

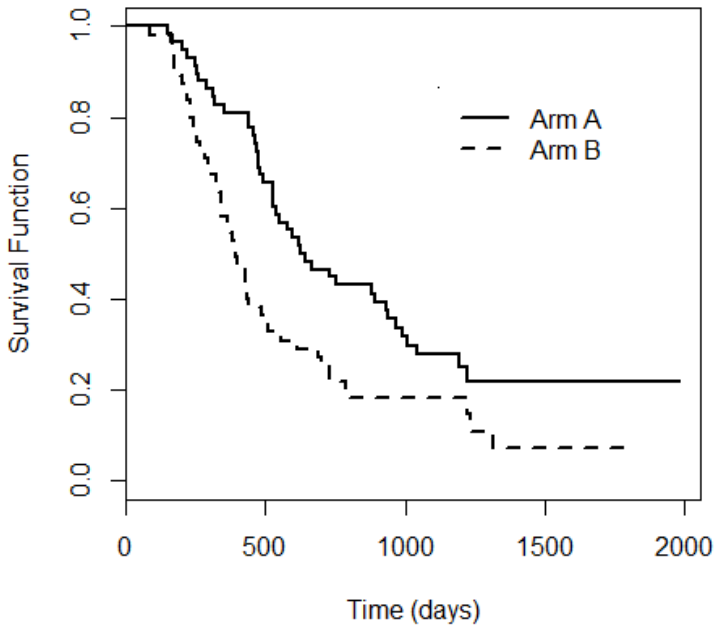

Figure 1. The Kaplan–Meier estimates for patients with small cell lung cancer.

The associated Kaplan–Meier (1958) estimates of the survival functions for patients in different arms are given in Figure 1. Patients in Arm A have higher survival rate than that in Arm B in general.

As noted in Ying et al. (1995) and Chen et al. (2015) that the Cox proportional hazards model is not appropriate for the data set when both treatment indicator and entry age are involved. Also, for each treatment group, the proportional hazards assumption is violated, as shown in Figure 2, the plot of $\log(-\log\hat{S}(t))$ vs. t. Therefore, we consider fitting the semiparametric transformation model into the two groups with entry age as the covariate.

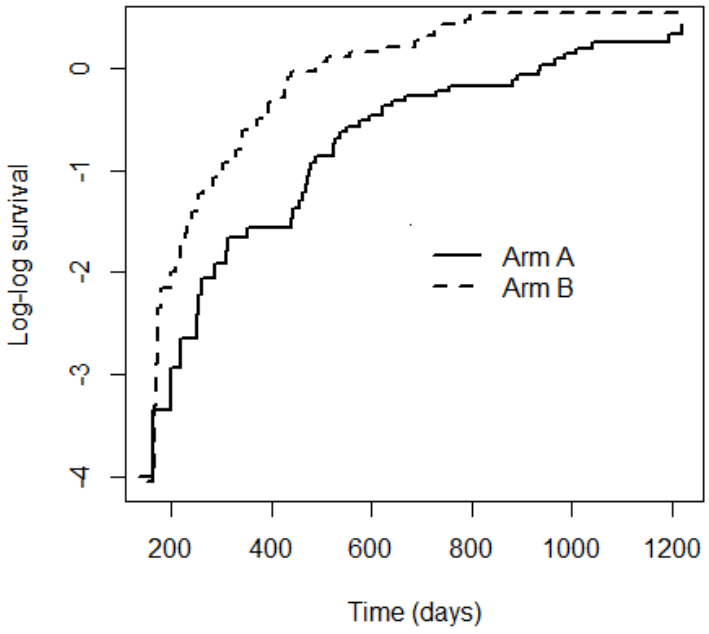

Figure 2. The plot of $\log(-\log\hat{S}(t))$ against time for the patients.

From Figures 3 and 4, under the semiparametric transformation model with $r = 0$, the proposed 95% (90%) confidence interval identifies the treatment difference among persons aged 50, 51, 53 and 57 (50 to 61, 63 to 66, 72 and 77 to 79) years, which details the information about the age-dependent difference between treatment and control. The median of survival times in Arm A are larger than that in Arm B. Under the semiparametric transformation model with $r = 0.5$, the proposed 95%

(90%) confidence interval identifies the difference among persons aged from 50, 52 to 55, 57 to 61 and 66 (50 to 68, 70, 75 and 79) years.

Under the semiparametric transformation model with $r = 1$, the proposed 95% (90%) confidence interval identifies the difference among persons aged from 50 to 61, 63, 67, 68, 70 and 79 (50 to 71, 73, 74, 77 to 79) years, and the difference among persons aged from 50, 52 to 62, 69, 73 and 79 (50 to 79) years under the model with $r = 2$.

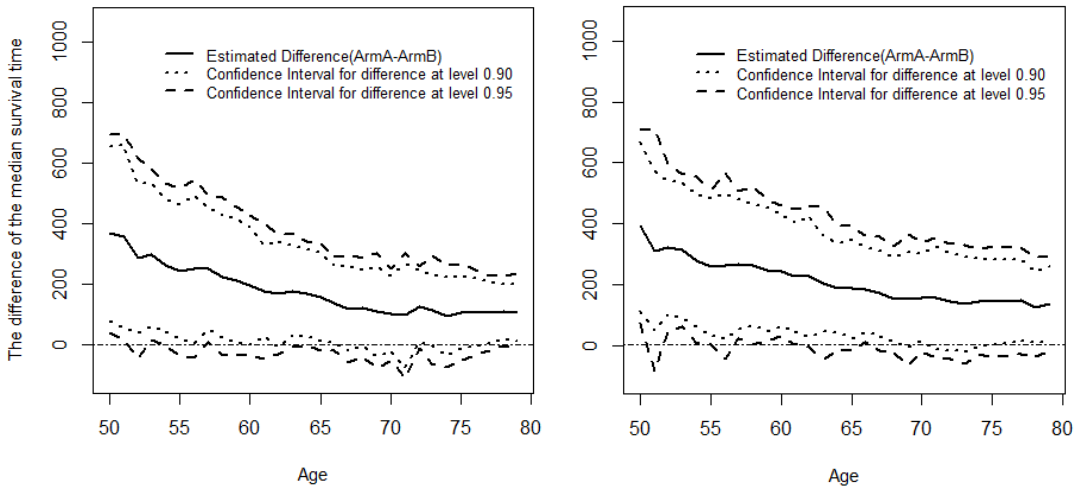

Figure 3. The confidence intervals for the difference in median survival time between Arms A and B under the semiparametric transformation model with $r = 0$ (left) and $r = 0.5$ (right) for patients with small cell lung cancer.

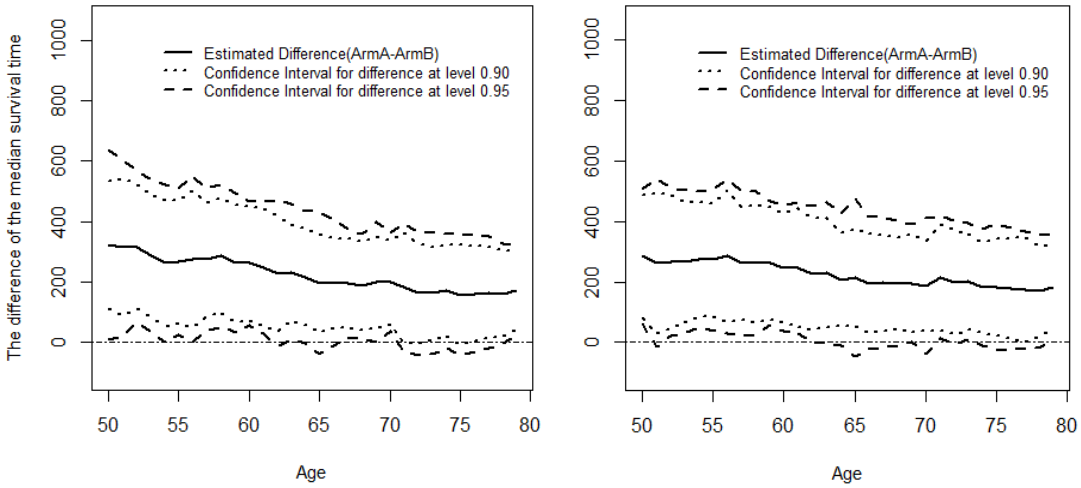

Figure 4. The confidence intervals for the difference in median survival time between Arms A and B under the semiparametric transformation model with $r = 1$ (left) and $r = 2$ (right) for patients with small cell lung cancer.

To select appropriate models, we consider the logarithmic transformations model with $r = 0$, 0.5, 1, 2. Using Equation (11), we obtain the best fit model based on $\sup_{\omega,t}|R_{tr}(\omega,t)|$ and $\sup_{\omega}|R_{tr}(\omega,\infty)|$.

Table 5. The cumulative sum of residuals $R_{tr}(\omega,t)$ and $R_{tr}(\omega,\infty)$ under the logarithmic transformation model with $r = 0$, 0.5, 1, 2.

	$\sup_{\omega,t}\|R_{tr}(\omega,t)\|$	$\sup_{\omega}\|R_{tr}(\omega,\infty)\|$
$r = 0$	0.682	0.417
$r = 0.5$	0.628	0.419
$r = 1$	0.581	0.410
$r = 2$	0.581	0.508

Table 5 indicates that a semiparametric transformation model with $r = 1$, that is the proportional odds model, fits the best among the four models considered. Fitting the data to the proportional odds model, we obtain the estimates $\hat{\beta}_1 = 1.016$ for treatment group and $\hat{\beta}_2 = 0.026$ for the effect of entry age, respectively, and the estimate of the difference in median survival time between the two arms given patients aged 60 is 262 days, which implies that patients in Arm A have higher survival rate than that in Arm B.

5. Discussions and conclusions

In this article, based on the EE estimator (Chen et al., 2002) and the MLE (Zeng and Lin, 2006, Chen, 2009), we have constructed the conditional confidence intervals for the difference of two median survival times given the covariates under semiparametric transformation model. Simulation results indicate that, in terms of coverage probabilities and interval lengths, the method based on the MLE performs better than that based on the EE estimator. When the sample sizes are small, most of the confidence intervals overestimate the nominal levels. As sample size increases, the coverage improves and some of the confidence intervals based on the MLE reach nominal level. The expected lengths of the intervals based on the MLE are slightly shorter than that based on the EE estimator.

In data analysis, according to Equations (11), we select the best-fit model: the semiparametric transformation model with $r = 1$. The 95% (90%) confidence intervals suggest that given at age 50 to 61, 63, 67, 68, 70 and 79 (50 to 71, 73, 74, 77 to 79) years old, the median of survival times in Arm A are larger than that in Arm B.

The proposed method can be generalized to construct intervals for the difference of two percentiles of survival times. In some situations, the survival time can be subject to interval censoring/truncation and covariates can be subject to mismeasurement errors. Further research is required to extent the propose method to these complex situations.

References

Bennett, S. (1983). Analysis of survival data by the proportional odds model. *Statistics in Medicine*, **2**, 273–277.

Burr, D. and Doss, H. (1993). Confidence bands for the median survival time as a function of the covariates in the Cox model. *Journal of the American Statistical Association*, **88**, 1330–1340.

Breslow, N. E. (1974). Covariance analysis of censored survival data. *Biometrics*, **30**, 89–99.

Cai, T., Wei, L. J. and Wicox, M. (2000). Semiparametric regression analysis for clustered failure time data. *Biometrika*, **87**, 867–878.

Cai, T., Cheng, S. C. and Wei, L. J. (2002). Semiparametric mixed-effects models for clustered failure time data. *Journal of the American Statistical Association*, **97**, 514–522.

Chen, K., Jin, Z. and Ying, Z. (2002). Semiparametric analysis of transformation models with censored data. *Biometrika*, **89**, 659–668.

Chen, L., Lin, D. Y. and Zeng, D. (2012). Checking semiparametric transformation models with censored data. *Biostatistics*, **13**, 18–31.

Chen, Y. H. (2009). Weighted Breslow-type and maximum likelihood estimation in semiparametric transformation models. *Biometrika*, **96**, 591–600.

Chen, Y. I. and Chang, Y. M. (2007). Covariates-dependent confidence intervals for the difference or ratio of two median survival times. *Statistics in Medicine*, **26**, 2203–2213.

Chen, Y. I., Chang, Y. M. and Lee, J. Y. (2015). Confidence bands for the difference between two median survival times as a function of covariates. *Lifetime Data Analysis*, **21**, 97–118.

Cheng, S. C., Wei, L. J. and Ying, Z. (1995). Analysis of transformation models with censored data. *Biometrika*, **82**, 835–845.

Cheng, S. C., Wei, L. J. and Ying, Z. (1997). Predicting survival probabilities with semiparametric transformation models. *Journal of the American Statistical Association*, **92**, 227–235.

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

Dabrowska, D. M. and Doksum, K. A. (1987). Estimates and confidence intervals for median and mean life in the proportional hazard model. *Biometrika*, **74**, 799–807.

Fine, J. P., Ying, Z. and Wei, L. J. (1998). On the linear transformation model with censored data. *Biometrika*, **85**, 980–6.

Fleming, T. R. and Harrington, D. P. (1991). Counting Processes and Survival Analysis. *John Wiley and Sons.*

Kaplan, E. L., Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.

Kim, J. (2001). Confidence intervals for the difference of median survival times using the stratified Cox proportional hazards model. *Biometrical Journal*, **43**, 781–790.

Kosorok, M. R., Lee, B. L. and Fine, J. P. (2004). Robust inference for univariate proportional hazards frailty regression models. *Annals of Statistics*, **32**, 1448–1491.

Lu, W. and Ying, Z. (2004). On semiparametric transformation cure models. *Biometrika*, **91**, 331–343.

Lu, W. and Tsiatis, A. A. (2006). Semiparametric transformation models for the case-cohort study. *Biometrika*, **93**, 207–214.

Murphy, S. A., Rossini, A. J. and van der Vaart, A. W. (1997). Maximal likelihood estimation in the proportional odds model. *Journal of the American Statistical Association*, **92**, 968–976.

Reid, N. (1981). Estimating the median survival time. *Biometrika*, **68**, 601–608.

Scharfstein, D. O., Tsiatis, A. A. and Gilbert, P. B. (1998). Semiparametric efficient estimation in the generalized odds-rate class of regression models for right-censored time-to-event data. *Lifetime Data Analysis*, **4**, 355–391.

Su, J. Q. and Wei, L. J. (1993). Nonparametric estimation for the difference or ratio of median failure times. *Biometrics*, **49**, 603–607.

Tsodikov, A. (2003). Semiparametric models: a generalized self-consistency approach. *Journal of the Royal Statistical Society, Series B*, **65**, 759–774.

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.

Zeng, D. and Lin, D. Y. (2006). Efficient estimation of semiparametric transformation models for counting processes. *Biometrika*, **93** 627–640.

Zeng, D. and Lin, D. Y. (2007). Maximum likelihood estimation in semiparametric regression models with censored data (with discussion). *Journal of the Royal Statistical Society, Series B*, **69**, 507–564.

Zeng, D., Lin, D. Y. and Lin, X. (2008). Semiparametric transformation models with random effects for clustered failure time data. *Statistica Sinica*, **18**, 355–377.

Zucker, D. M. (1998). Restricted mean life with covariates: modification and extension of a useful survival analysis method. *Journal of the American Statistical Association*, **93**, 702–709.