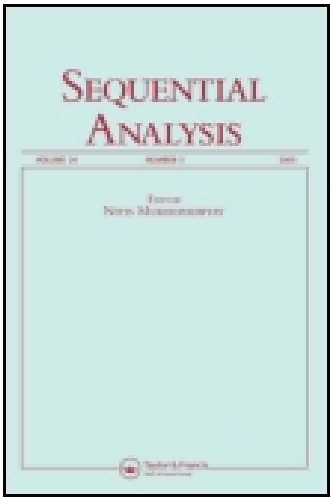

Sequential Analysis: Design Methods and Applications

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsqa20

Sequential Analysis of Censored Data with Linear Transformation Models

Lin Huang [a] & Zhezhen Jin [b]

[a] Children's Hospital Boston , Harvard Medical School , Boston , Massachusetts , USA

[b] Department of Biostatistics , Columbia University , New York , New York , USA

Published online: 10 Apr 2012.

Sequential Analysis, 31: 172–189, 2012

ISSN: 0747-4946 print/1532-4176 online
DOI: 10.1080/07474946.2012.665678

Sequential Analysis of Censored Data with Linear Transformation Models

Lin Huang[1] and Zhezhen Jin[2]
[1]Children's Hospital Boston, Harvard Medical School, Boston, Massachusetts, USA
[2]Department of Biostatistics, Columbia University, New York, New York, USA

Abstract: Sequential tests have been used commonly in clinical trials to compare treatments. For sequential analysis of right-censored survival data with covariate adjustment, several different methods have been studied based on either Cox proportional hazards model or accelerated failure time model. Here we propose a test process based on linear transformation models for staggered entry data. The proposed test process is motivated by Chen et al.'s (2002) estimating equations for linear transformation models. We show that the test process can be approximated by a mean 0 multidimensional Gaussian process. A consistent estimator of its covariance matrix function is provided. For given interim analysis time points, a repeated significant test is developed based on the boundaries procedure proposed by Slud and Wei (1982). Numerical studies show that the proposed test process performs well.

Keywords: Censoring; Gaussian process; Linear transformation models; Repeated significant test; Sequential analysis.

Subject Classifications: 62L10; 62N01; 62N03.

1. INTRODUCTION

In many clinical trials, it is often of interest to compare survival times between two groups after adjustment of other risk factors. Due to ethical and economical reasons, data are also often monitored and examined at different times of study period to see if early stopping is possible or necessary. In addition, these trials usually recruit patients sequentially. To deal with all of these considerations, several sequential methods have been proposed; for example, Jones and Whitehead (1979),

Received August 25, 2010, Revised June 9, 2011, December 16, 2011, January 26, 2012, Accepted February 3, 2012
Recommended by D. S. Coad
Address correspondence to Lin Huang, Clinical Research Center, Children's Hospital Boston, 300 Longwood Ave, Boston, MA 02115, USA; E-mail: lin.huang@childrens.harvard.edu

Slud and Wei (1982), Tsiatis (1982), Sellke and Siegmund (1983), Slud (1984), Tsiatis et al. (1985), Lin (1992), Gu and Ying (1993), Tsiatis et al. (1995), Gu and Ying (1995), and Jennison and Turnbull (1997). In particular, for covariate adjusted analysis, Tsiatis et al. (1985), Gu and Ying (1995), Scharfstein et al. (1997), and Bilias et al. (1997) developed sequential procedures based on the proportional hazards models, and Lin (1992) and Gu and Ying (1993) studied sequential procedures based on the accelerated failure time model.

In this article, we develop a sequential test procedure based on linear transformation models, which cover the proportional hazards model and the proportional odds model as special cases. The linear transformation models are specified as

$$H(T) = -\beta^T X + \epsilon, \tag{1.1}$$

where $H(\cdot)$ is an unknown monotone function satisfying $H(0) = -\infty$, β is an unknown vector of regression parameters, and the error term ϵ is assumed to follow a known distribution F. Different distributions of the error term ϵ would yield different models. The Cox roportional hazards model and proportional odds model are two special cases of linear transformation models.

The Cox proportional hazards model has the form of $\lambda(t) = \lambda_0(t)\exp(\beta^T X)$. Taking an integral from 0 to T on both sides of the Cox proportional hazards model and making a log transformation, we will get

$$\log(\Lambda_0(T)) = -\beta^T X + \log(\Lambda(T)),$$

where $\Lambda(\cdot)$ is the cumulative hazard function for the survival time T. Here $H(T) = \log(\Lambda_0(T))$, the logarithm of the cumulative baseline hazard function, and $\log(\Lambda(T))$ is the error term ϵ which follows the extreme value distribution.

Consider the proportional odds model $\log \frac{S(t)/[1-S(t)]}{S_0(t)/[1-S_0(t)]} = \beta^T X$, where $S(\cdot)$ is the survival function of survival time T and $S_0(\cdot)$ is the unknown baseline survival function. By some simple algebra, we will get

$$\log(S_0(T)/[1 - S_0(T)]) = -\beta^T X + \log(S(T)/[1 - S(T)]).$$

Here $H(T) = \log(S_0(T)/[1 - S_0(T)])$, the logarithm of the odds of baseline survival function, and the error term $\log(S(T)/[1 - S(T)])$ follows the logistic distribution.

Attempts to develop a unified estimation method for linear transformation models have been made by Cheng et al. (1995, 1997), Fine et al. (1998), and Cai et al. (2000), among others. These methods require the censoring variables to be independent and identically distributed and also be independent of covariates. Chen et al. (2002) were able to develop an estimation method which yields consistent estimators of regression parameters and transformation function, without the requirement of independence between censoring variable and covariates. Zeng and Lin (2006, 2007) also studied the properties of the non-parametric maximum likelihood estimator of a broad class of transformation models including the linear transformation models.

Motivated by the work of Chen et al. (2002), in this article we propose a test process based on linear transformation models for repeated significance tests of regression hypothesis for staggered entry data under general right censoring.

Notations and formulations are introduced in Section 2. In Section 3, the test process with its corresponding theoretical results and a repeated significance test based on the boundary procedure of Slud and Wei (1982) are presented. In Section 4, simulation studies and analysis of a real dataset on prostate cancer are provided. A further discussion is given in Section 5.

2. NOTATION AND FORMULATION

In many clinical trials, subjects enter the study sequentially, called *staggered entry*. Interim analysis is often scheduled at several fixed calendar times. Thus, two time scales appear in these trials. One is the observed survival time, which starts from subject's entry time point and ends at the last follow-up time point, which results in either subject's failure observed or censored. This timescale varies among subjects. The other is calendar time, measuring from the beginning of the trial to the scheduled analysis time. The relationship between the two timescales is illustrated by Figure 1.

Let τ_i denote the calendar time when the ith patient enters a clinical trial, T_i denote survival time since entry, and C_i denote censoring time since entry. Then the longest possible time this patient would stay in the trial is from τ_i to $\tau_i + T_i \wedge C_i$. When an interim analysis is scheduled at calendar time t, the censoring variable becomes $C_i \wedge (t - \tau_i)^+$, where $(t - \tau_i)^+ = \max(0, (t - \tau_i))$.

Define two variables $\widetilde{T}_i(t)$ and $\delta_i(t)$ as follows:

$$\widetilde{T}_i(t) = T_i \wedge C_i \wedge (t - \tau_i)^+, \quad \delta_i(t) = 1\{T_i \leq C_i \wedge (t - \tau_i)^+\}.$$

They represent respectively observed survival time or censoring time at time t and failure-censoring indicator up to calendar time t. Note that $\delta_i(t) = 1$ if and only if the patient enters the trial and dies before time t.

We use Z_i to denote treatment arms and X_i to denote other covariates which may also affect the distribution of survival time. It is assumed that $(T_i, C_i, Z_i, X_i, \tau_i)$ $(i = 1, 2, \ldots)$ are independent and identically distributed, and T_i and C_i are conditionally independent given Z_i and X_i. It is also assumed that Z_i is independent

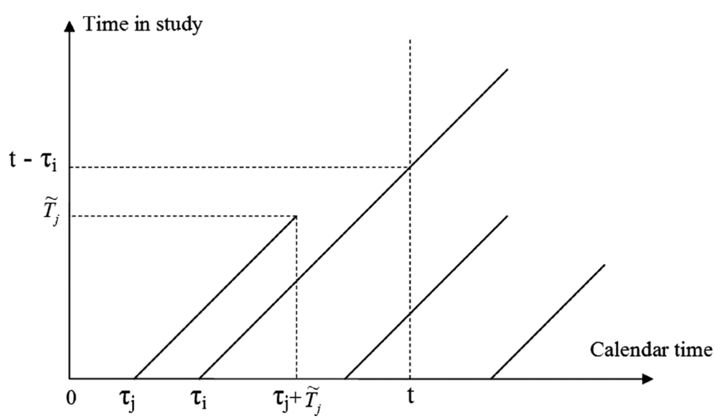

Figure 1. Lexis diagram with the two timescales in sequential analysis with staggered entry.

of X_i. With these notations, the data observed at time t consist of $(\widetilde{T}_i(t), \delta_i(t), Z_i, X_i)$ $(i\colon \tau_i < t)$.

The hypothesis to test is that the distribution of survival time is independent of treatment variable Z_i after adjustment of other covariates X_i. To reflect this, we rewrite the linear transformation model (1.1) as follows:

$$H(T_i) = -\gamma Z_i - \beta^T X_i + \epsilon_i, \quad (i = 1, \ldots, n) \tag{2.1}$$

where γ and β are unknown regression parameters. The null hypothesis is $H_0 : \gamma = 0$. For simplicity, we assume that Z is binary variable, with $Z = 1$ for treatment group and $Z = 0$ for the other group.

3. THEORETICAL RESULTS

For an analysis scheduled at calendar time t, we define the event indicator and the risk indictor at t with the usual counting process notations, as

$$\begin{aligned} Y_i(s; t) &= 1\{\widetilde{T}_i(t) \geq s\}, \\ N_i(s; t) &= \delta_i(t)1\{\widetilde{T}_i(t) \leq s\} \\ &= 1\{T_i \leq C_i \wedge (t - \tau_i)^+ \wedge s\}. \end{aligned}$$

Then for any fixed t, the process

$$M_i(s; t) = N_i(s; t) - \int_0^s Y_i(v; t) d\Lambda\{\gamma Z_i + \beta^T X_i + H(v)\} \tag{3.1}$$

is a martingale in s, where $\Lambda(\cdot)$ is the cumulative hazard functions of ϵ. Estimation equations from Chen et al. (2002) for the linear transformation models based on the data observered at time t are specified as:

$$\sum_{i:\tau_i<t} \int_0^\infty Z_i\big[dN_i(s; t) - Y_i(s; t)d\Lambda\{\gamma Z_i + \beta' X_i + H(s)\}\big] = 0, \tag{3.2a}$$

$$\sum_{i:\tau_i<t} \int_0^\infty X_i\big[dN_i(s; t) - Y_i(s; t)d\Lambda\{\gamma Z_i + \beta' X_i + H(s)\}\big] = 0, \tag{3.2b}$$

$$\sum_{i:\tau_i<t} \big[dN_i(s; t) - Y_i(s; t)d\Lambda\{\gamma Z_i + \beta' X_i + H(s)\}\big] = 0. \tag{3.2c}$$

Based on the estimation equation (3.2a), we propose to use the following process to test $H_0 : \gamma = 0$

$$U\{t; \beta, H(\cdot)\} = \sum_{i:\tau_i<t} \int_0^\infty Z_i\big[dN_i(s; t) - Y_i(s; t)d\Lambda\{\beta' X_i + H(s)\}\big], \tag{3.3}$$

An alternative expression of $U\{t; \beta, H(\cdot)\}$, without integral, is

$$U\{t; \beta, H(\cdot)\} = \sum_{i:\tau_i<t} Z_i\big[\delta_i(t) - \Lambda\{\beta' X_i + H(\widetilde{T}_i(t))\}\big]. \tag{3.4}$$

In practice, β and $H(\cdot)$ are unknown. An estimator of β and $H(\cdot)$ can be obtained by the estimating equations (3.2b) and (3.2c) under $H_0 : \gamma = 0$.

$$\sum_{i:\tau_i<t} \int_0^\infty X_i\big[dN_i(s;\,t) - Y_i(s;\,t)d\Lambda\{\beta' X_i + H(s,\beta)\}\big] = 0, \tag{3.5a}$$

$$\sum_{i:\tau_i<t} \big[dN_i(s;\,t) - Y_i(s;\,t)d\Lambda\{\beta' X_i + H(s,\beta)\}\big] = 0. \tag{3.5b}$$

Under the null hypothesis $H_0 : \gamma = 0$, solutions of (3.5a) and (3.5b) provide consistent estimators to β and $H(\cdot)$ as demonstrated in Chen et al. (2002). The estimating equations (3.5a) and (3.5b) can be solved with an iterative algorithm similar to the one in Chen et al. (2002).

We assume that there are $K(t)$ observed distinct failure times, $s_1 < s_2 < \cdots < s_{K(t)}$, at time t with $K(t)$ being a nondecreasing integer sequence depending on time t. Let d_k denote the number of failures at s_k, $(k = 1, 2, \ldots, K(t))$.

The iterative algorithm starts from picking an initial value for unknown parameter β, denoted as β^0.

- Step 1: With $\beta = \beta^0$, solve the equation

$$\sum_{i:\tau_i<t} Y_i(s_1;\,t)\Lambda\{\beta^T Z_i + H(s_1)\} = 1$$

to obtain $H^{(0)}(s_1)$, an estimation of $H(s_1)$. Then, solve the following equations one by one from $k = 2$ to $k = K(t)$,

$$\sum_{i:\tau_i<t} Y_i(s_k;\,t)\Lambda\{\beta^T Z_i + H(s_k)\} = d_k + \sum_{i:\tau_i<t} Y_i(s_k;\,t)\Lambda\{\beta^T Z_i + H(s_k-)\}$$

with $\beta = \beta^0$ to acquire $H^{(0)}(s_2), \ldots, H^{(0)}(s_{K(t)})$.
- Step 2: Set $H(\cdot) = H^{(0)}(\cdot)$ as a stepwise function. A new estimate of β, denoted as β^1, can be attained by solving the estimating equation (3.5a):

$$\sum_{i:\tau_i<t} \int_0^\infty X_i\big[dN_i(s;\,t) - Y_i(s;\,t)d\Lambda\{\beta' X_i + H(s,\beta)\}\big] = 0.$$

- Step 3: Reset $\beta^0 = \beta^1$, then repeat step 1 and step 2 until convergence.

Let $\hat\beta_t$ and $\widehat{H}(., \hat\beta_t)$ be the estimators obtained by the above iterative algorithm at time t. Plugging them into $U(t, \beta, H(\cdot))$, we obtain the test statistic $U\{t;\,\hat\beta_t, \widehat{H}(.;\,\hat\beta_t)\}$, for the null hypothesis $\gamma = 0$:

$$U\{t, \hat\beta_t, H(., \hat\beta_t))\} = \sum_{i:\tau_i<t} Z_i\big[\delta_i(t) - \Lambda\{\hat\beta_t' X_i + \hat{H}(\widetilde{T}_i(t), \hat\beta_t))\}\big].$$

For simplicity of notation, let $\widehat{U}(t)$ denote $U\{t, \hat\beta_t, \widehat{H}(.;\,\hat\beta_t)\}$.

3.1. Asymptotical Properties of $\widehat{U}(t)$

It is essential to know the distribution of the test statistic over time in order to conduct a sequential test. If β and $H(\cdot)$ were known, then the test statistic could be $U\{t; \beta, H(\cdot)\} = \sum_{i:\tau_i<t} Z_i\big[\delta_i(t) - \Lambda\{\beta^T X_i + H(\widetilde{T}_i(t))\}\big]$, which is a sum of identically and independently distributed random variables at any given time t and its asymptotic distribution can be obtained by the multivariate central limit theorem. However, β and $H(\cdot)$ are unknown in practice and their consistent estimators $\hat{\beta}_t$ and $\widehat{H}(.; \hat{\beta}_t)$ are needed.

We will show in Appendix A that when the two covariates Z and X are independent, $n^{-1/2}\widehat{U}(t; \hat{\beta}_t, \widehat{H}(.; \hat{\beta}_t))$ is asymptotically equivalent to $n^{-1/2}\widehat{U}(t; \beta, \widehat{H}(.; \beta))$. $n^{-1/2}\widehat{U}(t; \beta, \widehat{H}(.; \beta))$ can be written as two parts, one is a sum of integrals of identical and independent martingales and the other is a term which goes to 0 when sample size goes to infinity.

We follow the notations and regularity conditions in Chen et al.'s (2002) paper. Define β_0 and H_0 as the true value of β and H. Let $\lambda(\cdot)$ be the derivative of $\Lambda(\cdot)$, $\lambda'(\cdot)$ be the derivative of $\lambda(\cdot)$, and $\tau = \inf\{t : P(\widetilde{T} > t) = 0\}$. Define

$$
\begin{aligned}
B(s, u; t) &= \exp\left(\int_u^s \frac{E[\lambda'\{\beta_0^T X + H_0(x)\}Y(x; t)]}{E[\lambda\{\beta_0^T X + H_0(x)\}Y(x; t)]} dH_0(x)\right),\\
\mu_Z(s; t) &= \frac{E[Z\lambda'\{\beta_0^T X + H_0(s)\}Y(s; t)B(s, \tilde{T}; t)]}{E[\lambda\{\beta_0^T X + H_0(s)\}Y(s; t)]},\\
\sigma^2(t, t) &= \int_0^\tau E[\{Z - \mu_Z(s; t)\}^2 \lambda\{\beta_0^T X + H_0(s)\}Y(s; t)]dH_0(s),\\
\sigma^2(t_1, t_2) &= \int_0^\tau E[\{Z - \mu_Z(s; t_1)\}\{Z - \mu_Z(s; t_2)\}\lambda\{\beta_0^T X + H_0(s)\}Y(s; t_1)]dH_0(s).\\
&\quad (t_1 < t_2)
\end{aligned}
$$

The asymptotic properties of the process $\widehat{U}(\cdot)$ is given in following theorem.

Theorem 3.1. *Under the null hypothesis $H_0 : \gamma = 0$, for any given K and time sequence $t_1 < t_2 < \cdots < t_K$, $\{n^{-1/2}\widehat{U}(t_1), n^{-1/2}\widehat{U}(t_2), \ldots, n^{-1/2}\widehat{U}(t_K)\}$ converges weakly to a K-dimensional normal vector with 0 mean and covariance matrix $\{\sigma^2(t_l, t_m); l, m = 1, \ldots, K\}$, where $\sigma^2(t_l, t_m)$ can be consistently estimated by*

$$
\begin{aligned}
\hat{\sigma}^2(t_l, t_m) = \frac{1}{n}\sum_{i=1}^n \int_0^\tau &\big(Z_i - \bar{Z}(s; t_l)\big)\big(Z_i - \overline{Z}(s; t_m)\big)\\
&\times 1\{\widetilde{T}_i(t_l \wedge t_m) > s\}\lambda\{\hat{\beta}^T X_i + \widehat{H}(s)\}d\widehat{H}(s),
\end{aligned}
$$

where

$$
\begin{aligned}
\overline{Z}(s; t) &= \frac{\sum_i Z_i \lambda\{\hat{\beta}^T X_i + \widehat{H}(\tilde{T}_i(t))\}Y(s; t)\widehat{B}(s, \widetilde{T}_i(t))}{\sum_i Z_i \lambda\{\hat{\beta}^T X_i + \widehat{H}(\widetilde{T}_i(t))\}Y(s; t)},\\
\widehat{B}(s, \widetilde{T}_i(t)) &= \exp\left(\int_{\widetilde{T}_i(t)}^s \frac{\sum_j \lambda'\{\hat{\beta}^T X_j + \widehat{H}(x)\}Y(s; t)}{\sum_j \lambda\{\hat{\beta}^T X_j + \widehat{H}(x)\}Y(s; t)} d\widehat{H}(x)\right).
\end{aligned}
$$

Remark 3.1. Except for the special case of the Cox proportional hazards model when the underlying cumulative hazard function of ϵ is $\exp(t)$, the score process $n^{-1/2}\widehat{U}(t; \hat{\beta}_t, \widehat{H}(.; \hat{\beta}_t))$ do not have independent increment property in general. See the proof in Appendix B for more details.

3.2. Repeated Significance Test

Based on Theorem 3.1, a sequential test can be constructed with rejection boundaries proposed by Slud and Wei (1982).

Let W_i $(i = 1, 2, \ldots, K)$ be the standardized testing statistics of $n^{-1/2}\widehat{U}(t_i)$

$$W_i = n^{-1/2}\widehat{U}(t_i)/\hat{\sigma}(t_i, t_i).$$

Then, the variance of W_i is 1 and the covariance of W_i and W_j

$$Cov(W_i, W_j) = \hat{\sigma}^2(t_i, t_j)/\hat{\sigma}(t_i, t_i)\hat{\sigma}(t_j, t_j).$$

Let $t_1, \ldots, t_K$ denote the time sequence at which interim analyses will be conducted. Split the overall significance level α into K parts $\alpha_1, \ldots, \alpha_K$ with $\alpha = \alpha_1 + \cdots + \alpha_K$, where α_i acts as the significance level at interim time t_i.

At the ith $(i = 1, 2, \ldots, K)$ interim analysis, the Slud and Wei rejection boundary c_i can be calculated by solving the equation $P\{|W_1| < c_1, \ldots, |W_{i-1}| < c_{i-1}, |W_i| \geq c_i\} = \alpha_i$ under the null hypothesis. c_i is the only unknown variable in this equation. If the absolute value of W_i is greater than c_i, the trial should be ended and the null hypothesis is rejected; otherwise, the trial continues to the next interim analysis until the null hypothesis is rejected or the trial reaches the final analysis.

The Slud and Wei rejection boundaries $c_1, \ldots, c_K$ calculation equations can be expressed in the following way:

$$\begin{aligned} P_{H_0}\{|W_1| > c_1\} &= \alpha_1. \\ P_{H_0}\{|W_1| < c_1, |W_2| > c_2\} &= \alpha_2. \\ &\vdots \\ P_{H_0}\{|W_1| < c_1, \ldots, |W_{K-1}| < c_{K-1}, |W_K| \geq c_K\} &= \alpha_K. \end{aligned}$$

Solving these equations involve calculating multivariate normal probability. In our simulations and example presented in Section 4, we used the *QSIMVNV* function (Genz, 1992) in MATLAB.

3.3. $\widehat{U}(t)$ under the Cox Proportional Hazards Model

The Cox model is the most widely used model in survival analysis which is a special case of linear transformation models. As we mentioned in Section 1, several sequential methods have been developed based on the Cox proportional hazards model, to test treatment effect after adjustment of other covariates. It is interesting to study the behavior of the unified test statistic $\widehat{U}(t)$ under this special case.

Let the cumulative hazard function $\Lambda(t)$ of ϵ equal to $\exp(t)$. Then the linear transformation model (2.1) becomes equivalent to the Cox model $\lambda(s \mid Z, X) = \lambda_0(s)\exp(\gamma Z + \beta^T X)$. It is easy to verify that under $H_0 : \gamma = 0$, the estimating equation (3.5b) has the following form:

$$d\exp[H(s, \beta)] = \sum_{i:\tau_i<t} dN_i(s; t) \Big/ \sum_{i:\tau_i<t} Y_i(s; t)\exp\{\beta^T X_i\} \tag{3.6}$$

Plugging the above equation into the estimating equation (3.5a) under the Cox model, we obtain the equation

$$\sum_{i:\tau_i<t} \int_0^\infty \left\{ X_i - \frac{\sum_{j=1}^n X_j Y_i(s; t)\exp(\beta^T X_j)}{\sum_{j=1}^n Y_i(s; t)\exp(\beta^T X_j)} \right\} dN_i(s; t) = 0. \tag{3.7}$$

Equation (3.7) involves unknown parameter, β. It is easy to see that the left side of this equation is the same as the Cox's partial likelihood score with respect to β. Thus, the solution of equation (3.7), $\hat{\beta}_t$, is the maximum partial likelihood estimator under H_0 at time t.

Plugging equation (3.6) and $\hat{\beta}_t$ into the test statistics $\widehat{U}(t)$, we obtain

$$\begin{aligned}
\widehat{U}(t) &= \sum_{i:\tau_i<t} \int_0^\infty Z_i \left[dN_i(s; t) - Y_i(s; t)\exp\{\hat{\beta}_t^T X_i\} d\exp\{H(s)\} \right] \\
&= \sum_{i:\tau_i<t} \int_0^\infty \left\{ Z_i - \frac{\sum_{j=1}^n Z_j Y_i(s; t)\exp(\hat{\beta}_t^T X_j)}{\sum_{j=1}^n Y_i(s; t)\exp(\hat{\beta}_t^T X_j)} \right\} dN_i(s; t) \\
&= \sum_{i:\tau_i<t} \delta_i(t) \left\{ Z_i - \frac{\sum_{j=1}^n Z_j Y_i(s; t)\exp(\hat{\beta}_t^T X_j)}{\sum_{j=1}^n Y_i(s; t)\exp(\hat{\beta}_t^T X_j)} \right\}.
\end{aligned} \tag{3.8}$$

Note that the log partial likelihood function for a Cox model at time t has the form

$$l(t; \gamma, \beta) = \sum_{i:\tau_i<t} \delta_i(t) \left[\gamma Z_i + \beta^T X_i - \log\left\{ \sum_{j=1}^n Y_i(s; t)\exp(\gamma Z_i + \beta^T X_j) \right\} \right] \tag{3.9}$$

By some simple algebra, it can be verified that the test statistic (3.8) is the same as Cox's partial likelihood score with respect to γ with $\beta = \hat{\beta}_t$.

$$\widehat{U}(t) = \frac{\partial}{\partial \gamma} l(t; \gamma, \beta) \bigg|_{\gamma=0, \beta=\hat{\beta}_t} \tag{3.10}$$

The result reveals that the test statistics have independent increments under the special case of Cox model; thus Pocock's (1977) and O'Brien and Fleming's (1979) boundaries can be validly used and multivariate normal integral can be avoided.

4. SIMULATIONS AND EXAMPLES

4.1. Simulations

Simulations were based on model $H(T_i) = -\gamma Z_i - \beta^T X_i + \epsilon_i$. We chose $H_0(t) = \log(t)$ and the hazard function of ϵ being $\lambda(t) = \exp(t)/\{1 + r\exp(t)\}$. The treatment indicator Z was generated from a Bernoulli distribution with $p = 0.5$ and covariate X was generated from a standard normal distribution. The entry time τ was from a uniform distribution $U(0, 5)$. The censoring time was independently generated from $U(0, 10)$. Five interim analyses were conducted at times 3, 4, 5, 6, and 7. The overall type I error was set to 0.05, and an equal nominal type I error 0.01 was assigned to each interim analysis. All simulations were based on 10,000 replications. The maximum possible sample size was set to be 200. We conducted nine scenarios of simulations where the value of β being one of $(0, 1, 2)$ and γ being one of $(0, -0.5, -1)$. Three models were used in simulations with $r = (0, 0.5, 1)$. Note that $r = 0$ specifies the Cox proportional hazards model and $r = 1$ specifies the proportional odds model. Repeated significance test was carried out with the boundary of Slud and Wei (1982). The MATLAB function *QSIMVNV* developed by Genz (1992) was used to calculate multivariate normal probabilities for obtaining Slud and Wei boundaries.

To assess the performance of the proposed method, each simulated dataset was analyzed by three methods: the $\widehat{U}(t)$ proposed in Section 3, traditional logrank test without covariates adjustment, and the test based on Cox's partial likelihood with covariates adjustment.

The simulation results are presented in Tables 1–3.

The simulations showed that when the underlying model is the Cox proportional hazard model, the proposed method provided very similar results to

Table 1. Operating characteristic in the simulation studies

		$\beta = 0$			$\beta = 1$			$\beta = 2$		
r	γ	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox
0	0	0.0428	0.0400	0.0396	0.0404	0.0419	0.0424	0.0413	0.0398	0.0381
		(0.0020)	(0.0020)	(0.0019)	(0.0020)	(0.0020)	(0.0020)	(0.0020)	(0.0019)	(0.0019)
	−0.5	0.8251	0.8337	0.8292	0.7896	0.4920	0.7965	0.7318	0.2328	0.7437
		(0.0038)	(0.0037)	(0.0038)	(0.0041)	(0.0050)	(0.0040)	(0.0044)	(0.0042)	(0.0044)
	−1	1.0000	1.0000	1.0000	0.9998	0.9780	0.9998	0.9995	0.6991	0.9993
		(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0015)	(0.0001)	(0.0002)	(0.0046)	(0.0003)
0.5	0	0.0482	0.0432	0.0425	0.0521	0.0433	0.0481	0.0615	0.0406	0.0653
		(0.0021)	(0.0020)	(0.0020)	(0.0022)	(0.0020)	(0.0021)	(0.0024)	(0.0020)	(0.0025)
	−0.5	0.6109	0.5637	0.5585	0.5944	0.3830	0.5489	0.5839	0.2052	0.5211
		(0.0049)	(0.0050)	(0.0050)	(0.0049)	(0.0049)	(0.0050)	(0.0049)	(0.0040)	(0.0050)
	−1	0.9928	0.9911	0.9902	0.9904	0.9222	0.9839	0.9851	0.6396	0.9719
		(0.0008)	(0.0009)	(0.0010)	(0.0010)	(0.0027)	(0.0013)	(0.0012)	(0.0048)	(0.0016)
1	0	0.0461	0.0424	0.0410	0.0509	0.0409	0.0504	0.0576	0.0412	0.0735
		(0.0021)	(0.0020)	(0.0020)	(0.0022)	(0.0020)	(0.0022)	(0.0023)	(0.0020)	(0.0026)
	−0.5	0.4494	0.4040	0.3940	0.4319	0.3053	0.3883	0.4123	0.1812	0.3774
		(0.0050)	(0.0049)	(0.0049)	(0.0049)	(0.0046)	(0.0049)	(0.0049)	(0.0038)	(0.0048)
	−1	0.9548	0.9371	0.9347	0.9400	0.8378	0.9115	0.9205	0.5729	0.8698
		(0.0021)	(0.0024)	(0.0025)	(0.0024)	(0.0037)	(0.0028)	(0.0027)	(0.0049)	(0.0034)

The values without parentheses are empirical type I errors when $\gamma = 0$ and are empirical powers when $\gamma = -0.5$ or $\gamma = -1$ respectively.

The values within parentheses are empirical standard errors of the type I errors and powers of simulation.

Table 2. Average sample size of simulation studies

		$\beta = 0$			$\beta = 1$			$\beta = 2$		
r	γ	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox
0	0	198	198	198	198	198	198	198	199	198
		(11.3)	(10.7)	(10.3)	(11.1)	(11.0)	(10.9)	(11.2)	(10.6)	(10.3)
	−0.5	163	163	164	165	181	165	168	191	167
		(35.7)	(35.4)	(35.4)	(35.6)	(31.4)	(35.7)	(35.1)	(23.1)	(35.4)
	−1	124	124	124	124	138	124	125	167	125
		(14.6)	(14.2)	(14.7)	(15.2)	(29.3)	(15.4)	(16.7)	(35.9)	(16.4)
0.5	0	199	198	198	199	198	198	199	198	197
		(9.5)	(11.0)	(10.6)	(9.0)	(11.0)	(11.7)	(9.0)	(10.5)	(13.3)
	−0.5	181	178	179	182	185	178	183	192	179
		(31.3)	(32.6)	(32.4)	(29.9)	(28.5)	(32.4)	(28.5)	(21.7)	(32.4)
	−1	137	134	135	140	148	136	143	171	138
		(27.6)	(26.3)	(26.9)	(28.9)	(34.0)	(27.9)	(30.1)	(35.1)	(29.6)
1	0	199	198	198	199	198	198	199	198	197
		(9.1)	(10.9)	(10.7)	(9.1)	(10.6)	(11.8)	(8.8)	(10.7)	(14.2)
	−0.5	187	185	186	188	189	185	189	193	185
		(27.3)	(29.0)	(28.3)	(25.9)	(25.8)	(28.5)	(24.1)	(20.8)	(28.5)
	−1	150	146	147	153	158	149	158	175	153
		(33.7)	(33.2)	(33.4)	(34.1)	(35.9)	(34.4)	(34.2)	(34.0)	(35.5)

The maximal sample size for each simulation is 200.
The values within parentheses are the empirical standard deviation for sample size.

Table 3. Average stopping time of simulation studies

		$\beta = 0$			$\beta = 1$			$\beta = 2$		
r	γ	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox	$\widehat{U}(t)$	Logrank	Cox
0	0	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9
		(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)	(0.6)
	−0.5	2.5	2.5	2.6	2.7	3.6	2.7	2.9	4.4	2.8
		(1.5)	(1.5)	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)	(1.3)	(1.6)
	−1	1.1	1.1	1.1	1.1	1.5	1.1	1.1	2.9	1.1
		(0.3)	(0.3)	(0.3)	(0.4)	(1.0)	(0.4)	(0.4)	(1.7)	(0.4)
0.5	0	4.9)	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.8
		(0.5)	(0.6)	(0.6)	(0.5)	(0.6)	(0.7)	(0.6)	(0.6)	(0.8)
	−0.5	3.5	3.5	3.5	3.6	4.0	3.5	3.6	4.4	3.5
		(1.6)	(1.6)	(1.6)	(1.5)	(1.5)	(1.6)	(1.5)	(1.2)	(1.7)
	−1	1.5	1.4	1.4	1.5	1.9	1.5	1.6	3.1	1.5
		(0.9)	(0.8)	(0.8)	(0.9)	(1.3)	(0.9)	(0.9)	(1.7)	(1.0)
1	0	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.8
		(0.5)	(0.6)	(0.6)	(0.5)	(0.6)	(0.7)	(0.5)	(0.6)	(0.8)
		3.9	3.9	3.9	4.0	4.2	3.9	4.1	4.5	3.9
	−0.5	(1.5)	(1.5)	(1.5)	(1.4)	(1.4)	(1.5)	(1.4)	(1.2)	(1.5)
	−1	1.9	1.8	1.9	2.0	2.4	2.0	2.2	3.3	2.2
		(1.2)	(1.2)	(1.2)	(1.2)	(1.5)	(1.3)	(1.3)	(1.7)	(1.4)

Total five analyses (four interim and one final) were planned.
The values within parentheses are empirical standard deviation for the number of analyses conducted.

the test based on the Cox's partial likelihood, which has been proved theoretically. When the underlying model is not the Cox proportional hazards model, the proposed method showed some increased power than the Cox model. The biggest increase in our simulation cases was about 0.06 (0.5839 vs. 0.5211). When there was no other covariate in the model ($\beta = 0$), the three methods performs similar. When $\beta \neq 0$ (i.e., covariate X has effect on survival time), the proposed method had better performance compared to the logrank test: the larger β was, the better performance was in the sense of higher powers (Table 1) and the smaller sample sizes (Table 2). Table 3 shows that the average stopping time were similar for the three methods.

4.2. An Example

We applied the proposed sequential method to the prostatic cancer data in Byar (1985). The data were from a double-blinded randomized clinical trial. The aim of the study was to compare four treatments for patients with prostatic cancer in Stage 3 and Stage 4. The four treatments were placebo, 0.2 mg diethylstilbestrol (DES), 1.0 mg of DES, and 5 mg of DES. Total 506 patients were recruited into the trial sequentially from 1967 to 1969. We chose the subgroup of 253 patients who were in either placebo group or 1.0 mg of DES group and no missing covariates. Events were defined as death from all causes.

We considered the linear transformation model $H(T_i) = -\gamma Z_i - \beta X_i + \epsilon_i$, where Z_i is the indicator of the treatment and X_i is the indicator of Stage 3 or 4. The hazard function of ϵ was chosen to be $\lambda(t) = \exp(t)/\{1 + r\exp(t)\}$. We fitted three models for the data with $r = 0$, $r = 0.5$, and $r = 1$. A total of six analyses (five interim analyses plus a final analysis) were planned at the end of each year from 1969 to 1974. The overall type I error was set to be 0.05, which was distributed evenly among the six analyses. The repeated significance test was conducted with the rejection boundaries based on the method of Slud and Wei (1982).

Table 4 presents the analysis results. Analysis under the three models had consistent results that the trial would stop at the end of 1973 with the null hypothesis rejected. In this example, early stopping did not save any sample size, but it would save one year of the study time.

Table 4. Repeated significant test for prostatic cancer data

			$r = 0$		$r = 0.5$		$r = 1$	
Date of analysis	No. of entered	No. of deaths	Std. test stats	Boundary	Std. test stats	Boundary	Std. test stats	Boundary
Dec. 69	253	74	0.16	2.64	0.16	2.64	0.17	2.64
Dec. 70	253	117	−1.13	2.53	−1.12	2.53	−1.11	2.53
Dec. 71	253	138	−1.17	2.44	−1.16	2.44	−1.15	2.44
Dec. 72	253	158	−2.28	2.33	−2.26	2.33	−2.25	2.33
Dec. 73	253	166	−2.66	2.22	−2.65	2.22	−2.64	2.22
Dec. 74	253	166	—	—	—	—	—	—

5. DISCUSSION

The proposed method is justified with the assumption that the treatment variable is independent of other covariates, which is valid for randomized clinical trials but might not be always true for nonrandomized clinical trials and observational studies, which is of interest for further investigation.

APPENDIX A: PROOF OF ASYMPTOTIC EQUIVALENCE OF $n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\}$ AND $n^{-1/2}U\{t, \beta_0, \widehat{H}(., \beta_0)\}$

We assume regularity conditions as those assumed in Chen et al. (2002) for ensuring the consistency and asymptotical normality of estimations.

First, expand $n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\}$ with respect to β at the true value β_0,

$$\begin{aligned}
&n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\}\\
&\quad = n^{-1/2}U\{t, \beta_0, \widehat{H}(., \beta_0)\} + \frac{1}{n}\frac{\partial}{\partial\beta}U\{t, \beta, \widehat{H}(., \beta)\}\big|_{\beta=\beta_0}\big[n^{1/2}(\hat{\beta}_t - \beta_0)\big]\\
&\qquad + o_p(\max\{1, n^{1/2}|\hat{\beta}_t - \beta_0|\}),
\end{aligned}$$

where

$$\begin{aligned}
&\frac{1}{n}\frac{\partial}{\partial\beta}U\{t, \beta, \widehat{H}(., \beta)\}\big|_{\beta=\beta_0}\\
&\quad = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau} -Z_iY_i(s; t)d\lambda\{\beta_0'X_i + \widehat{H}(s, \beta_0)\}\left[X_i + \frac{\partial}{\partial\beta}\widehat{H}(s, \beta)|_{\beta=\beta_0}\right]^T\\
&\quad \to -E\left[\int_0^{\tau} ZY(s; t)d\lambda\{\beta_0'X + \widehat{H}(s, \beta_0)\}\left[X + \frac{\partial}{\partial\beta}\widehat{H}(s, \beta)|_{\beta=\beta_0}\right]^T\right]\\
&\quad = -(EZ)E\left[\int_0^{\tau} Y(s; t)d\lambda\{\beta_0'X + \widehat{H}(s, \beta_0)\}\left[X + \frac{\partial}{\partial\beta}\widehat{H}(s, \beta)|_{\beta=\beta_0}\right]^T\right]
\end{aligned}$$

The last equation holds because our assumption that under H_0, the covariate Z is independent with all other random variables.

Under $H_0 : \gamma = 0$, from the third estimation equation, we have

$$\sum_{i=1}^{n}[dN_i(s; t) - Y_i(s; t)d\Lambda\{\beta X_i + \widehat{H}(s, \beta)\}] = 0.$$

Differentiating the above equation with respect to β, we have

$$\sum_{i=1}^{n} Y_i(s; t)\left\{d\lambda\{\beta_0'X_i + \widehat{H}(s, \beta_0)\}\left[X_i + \frac{\partial}{\partial\beta}\widehat{H}(s, \beta)|_{\beta=\beta_0}\right]^T\right\} = 0$$

Thus,

$$E\left[Y(s; t)d\lambda\{\beta_0'X + \widehat{H}(s, \beta_0)\}\left[X + \frac{\partial}{\partial\beta}\widehat{H}(s, \beta)|_{\beta=\beta_0}\right]^T\right] = 0.$$

Plugging the above result into the relation (2.6) above, we have that

$$\frac{1}{n}\frac{\partial}{\partial\beta}U\{t,\beta,\widehat{H}(.,\beta)\}\big|_{\beta=\beta_0}\xrightarrow{P}0.$$

Then we get

$$\begin{aligned}&n^{-1/2}U\{t,\hat{\beta}_t,\widehat{H}(.,\hat{\beta}_t)\}\\&\quad=n^{-1/2}U\{t,\beta_0,\widehat{H}(.,\hat{\beta}_0)\}+o_p(\max\{1,n^{1/2}|\hat{\beta}_t-\beta_0|\})\end{aligned}$$

APPENDIX B: PROOF OF THE THEOREM

Let $a>0$ and b be fixed finite numbers. Define

$$\begin{aligned}B(u,s;t)&=\exp\left(\int_u^s\frac{E[\lambda'\{\beta_0'X+H_0(x)\}Y(x;t)]}{E[\lambda\{\beta_0'X+H_0(x)\}Y(x;t)]}dH_0(x)\right)\\\lambda^*\{H_0(s);t\}&=B(a,s;t)\\B_1(s;t)&=\int_a^s E[\lambda'\{\beta_0'X+H_0(x)\}Y(x;t)]dH_0(x)\\B_2(s;t)&=E[\lambda\{\beta'X+H_0(s)\}Y(s;t)]\\\Lambda^*(x;t)&=\int_b^x\lambda^*(s;t)ds.\end{aligned}$$

Then we will see easily that

$$B(u,s;t)=\frac{\lambda^*\{H_0(u);t\}}{\lambda^*\{H_0(s);t\}},$$

and

$$\begin{aligned}d\lambda^*\{H_0(s);t\}&=dB(a,s;t)\\&=\lambda^*\{H_0(s);t\}\frac{E[\lambda'\{\beta_0'X+H_0(s)\}Y(s;t)]}{E[\lambda\{\beta_0'X+H_0(s)\}Y(s;t)]}dH_0(s)\\&=\frac{\lambda^*\{H_0(s);t\}}{B_2(s;t)}dB_1(s;t)\end{aligned}$$

Then following Step A2 in the Appendix of Chen et al. (2002), we have

$$\begin{aligned}-\frac{1}{n}\sum_{i=1}^n M_i(s;t)&=-\frac{1}{n}\sum_{i=1}^n\int_0^s[dN_i(x;t)-Y_i(x;t)d\Lambda\{\beta_0'X_i+H_0(x)\}\\&=-\frac{1}{n}\sum_{i=1}^n\int_0^s Y_i(x;t)[d\Lambda\{\beta_0'X_i+\widehat{H}(x)\}-d\Lambda\{\beta_0'X_i+H_0(x)\}]\end{aligned}$$

According to the estimation equation

$$\sum_{i=1}^n dN_i(x;t)-Y_i(x;t)d\Lambda\{\beta_0'X_i+\widehat{H}(x)\}=0)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\int_0^s Y(x;t)d\left(\frac{\lambda\{\beta_0'X_i + H_0(x)\}}{\lambda^*\{H_0(x);t\}}\left[\Lambda^*\{\widehat{H}(x);t\} - \Lambda^*\{H_0(x);t\}\right]\right)$$
$$+ o_p\left(\frac{1}{\sqrt{n}}\right)$$
$$= -\int_0^s \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i(x;t)\lambda\{\beta_0'X_i + H_0(x)\}}{\lambda^*\{H_0(x);t\}}\left[d\Lambda^*\{\widehat{H}(x);t\} - d\Lambda^*\{H_0(x);t\}\right]$$
$$- \int_0^s \left[\Lambda^*\{\widehat{H}(x);t\} - \Lambda^*\{H_0(x);t\}\right]\frac{1}{n}\sum_{i=1}^{n}Y_i(x;t)d\left[\frac{\lambda\{\beta_0'X_i + H_0(x)\}}{\lambda^*\{H_0(x);t\}}\right]$$
$$+ o_p\left(\frac{1}{\sqrt{n}}\right)$$
$$= -\int_0^s \frac{B_2(x;t)}{\lambda^*\{H_0(x);t\}}\left[d\Lambda^*\{\widehat{H}(x);t\} - d\Lambda^*\{H_0(x);t\}\right]$$
$$+ \int_0^s \left[\Lambda^*\{\widehat{H}(x);t\} - \Lambda^*\{H_0(x);t\}\right]$$
$$\times \frac{\lambda^*\{H_0(x);t\}dB_1(x;t) - B_2(x;t)d\lambda^*\{H_0(x);t\}}{\left[\lambda^*\{H_0(x);t\}\right]^2} + o_p\left(\frac{1}{\sqrt{n}}\right) \tag{B.1}$$

Recall that

$$d\lambda^*\{H_0(s);t\} = \frac{\lambda^*\{H_0(s);t\}}{B_2(s;t)}dB_1(s;t) \tag{B.2}$$

Plugging (B.2) into (B.1) we get the following equation

$$-\frac{1}{n}\sum_{i=1}^{n}M_i(s;t) = -\int_0^s \frac{B_2(x;t)}{\lambda^*\{H_0(x);t\}}\left[d\Lambda^*\{\widehat{H}(x);t\} - d\Lambda^*\{H_0(x);t\}\right] + o_p\left(\frac{1}{\sqrt{n}}\right).$$

Thus, we have

$$\Lambda^*\{\widehat{H}(s);t\} - d\Lambda^*\{H_0(s);t\} = \frac{1}{n}\sum_{1}^{n}\int_0^s \frac{\lambda^*\{H_0(x);t\}}{B_2(x;t)}dM_i(x;t) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

After all of these preparations, we now study the asymptotic behaviors of the test process $n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\}$. Since we have proved that $n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\}$ and $n^{-1/2}U\{t, \beta_0, \widehat{H}(., \beta_0)\}$ are asymptotically equivalent, then we can study $n^{-1/2}U\{t, \beta_0, \widehat{H}(., \beta_0)\}$ instead.

Start from the definition of $U\{t, \beta_0, H(., \beta_0)\}$,

$$U\{t, \beta_0, \widehat{H}(., \beta_0)\} = \sum_{i=1}^{n}\int_0^\tau Z_i\left[dN_i(s;t) - Y_i(s;t)d\Lambda\{\beta'X_i + \widehat{H}(s;\beta_0)\}\right]$$
$$= \sum_{i=1}^{n}\int_0^\tau Z_i dM_i(x;t) - \sum_{i=1}^{n}Z_i\bigg(\Lambda\{\beta_0'X_i + \widehat{H}(\widetilde{T}_i(t), \beta_0)\}$$
$$- \Lambda\{\beta_0'X_i + H_0(\widetilde{T}_i(t))\}\bigg)$$

$$
\begin{aligned}
&= \sum_{i=1}^{n} \int_0^{\tau} Z_i dM_i(x; t) - \sum_{i=1}^{n} \frac{Z_i \lambda\{\beta_0' X_i + H_0(\tilde{T}_i(t))\}}{\lambda^*\{H_0(\widetilde{T}_i(t)); t\}} \big[\Lambda^*\{\widehat{H}(\widetilde{T}_i(t)); t\} \\
&\quad - \Lambda^*\{H_0(\widetilde{T}_i(t)); t\}\big] + o_p(\sqrt{n}) \\
&= \sum_{i=1}^{n} \int_0^{\tau} Z_i dM_i(x; t) - \sum_{i=1}^{n} \frac{Z_i \lambda\{\beta_0' X_i + H_0(\tilde{T}_i(t))\}}{\lambda^*\{H_0(\widetilde{T}_i(t)); t\}} \\
&\quad \frac{1}{n} \sum_{j=1}^{n} \int_0^{\widetilde{T}_i(t)} \frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)} dM_j(s; t) + o_p\Big(\frac{1}{\sqrt{n}}\Big) \\
&= \sum_{i=1}^{n} \int_0^{\tau} Z_i dM_i(x; t) - \sum_{j=1}^{n} \int_0^{\tau} E\left[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t))\}}{\lambda^*\{H_0(\tilde{T}(t)); t\}}\right] \\
&\quad \times \frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)} dM_j(s; t) + o_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \sum_{i=1}^{n} \int_0^{\tau} \left(Z_i - E\left[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t))\}}{\lambda^*\{H_0(\tilde{T}(t)); t\}}\right] \frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)}\right) dM_i(s; t) \\
&\quad + o_p\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&n^{-1/2} U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\} \\
&\quad = n^{-1/2} U\{t, \beta_0, \widehat{H}(., \beta_0)\} + o_p(\max\{1, n^{1/2}|\hat{\beta}_t - \beta_0|\}) \\
&\quad = n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} \left(Z_i - E\left[\frac{Z\lambda\{\beta_0' X + H_0(\widetilde{T}(t))\} Y(s; t)}{\lambda^*(H_0(\tilde{T}(t)))}\right] \frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)}\right) dM_i(s; t) \\
&\qquad + o_p(\max\{1, n^{1/2}|\hat{\beta}_t - \beta_0|\}),
\end{aligned}
$$

which means that the proposed test process is asymptotically equivalent to a sum of independent and identically distributed random variables for a fixed t. Let $\xi_i(t)$ be a single term of the summation; that is,

$$
\xi_i(t) = \int_0^{\tau} \left(Z_i - E\left[\frac{Z\lambda\{\beta_0' X + H_0(\widetilde{T}(t))\} Y(s; t)}{\lambda^*(H_0(\tilde{T}(t)))}\right] \frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)}\right) dM_i(s; t).
$$

Then $n^{-1/2} U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)\} = n^{-1/2} \sum_{i=1}^{n} \xi_i(t) + o_p(1)$ since $n^{1/2}(\hat{\beta}_t - \beta_0) = O_p(1)$.

Therefore, for any given time points $t_1, t_2, \ldots, t_K$,

$$
n^{-1/2} \begin{pmatrix} U\{t_1, \hat{\beta}_{t_1}, \widehat{H}(., \hat{\beta}_{t_1})\} \\ U\{t_2, \hat{\beta}_{t_2}, \widehat{H}(., \hat{\beta}_{t_2})\} \\ \cdot \\ \cdot \\ \cdot \\ U\{t_K, \hat{\beta}_{t_K}, \widehat{H}(., \hat{\beta}_{t_K})\} \end{pmatrix} = n^{-1/2} \sum_{i=1}^{n} \begin{pmatrix} \xi_i(t_1) \\ \xi_i(t_2) \\ \cdot \\ \cdot \\ \cdot \\ \xi_i(t_K) \end{pmatrix} + o_p(1).
$$

By the standard multivariate central limit theorem, $n^{-1/2} \sum_{i=1}^{n} (\xi_i(t_1), \xi_i(t_2), \ldots, \xi_i(t_K))^{\mathrm{T}}$ converges weakly to a K-dimensional normal vector with mean 0 and

$K \times K$ covariance matrix $\{\sigma^2(t_l, t_m); l, m = 1, \ldots, K\}$, where $\sigma^2(t_l, t_m) = cov(\xi_1(t_l), \xi_1(t_m))$. Consequently, $(n^{-1/2}U\{t_1, \hat{\beta}_{t_1}, \widehat{H}(., \hat{\beta}_{t_1})\}, n^{-1/2}U\{t_2, \hat{\beta}_{t_2}, \widehat{H}(., \hat{\beta}_{t_2})\}, \ldots, n^{-1/2}U\{t_K, \hat{\beta}_{t_K}, \widehat{H}(., \hat{\beta}_{t_K})\})$ converges weakly to a K-dimensional normal vector.

The variance of $\xi_i(t)$ is

$$Var(\xi_i(t)) = \int_0^\tau E\left[\left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\widetilde{T}(t))\}Y(s;t)}{\lambda^*(H_0(\widetilde{T}(t)))}\right]\frac{\lambda^*\{H_0(s);t\}}{B_2(s;t)}\right)^2 1\{\widetilde{T}(t) > s\}\lambda\{\beta_0'X + H_0(s)\}\right] dH_0(s).$$

Since for different i and j, ξ_i and ξ_j are independent, the variance of $n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)$ equals to the variance of ξ_i,

$$Var(n^{-1/2}U\{t, \hat{\beta}_t, \widehat{H}(., \hat{\beta}_t)) = \int_0^\tau E\left[\left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\tilde{T}(t))\}Y(s;t)}{\lambda^*(H_0(\tilde{T}(t)))}\right]\frac{\lambda^*\{H_0(s);t\}}{B_2(s;t)}\right)^2 1\{\widetilde{T}(t) > s\}\lambda\{\beta_0'X + H_0(s)\}\right] dH_0(s).$$

A consistent estimator of this variance can be

$$\hat{\sigma}^2(t,t) = \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left(Z_i - \bar{Z}(s;t)\right)^2 1_{\{\widetilde{T}_i(t) > s\}}\lambda\{\hat{\beta}'X_i + \widehat{H}(s)\}d\widehat{H}(s),$$

where

$$\overline{Z}(s;t) = \frac{\sum_{i=1}^n \lambda\{\hat{\beta}'X_i + \widehat{H}(\tilde{T}_i(t))\}Y(s;t)\widehat{B}(s, \widetilde{T}_i(t))}{\sum_{i=1}^n \lambda\{\hat{\beta}'X_i + \widehat{H}(\widetilde{T}_i(t))\}Y(s;t)},$$

$$\widehat{B}(s, \widetilde{T}_i(t)) = \exp\left(\int_{\widetilde{T}_i(t)}^s \frac{\sum_{j=1}^n Z_i\lambda'\{\hat{\beta}'X_i + \widehat{H}(x)\}Y(s;t)}{\sum_{j=1}^n Z_i\lambda\{\hat{\beta}'X_i + \widehat{H}(x)\}Y(s;t)}d\widehat{H}(x)\right).$$

For any two different time points t_1 and t_2, the $cov(\xi_i(t_1), \xi_i(t_2))$ is

$$cov(\xi_i(t_1), \xi_i(t_2)) = \int_0^\tau E\Bigg\{\left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\tilde{T}(t_1))\}Y(s;t_1)}{\lambda^*(H_0(\tilde{T}(t_1)))}\right]\frac{\lambda^*\{H_0(s);t_1\}}{B_2(s;t_1)}\right) \left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\tilde{T}(t_2))\}Y(s;t_2)}{\lambda^*(H_0(\tilde{T}(t_2)))}\right]\frac{\lambda^*\{H_0(s);t_2\}}{B_2(s;t_2)}\right) 1\{\widetilde{T}(t_1 \wedge t_2) > s\}\lambda\{\beta_0'X + H_0(s)\}\Bigg\} dH_0(s)$$

So the covariance of $n^{-1/2}U(t_1)$ and $n^{-1/2}U(t_2)$ equals

$$cov(\xi_i(t_1), \xi_i(t_2)) = \int_0^\tau E\Bigg\{\left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\tilde{T}(t_1))\}Y(s;t_1)}{\lambda^*(H_0(\tilde{T}(t_1)))}\right]\frac{\lambda^*\{H_0(s);t_1\}}{B_2(s;t_1)}\right) \left(Z - E\left[\frac{Z\lambda\{\beta_0'X + H_0(\tilde{T}(t_2))\}Y(s;t_2)}{\lambda^*(H_0(\tilde{T}(t_2)))}\right]\frac{\lambda^*\{H_0(s);t_2\}}{B_2(s;t_2)}\right) 1\{\widetilde{T}(t_1 \wedge t_2) > s\}\lambda\{\beta_0'X + H_0(s)\}\Bigg\} dH_0(s)$$

A consistent estimator of $cov(\xi_i(t_1), \xi_i(t_2))$ is

$$\hat{\sigma}^2(t_1, t_2) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\Big(Z_i - \bar{Z}(s; t_1)\Big)\Big(Z_i - \overline{Z}(s; t_2)\Big)1_{\{\tilde{T}_i(t_1 \wedge t_2) > s\}}\lambda\{\hat{\beta}' X_i + \widehat{H}(s)\}d\widehat{H}(s).$$

The investigation of whether the process has independent increments also relies on the structure of $\xi_i(t)$. For any two different time points t_1 and t_2 (assume $t_2 > t_1$), the covariance of $n^{-1/2}U(t_1)$ and $n^{-1/2}(U(t_2) - U(t_1))$ equals

$$\begin{aligned} cov(\xi_i(t_1), \xi_i(t_2) - \xi_i(t_1)) = \int_0^{\tau} E\Bigg\{\Bigg(Z - E\Bigg[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t_1))\}Y(s; t_1)}{\lambda^*(H_0(\tilde{T}(t_1)))}\Bigg]\frac{\lambda^*\{H_0(s); t_1\}}{B_2(s; t_1)}\Bigg) \\ \Bigg(E\Bigg[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t_1))\}Y(s; t_1)}{\lambda^*(H_0(\tilde{T}(t_1)))}\Bigg]\frac{\lambda^*\{H_0(s); t_1\}}{B_2(s; t_1)} \\ - E\Bigg[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t_2))\}Y(s; t_2)}{\lambda^*(H_0(\tilde{T}(t_2)))}\Bigg]\frac{\lambda^*\{H_0(s); t_2\}}{B_2(s; t_2)}\Bigg) \\ 1\{\widetilde{T}(t_1) > s\}\lambda\{\beta_0' X + H_0(s)\}\Bigg\}dH_0(s) \end{aligned}$$

Under this general form, $cov(\xi_i(t_1), \xi_i(t_2) - \xi_i(t_1))$ doesnot always equal 0. However, under the special case of the Cox proportional hazards model, the hazard function $\lambda(\cdot)$ of ϵ is $\exp(\cdot)$, which leads to $B_2(s; t) = E(\exp(\beta^T + H_0(s))Y(s; t))$ and $\lambda^*\{H_0(s); t\} = \exp(H_0(s) - H_0(a))$. Thus, it is easy to see that

$$\begin{aligned} & E\Bigg[\frac{Z\lambda\{\beta_0' X + H_0(\tilde{T}(t))\}Y(s; t)}{\lambda^*(H_0(\tilde{T}(t)))}\Bigg]\frac{\lambda^*\{H_0(s); t\}}{B_2(s; t)} \\ &= E\Bigg[\frac{Z\exp\{\beta_0' X + H_0(\tilde{T}(t))\}Y(s; t)}{\exp(H_0(\tilde{T}(t)) - H_0(a))}\Bigg]\frac{\exp(H_0(s) - H_0(a))}{E(\exp(\beta^T + H_0(s))Y(s; t))} \\ &= E[Z\exp\{\beta_0' X + H_0(a)\}Y(s; t)]\frac{1}{E(\exp(\beta^T + H_0(a))Y(s; t))} \\ &= E(Z) \end{aligned}$$

The last equation is based on the assumption that treatment indicator Z is independent of X, and under the null hypothesis, Z is also independent of $Y(s; t)$. Thus, we have, under the Cox proportional hazards model,

$$\xi_i(t) = \int_0^{\tau}(Z_i - E(Z_i))dM_i(s; t).$$

It is easy to see that $cov(\xi_i(t_1), \xi_i(t_2) - \xi_i(t_1)) = 0$.

ACKNOWLEDGMENT

We would like thank the Associate Editor and two anonymous referees for insightful comments which helped improve this article.

REFERENCES

Bilias, Y., Gu, M., and Ying, Z. (1997). Towards a General Asymptotic Theory for Cox Model with Staggered Entry, *Annals of Statistics* 25: 662–682.

Byar, D. P. (1985). *Prognostic Variables for Survival in a Randomized Comparison of Treatments for Prostatic Cancer*, New York: Springer.

Cai, T., Wei, L. J., and Wilcox, M. (2000). Semiparametric Regression Analysis for Clustered Failure Time Data, *Biometrika* 87: 867–878.

Cheng, S. C., Wei, L. J., and Ying, Z. (1995). Analysis of Transformation Models with Censored Data, *Biometrika* 87: 835–843.

Cheng, S. C., Wei, L. J., and Ying, Z. (1997). Prediction of Survival Probabilities with Semi-Parametric Transformation Models, *Journal of American Statistical Association* 92: 227–235.

Chen, K., Jin, Z., and Ying, Z. (2002). Semiparametric Analysis of Transformation Models with Censored Data, *Biometrika* 89: 659–668.

Fine, J. P., Ying, Z., and Wei, L. J. (1998). On the Linear Transformation Model with Censored Data, *Biometrika* 85: 980–986.

Genz, A. (1992). Numerical Computation of Multivariate Normal Probabilities, *Journal of Computational and Graphical Statistics* 1: 141–149.

Gu, M. and Ying, Z. (1993). Sequential Analysis for Censored Regression Data, *Journal of American Statistical Association* 88: 890–898.

Gu, M. and Ying, Z. (1995). Group Sequential Methods for Survival Data Using Partial Likelihood Score Processes with Covariate Adjustment, *Statistica Sinica* 5: 793–804.

Jennison, C. and Turnbull, B. (1997). Group-Sequential Analysis Incorporating Covariate Information, *Journal of American Statistical Association* 92: 1330–1341.

Jones, D. and Whitehead, J. (1979). Sequential Forms of the Logrank and Modified Wilcoxon Tests for Censored Data, *Biometrika* 66: 105–113.

Lin, D. Y. (1992). Sequential Log Rank Tests Adjusting for Covariates with the Accelarated Life Model, *Biometrika* 79: 523–529.

O'Brien, P. and Fleming, T. (1979). A Multiple Testing Procedure for Clinical Trials, *Biometrics* 35: 549–556.

Pocock, S. (1977). Group Sequential Methods in the Design and Analysis of Clinical Trials, *Biometrika* 64: 191–199.

Scharfstein, D. O., Tsiatis, A. A., and Robins, J. M. (1997). Semiparametric Efficiency and Its Implication on the Design and Analysis of Group Sequential Studies, *Journal of American Statistical Association* 92: 1342–1350.

Sellke, T. and Siegmund, D. (1983). Sequential Analysis of the Proportional Hazards Model, *Biometrika* 70: 315–326.

Slud, E. V. (1984). Sequential Linear Rank Tests for Two-Sample Censored Survival Data, *Annals of Statistics* 12: 551–571.

Slud, E. and Wei, L. J. (1982). Two-Sample Repeated Significance Tests Based on the Modified Wilcoxon Statistic, *Journal of American Statistical Association* 77: 862–868.

Tsiatis, A. A., Boucher, H., and Kim, K. (1995). Sequential Methods for Parametric Survival Models, *Biometrika* 82: 165–173.

Tsiatis, A. A., Rosner, G. L., and Tritchler, D. L. (1985). Group Sequential Tests with Censored Survival Data Adjusting for Covariates, *Biometrika* 72: 365–373.

Tsiatis, A. A. (1982). Repeated Signigicance Testing for a General Class of Statistics Used in Censored Survival Analysis, *Journal of American Statistical Association* 77: 855–861.

Zeng, D. and Lin, D. Y. (2006). Efficient Estimation of Semiparametric Transformation Models for Counting Process, *Biometrika* 93: 627–640.

Zeng, D. and Lin, D. Y. (2007). Maximum Likelihood Estimation in Semiparametric Regression Models with Censored Data, *Journal of Royal Statistical Society, Series B* 69: 507–564.