Published in final edited form as:
Biometrics. 2013 December ; 69(4): . doi:10.1111/biom.12097.

Efficient Semiparametric Estimation of Short-term and Long-term Hazard Ratios with Right-Censored Data

Guoqing Diao[1,*], **Donglin Zeng**[2], and **Song Yang**[3]

[1]Department of Statistics, George Mason University, Fairfax, Virginia, U.S.A

[2]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, U.S.A

[3]Office of Biostatistics Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, U.S.A

Summary

The proportional hazards assumption in the commonly used Cox model for censored failure time data is often violated in scientific studies. Yang and Prentice (2005) proposed a novel semiparametric two-sample model that includes the proportional hazards model and the proportional odds model as sub-models, and accommodates crossing survival curves. The model leaves the baseline hazard unspecified and the two model parameters can be interpreted as the short-term and long-term hazard ratios. Inference procedures were developed based on a pseudo score approach. Although extension to accommodate covariates was mentioned, no formal procedures have been provided or proved. Furthermore, the pseudo score approach may not be asymptotically efficient. We study the extension of the short-term and long-term hazard ratio model of Yang and Prentice (2005) to accommodate potentially time-dependent covariates. We develop efficient likelihood-based estimation and inference procedures. The nonparametric maximum likelihood estimators are shown to be consistent, asymptotically normal, and asymptotically efficient. Extensive simulation studies demonstrate that the proposed methods perform well in practical settings. The proposed method successfully captured the phenomenon of crossing hazards in a cancer clinical trial and identified a genetic marker with significant long-term effect missed by using the proportional hazards model on age-at-onset of alcoholism in a genetic study.

Keywords

Semiparametric hazards rate model; Non-parametric likelihood; Proportional hazards model; Proportional odds model; Semiparametric efficiency

1. Introduction

Much of the modern statistical methodology for survival analysis originates from the seminal work of Cox (1972). The Cox proportional hazards model specifies that the hazard function of the event time T given a $p \times 1$ covariate vector $\mathbf{X}$ takes the form

*gdiao@gmu.edu.

Supplementary Materials

Web Appendices A-G, referenced in Sections 2 and 5, are available with this paper at the Biometrics website on Wiley Online Library.

$$\lambda(t|\mathbf{X})=\lambda(t)e^{\beta^T\mathbf{X}}, \quad (1)$$

where $\lambda(t)$ is an unspecified baseline hazard function and $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown regression parameters. The assumption of constant relative risks over time in the Cox model, however, is often violated in many biomedical and genetic studies. For instance, crossing hazards may be commonly observed in certain clinical trials, in which a particularly aggressive treatment such as surgery may have adverse effects initially but may show beneficial results in the long run. In genetic studies, a certain gene may have a large impact on the hazard for children shortly after birth, but may have a relatively small impact later in life. In some other studies, genes related to susceptibility for a certain disease may affect older people more than younger people.

A motivating example is from the Collaborative Study on the Genetics of Alcoholism (COGA), a genetic family study with the aim of identifying and characterizing genetic factors that affect the susceptibility to alcohol dependence and related phenotypes (Hasin, 2003). The investigators were particularly interested in assessing genetic effects on the age at onset of ALDX1, the DSM-III-R+Feighner classification status for alcohol dependence. Recent studies by Wang et al. (2006) and Diao and Lin (2010) suggested that SNP rs1972373 on chromosome 14 might be a disease susceptibility locus. There are three possible genotypes, '1/1', '1/2', and '2/2', at SNP rs1972373. Kaplan-Meier estimates of survival curves for the three genotype groups presented in Figure 1 appear to be overlapping with each other before age of around 25 whereas allele '2' appears to be associated with the risk of alcoholism at older ages. In such situations, the proportional hazards model cannot distinguish short-term and long-term genetic effects. Another interesting example involves data from a randomized clinical trial on the treatment of locally unresectable gastric cancer (Gastrointestinal Tumor Study Group, 1982). The aim of this trial was to compare chemotherapy with the combined chemotherapy and radiotherapy. As shown in Yang and Prentice (2005) and Zeng and Lin (2007), the Kaplan-Meier survival curves for the two treatment groups cross at around 1000 days indicating crossing hazards. The proportional hazards model cannot capture crossing hazards and could yield very misleading results in such situations.

When the assumption of proportional hazards is questionable, an alternative to the Cox model is the proportional odds model (Bennett, 1983; Murphy et al., 1997), which assumes that the relative risk converges to one rather than remaining constant as time increases. The survival function of T given covariates $\mathbf{X}$ under the proportional odds model takes the form

$$S(t|\mathbf{X})=\frac{e^{-\beta^T\mathbf{X}}}{G(t)+e^{-\beta^T\mathbf{X}}}, \quad (2)$$

where $G(\cdot)$ is a strictly increasing function with $G(0) = 0$. Both the proportional hazards and proportional odds models belong to the class of linear transformation models which relate an unknown monotone transformation of the failure time T linearly to the covariates $\mathbf{X}$ (Bickel et al., Ch 3; Zeng and Lin, 2007). The phenomenon of crossing hazards, however, cannot be directly captured by linear transformation models.

To accommodate time-varying covariate effects on survival outcomes, one option is to use manufactured time-dependent covariates involving interactions between covariates and time in the standard Cox model (Hess, 1994; Therneau and Grambsch, 2000). Specifying the right form of the interaction terms can be challenging, particularly when there are multiple

continuous covariates. Alternatively, one can extend the Cox model (1) through the use of time-varying regression coefficients such that

$$\lambda(t|\mathbf{X})=\lambda(t)e^{\beta^T(t)\mathbf{X}},$$

where $\boldsymbol{\beta}(t)$ is a $p \times 1$ vector of unspecified functions of t. Estimation and inference procedures for this so-called varying-coefficient Cox model have been investigated by several authors, including Zucker and Karr (1990), Murphy and Sen (1991), Murphy (1993), Martinussen et al. (2002), Winnett and Sasieni (2003), Cai and Sun (2003), Tian et al. (2005), and Peng and Huang (2007), among others. In general, nonparametric smoothing is required to estimate the time varying coefficients. Other methods that allow for non-constant hazard ratios also include the heteroscedastic version of linear transformation models (Zeng and Lin, 2007) and the gamma frailty model with scale and shape parameters dependent on covariates, which includes the proportional hazards and the proportional odds models as nested models (Hougaard, 2000; Kosorok et al., 2004; Zeng and Lin, 2007). It is worthwhile to note that the interpretations of the parameters in these models can be different.

Yang and Prentice (2005) proposed a novel semiparametric two-sample hazards rate model that accommodates crossing survival curves. Their model leaves the baseline distribution unspecified and the two model parameters have appealing interpretations of the short-term and the long-term hazard ratios, respectively. The authors developed inference procedures based on a pseudo score approach and showed that the estimators are consistent and asymptotically normal. Although extension to accommodate covariates was mentioned, no formal procedures have been provided or proved. In addition, the pseudo score approach may not be asymptotically efficient.

In this paper, we study the extension of the two-sample semiparametric hazards rate model of Yang and Prentice (2005) to accommodate potentially time-dependent covariates. We develop efficient likelihood-based estimation and inference procedures. The estimators are shown to be consistent, asymptotically normal, and asymptotically efficient. The rest of the paper is organized as follows. In section 2, we introduce the semiparametric hazards rate model accommodating potentially time-dependent covariates and formulate the non-parametric likelihood function. In Section 3, we describe the model assumptions and derive the asymptotic theories. Extensive simulations studies are presented in Section 4 to examine the finite sample properties of the proposed method. In Section 5, we illustrate the new model through the applications to a cancer clinical trial and a genetic study on alcoholism. We conclude with a brief discussion in Section 6.

2. Models and Inference

Suppose that there is a random sample of n independent subjects. For the ith subject, let T_i be the failure time, C_i be the censoring time, and $\mathbf{X}_i$ be a $p \times 1$ vector of (time invariant) covariates. The data consist of $\{Y_i = \min(T_i, C_i), \Delta_i = I(T_i \leq C_i), \mathbf{X}_i, i = 1, \ldots, n\}$, where $I(\cdot)$ is the indicator function. Let τ be a constant denoting the end of the study. We assume that T_i and C_i are independent given $\mathbf{X}_i$. We also assume that $P(C_i \geq \tau|\mathbf{X}_i) = P(C_i = \tau|\mathbf{X}_i) > 0$.

To incorporate short-term and long-term covariate effects, Yang and Prentice (2005) discussed the following semiparametric hazards rate model

$$\lambda(t|\mathbf{X}_i)=\frac{e^{(\beta+\gamma)^T\mathbf{X}_i}}{e^{\beta^T\mathbf{X}_i}F(t)+e^{\gamma^T\mathbf{X}_i}S(t)}\lambda(t), \quad (3)$$

where $\lambda(t|\mathbf{X}_i)$ is the hazard function of the event time T_i given $\mathbf{X}_i$, $\lambda(t)$ is the baseline hazard function, $S(t)=\exp\{-\int_0^t\lambda(s)ds\}$ is the baseline survival function, $F(t) = 1 - S(t)$ is the baseline cumulative distribution function, and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are two vectors of unknown regression parameters. The baseline cumulative hazard function $\Lambda(t)\equiv\int_0^t\lambda(s)ds$ is left unspecified. Under this model, the hazard ratios between two sets of covariate values are allowed to be non-constant over time. Particularly, we can show that

$$\lim_{t\to 0}\frac{\lambda(t|\mathbf{X}_1)}{\lambda(t|\mathbf{X}_2)}=e^{\beta^T(\mathbf{X}_1-\mathbf{X}_2)}, \quad \lim_{t\to\tau 0}\frac{\lambda(t|\mathbf{X}_1)}{\lambda(t|\mathbf{X}_2)}=e^{\gamma^T(\mathbf{X}_1-\mathbf{X}_2)},$$

assuming the existence of the limits, where $\tau_0 = \sup\{t : S(t) > 0\}$. Therefore, the parameters $e^{\boldsymbol{\beta}}$ and $e^{\boldsymbol{\gamma}}$ can be interpreted as the short-term and long-term hazard ratios, respectively. Moreover, model (3) includes the proportional hazards and proportional odds models as two sub-models. When $\boldsymbol{\beta} = \boldsymbol{\gamma}$, model (3) reduces to the proportional hazards model (1); and when $\boldsymbol{\gamma} = \mathbf{0}$, model (3) reduces to the proportional odds model (2) with $G(t) = e^{\boldsymbol{\Lambda}(t)} - 1$ and $\boldsymbol{\beta}$ gives the log-odd ratios.

We extend model (3) to allow time-dependent covariates. Let $\mathbf{X}_i(\cdot)$ be a $p \times 1$ vector of (possibly time-dependent) covariates. Also let $\bar{\mathbf{X}}_i(t)$ denote the history of $\mathbf{X}_i(\cdot)$ over $[0, t]$. We assume that the time dependent covariates are external and that $\mathbf{X}_i(\cdot)$ are bounded right-continuous functions with bounded right derivatives in $[0, \tau]$ with probability one. We specify that the cumulative hazard function conditional on $\bar{\mathbf{X}}_i(t)$ takes the form

$$\Lambda(t|\overline{\mathbf{X}}_i(t))=\int_0^t\frac{e^{(\beta+\gamma)^T\mathbf{X}_i(s)}}{e^{\beta^T\mathbf{X}_i(s)}F(s)+e^{\gamma^T\mathbf{X}_i(s)}S(s)}d\Lambda(s), \quad (4)$$

where $\Lambda(t)$, $S(t)$, $F(t)$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ have the same interpretation as those under model (3).

Our goal is to make inference about parameters $\boldsymbol{\theta}\equiv(\boldsymbol{\beta}, \boldsymbol{\gamma})$ and the function $\Lambda(t)$. Under the assumption of conditional independent censoring, the likelihood for $(\boldsymbol{\theta}, \Lambda)$ takes the form

$$\prod_{i=1}^n\left[\frac{e^{(\beta+\gamma)^T\mathbf{X}_i(Y_i)}\Lambda'(Y_i)}{e^{\beta^T\mathbf{X}_i(Y_i)}F(Y_i)+e^{\gamma^T\mathbf{X}_i(Y_i)}S(Y_i)}\right]^{\Delta_i}e^{-\Lambda(Y_i|\overline{\mathbf{X}}_i(Y_i))},$$

where $\Lambda'(t)$ is the first derivative of $\Lambda(t)$.

In order to estimate the unknown parameters, we need to maximize the observed-data likelihood. However, this maximum does not exist because one can always choose $\Lambda'(Y_i) = \infty$ for some Y_i with $\Delta_i = 1$. Thus, we take a nonparametric maximum likelihood approach, in which Λ is allowed to be a right-continuous function. Specifically, we replace $\Lambda'(Y_i)$ with $\Lambda\{Y_i\}$, the jump size of $\Lambda(\cdot)$ at Y_i. Therefore, we obtain the following nonparametric likelihood function

$$L_n(\theta,\Lambda)=\prod_{i=1}^{n}\left[\frac{e^{(\beta+\gamma)^T\mathbf{X}_i(Y_i)}\Lambda\{Y_i\}}{e^{\beta^T\mathbf{X}_i(Y_i)}F(Y_i)+e^{\gamma^T\mathbf{X}_i(Y_i)}S(Y_i)}\right]^{\Delta_i}e^{-\Lambda(Y_i|\overline{\mathbf{X}}_i(Y_i))}. \quad (5)$$

We maximize the nonparametric log-likelihood function $l_n(\phi) \equiv \log L_n(\phi)$. The resultant nonparametric maximum likelihood estimators (NPMLEs) are denoted by $(\hat{\theta}_n, \hat{\Lambda}_n)$. It is easy to show that $\hat{\Lambda}_n$ must be a step function with positive jumps only at the Y_is for which $\Delta_i = 1$. We order the distinct observed failure time as $(Y_{(1)}, \ldots, Y_{(m)})$, where m is the total number of distinct observed failure times. Therefore, the above maximization should be performed over the parameters θ and these positive jumps. The cumulative hazard function $\Lambda(t|\bar{\mathbf{X}}_i(t))$ in (5) takes the form

$$\sum_{k:Y_{(k)}\le t}\frac{e^{(\beta+\gamma)^T\mathbf{X}_i(Y_{(k)})}}{e^{\beta^T\mathbf{X}_i(Y_{(k)})}F(Y_{(k)})+e^{\gamma^T\mathbf{X}_i(Y_{(k)})}S(Y_{(k)})}\Lambda\{Y_{(k)}\}.$$

To compute the NPMLEs, we use the quasi-Newton algorithm described in Chapter 10 of Press et al. (1992). Specifically, we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, which is one of the most efficient method for solving nonlinear optimization problems, and was proposed by Broyden (1970), Fletcher (1970), Goldfarb (1970), and Shanno (1970) individually. The BFGS method and its variants have been implemented in standard software such as SAS, R, and Matlab and have been successfully used in literature. To ensure the stability of the quasi-Newton algorithm, we suggest to center covariates at their means. When we constrain the regression parameters such that $\beta = \gamma$, the quasi-Newton algorithm yields the exactly the same parameter estimates as those from the procedure *phreg* in SAS software and R routine *coxph* under the proportional hazards model; when we constrain $\gamma = \mathbf{0}$, the NPMLEs obtained from the quasi-Newton algorithm are the same as those from R routine *nltm* under the proportional odds model. These results provide an empirical validation of the quasi-Newton algorithm.

In Web Appendix A, we establish consistency and asymptotic normality of the NPMLEs. We show that the asymptotic covariance matrix for $\hat{\theta}_n$ attains the semiparametric efficiency bound and can be consistently estimated using the inverse of the observed Fisher information matrix for all parameters including θ and the jump sizes of $\hat{\Lambda}_n$. Alternatively, following the argument of Murphy and van der Vaart (2000), we can estimate the covariance matrix of $\hat{\theta}_n$ by using the profile likelihood function for θ, which is defined as the maximum likelihood of $L_n(\theta, \Lambda)$ for any fixed θ. Our simulation studies indicated that both approaches work very well in practical situations. Detailed proofs of the asymptotic properties of the NPMLEs are provided in Web Appendices B–F.

The formulation of the semiparametric hazards rate model provides an appealing diagnostic tool for testing the proportional hazards and proportional odds models since the latter two models are embedded in the former. Specifically, we can check the proportional hazards and proportional odds assumptions by testing $H_0 : \beta = \gamma$ and $H_0 : \gamma = \mathbf{0}$, respectively. This can be done by the Wald, score or likelihood ratio statistics.

To judge the goodness of fit of the proposed model and compare fits of different models, we propose to use a Cramér-von Mises type criterion. Specifically, we first define some strata based on the covariate values. We then define the following Cramér-von Mises type criterion (CMC)

$$\sum_{k=1}^{K} n_k \int_0^\tau |\hat{F}_k(t) - \tilde{F}_k(t)|^2 d\tilde{F}_k(t),$$

where K is the number of strata, n_k is the number of subjects in the kth stratum, and $\hat{F_k}(t)$ and $\tilde{F_k}(t)$ are the model fit and the nonparametric estimators of the cumulative distribution function of the failure time for the kth stratum, respectively. Smaller values of CMC indicate a better fit to the data.

3. Simulation Studies

We conducted extensive simulation studies to evaluate the finite sample performance of the proposed methodology using 1000 replicates. We generated failure times from the following model

$$\Lambda(t|X_i) = \int_0^t \frac{e^{(\beta+\gamma)X_i}}{e^{\beta X_i} F(s) + e^{\gamma X_i} S(s)} d\Lambda(s), \quad (6)$$

where X_i is a uniform(−1, 1) variable. The baseline cumulative hazard function is set to be $\Lambda(t) = t$. We consider four scenarios for the values of regression parameters: (a) (β, γ) = (−0.5, 0.5); (b) (β, γ) = (−0.5, 0); (c) (β, γ) = (0, 0.5); and (d) (β, γ) = (0.5, 0.5). Under scenario (a), the short-term and long-term hazard ratios are on opposite directions; under scenario (b), the long-term hazard ratio is 1 corresponding to a true proportional odds model; under scenario (c), the short-term hazard ratio is 1; and under scenario (d), the short-term and long-term hazard ratios are equal corresponding to a true proportional hazards model. The censoring time is set to be the minimum of 2 and a uniform(0, 4) variable, producing approximately 29% censoring under all four scenarios. We used the quasi-Newton algorithm (Press et al., 1992) to calculate the NPMLEs. There is little difference between the standard error estimates through the Fisher information matrix and those from the profile likelihood approach. We present the standard error estimates based on the observed Fisher information matrix throughout the simulation studies and real data applications.

Table 1 summarizes the results for β, γ, and $\Lambda(t)$ with n = 100 and n = 200. For the nonparametric estimation of $\Lambda(t)$, we evaluated its estimates at t = 0.5 and t = 1.0. For comparison, we also fit the proportional hazards and proportional odds models, for which the regression parameters were denoted as β_{PH} and β_{PO}, respectively. The results in Table 1 indicate that the proposed method performs well for small sample sizes. In particular, the proposed estimators appear to be unbiased. The standard error estimator reflects accurately the true variation, and the confidence intervals have proper coverage probabilities. When the proportional hazards assumption is violated, the Cox model leads to biased estimates. Particularly, the results based on the Cox model can be very misleading when the short-term and long-term covariate effects are in opposite directions. Similar results were observed for the proportional odds model when the model assumption is not true. When the Cox model or the proportional odds model holds, as expected, the proposed NPMLEs are less efficient than those obtained under the true sub-model.

We also carried out simulation studies to incorporate a time-dependent covariate. Specifically, we generated failure times from the following model

$$\Lambda(t|\overline{\mathbf{X}}_i(t))=\int_0^t \frac{e^{(\beta+\gamma)X_i(s)}}{e^{\beta X_i(s)}F(s)+e^{\gamma X_i(s)}S(s)}d\Lambda(s),$$

where $X_i(t) = u_i(1+t)$ and u_i is a uniform(−1, 1) variable. We considered the same simulation setting as above regarding the parameters (β, γ, Λ). The summary statistics of the estimation of (β, γ, Λ) are provided in Table 2. The NPMLEs appear to continue to have good properties in the presence of time-dependent covariates.

Our next set of studies evaluated the proposed inference procedures for the testing of covariate effects and the assumptions of proportional hazards and proportional odds. Specifically, we considered Wald tests for the following null hypotheses: (H1) $H_0 : \beta = 0$; (H2) $H_0 : \gamma = 0$; (H3) $H_0 : \beta = \gamma = 0$; and (H4) $H_0 : \beta = \gamma$. Note that testing the long-term hazard ratio is equivalent to testing the proportional odds model. For comparison, we also considered the testing of covariate effects under the proportional hazards model: (H5) $H_0 : \beta_{PH} = 0$. We used the same simulation setting as above with $n = 200$. Table 3 presents the sizes/powers of the Wald tests at the nominal levels of 0.05. In all cases, the proposed tests have accurate control of type I error rates and reasonable powers under the alternative. The proposed tests of short-term, long-term and overall covariate effects tend to be more powerful than the Cox model when the proportional hazards assumption is violated. When our interest is to test the short-term or long-term hazard ratio only, the Cox model tends to yield inflated type I error rates under model mis-specifications.

We carried out additional simulation studies to compare the efficiency of the proposed NPMLEs relative to the pseudo-maximum likelihood estimators for two-sample data as implemented by Yang and Prentice (2005). We considered the same simulation settings as above except that X_i is a binary variable taking values −0.5 and 0.5 with equal probabilities. Table 4 presents the empirical mean squared errors for estimating β and γ based on 1,000 repetitions. As expected, under almost all situations the proposed estimators are more efficient than the pseudo-maximum likelihood estimators.

Finally, we conducted simulation studies to compare the new method with the time-varying coefficient Cox model (TimeCox) of Martinussen and Scheike (2006) and the Cox proportional hazards model incorporating time-dependent covariates (CoxPH$_t$), under both the true model and the mis-specfied model. The latter two models were fit by using R packages *timereg* (Martinussen and Scheike, 2006) and *survival*, respectively. We first generated data under the true model (6). To evaluate the performance of the new method under model mis-specification, we also generated data from the following model

$$\lambda(t|X_i)=\lambda(t)e^{\beta X_i+\gamma X_i t}. \quad (7)$$

In both cases, β and γ were set to be −0.5 and 0.5. The simple time-dependent covariate in the Cox proportional hazards model was defined as the interaction term between time and X_i. Table 5 displays the mean squared errors of three different estimators of the cumulative hazard function $\Lambda(t|x)$ with $n = 400$ based on 1,000 replicates. Under the correct model specification, as expected, the proposed estimators are more efficient than the other two estimators. When data are generated from model (7), the proposed estimators still perform reasonably well compared to the varying-coefficient Cox model and the Cox proportional hazards model with time-dependent covariates, both of which were fit under the correct model specification.

4. Real Data Examples

4.1 COGA study

In the COGA study mentioned previously, 643 individuals were affected with alcoholism and 971 individuals were disease-free at the time of interview. After excluding individuals with missing genotype at the target gene locus or phenotype data, the final data set for our analysis consisted of 1,371 individuals, including 626 affected individuals and 745 unaffected individuals.

Preliminary analysis revealed that gender was a risk factor for alcoholism; males were at a higher risk than females. Of the 626 affected individuals, 424 were males, as opposed to 229 males in the unaffected individuals. Previous linkage analysis showed a linked region on chromosome 14 (Palmer et al., 1999). Several recent studies on the genetic association analysis of ordinal traits (Wang et al., 2006; Diao and Lin, 2010) and age-at-onset data using cure models (Diao and Yin, 2012) suggested that SNP rs1972373 on chromosome 14 might be a disease susceptibility locus. Based on the Kaplan-Meier estimates of survival curves for the three genotype groups at SNP rs1972373 presented in Figure 2, allele ‘2’ appeared to have little short-term impact but strong long-term impact on the risk of alcoholism.

In our analysis, we fit the proposed model (4) and included gender and genotype score at SNP rs1972373 as covariates. The gender of an individual was coded as 1 for male and 0 for female, and the genotype score was coded as the numbers of allele type ‘2’. Both covariates were then centered at their means. The tests of the proportional hazards assumption for gender and genotype score at SNP rs1972373 were significant with p-values of 0.016 and 0.027. Gender appeared to have significant short-term and long-term effects on the age-at-onset of alcoholism. The short-term and long-term log-hazard ratios of male versus female are estimated at 0.866 and 1.9932 with standard error estimates of 0.147 and 0.367, both leading to *p*-values less than 0.0001. As expected, SNP rs1972373 appeared to have no short-term effect but significant long-term effect on the age-at-onset of alcoholism. The short-term log-hazard ratio of allele type ‘2’ versus allele type ‘1’ is estimated at −0.06 with a p-value of 0.479 whereas the long-term log-hazard ratio is estimated at 0.683 with a p-value of 0.015. One copy of allele type ‘2’ in the genotype at SNP rs1972373 is expected to increase the long-term hazard of alcoholism by 98% with a 95% confidence interval of (14%, 243%). Figure 1 plots the separate Kaplan-Meier and the model-fitted survival curves for each genotype group. The model-fitted survival function is calculated as the empirical average of the predicted survival functions. That the predicted survival functions agree well with the nonparametric Kaplan-Meier estimates of the survival curves indicates a good fit of the model. In contrast, the Cox model failed to detect the long-term effect of SNP rs1972373. The log-hazard ratio estimated from the Cox model is 0.083 with a standard error estimate of 0.058, corresponding to a p-value of 0.153.

We also fit the CoxPH_t model incorporating time by gender and time by gene interaction terms and the TimeCox model to the COGA data. The CoxPH_t model failed to detect the main genetic effect or the time by gene interaction effect with p-values of 0.279 and 0.094. The test of the proportional hazards assumption for gender using the CoxPH_t was also non-significant with a p-value of 0.509. The TimeCox detected patterns of the time-varying covariate effects similar to the proposed model, however, the Kolmogorov-Smirnov tests of the proportional hazards assumption were not significant for gender or the major gene with p-values 0.160 and 0.216, respectively, based on 10,000 resampling samples. To compare the fits of different models, we define six strata according to the different combinations of gender and genotype score and calculated the CMC. The proposed model fit the COGA very well with a CMC value of 0.388, compared to the CMC values of 1.515 and 0.777 for the

CoxPH$_t$ and the TimeCox model. The standard Cox proportional hazards model fit the data extremely poorly with a large CMC value of 19.482.

4.2 Gastrointestinal tumor study

As mentioned in the Introduction section, the gastrointestinal tumor study compared chemotherapy with the combined chemotherapy and radiotherapy on the treatment of locally unre-sectable gastric cancer. There were 45 patients randomly assigned to each treatment arm. Two observations were censored in the chemotherapy group and six were censored in the combined therapy group. Under the two-sample proportional hazards model, the log-hazard ratio of chemotherapy versus the combined therapy is estimated at 0.106 with a standard error estimate of 0.223, yielding a *p*-value of 0.635. The use of proportional hazards model failed to capture the phenomenon of crossing survival curves shown in Figure 1 and the results were meaningless in this situation.

We fit the proposed model (4) by letting $X_i = 0.5$ for the combined therapy group and $X_i = -0.5$ for the chemotherapy group. The test of the proportional hazards assumption is highly significant with a p-value of 6.0×10^{-4}. The new method successfully captured the phenomenon of crossing hazards. The short-term log-hazard ratio β and long-term log-hazard ratio γ are on opposite directions and estimated at 1.76 and −1.59 with standard error estimates of 0.582 and 0.509, leading to *p*-values of 0.0025 and 0.0018, respectively. The 95% confidence intervals are (0.62, 2.90) for β and (−2.59, −0.59) for γ. The estimated short-term and long-term hazard ratios are 5.81 and 0.20 with 95% confidence intervals (1.86, 18.17) and (0.075, 0.553). As evident in Figure 2, the model fitted survival curves agree well with the nonparametric Kaplan-Meier survival estimates very well indicating a good model fit. Our results are also consistent with the results from the two-sample model of Yang and Prentice (2005) using the pseudo-likelihood approach.

For comparison, we also fit the CoxPH$_t$ model incorporating the time by treatment interaction term and the TimeCox model. Both the CoxPH$_t$ and the TimeCox models detected the short-term and long-term treatment effects in the right direction. The CoxPH$_t$ detected significant time by treatment interaction effect with a p-value of 0.0035, whereas the Kolmogorov-Smirnov test of proportional hazards assumption in the TimeCox model was not significant with a p-value of 0.334 based on 10,000 resampling samples. To compare the fits of different models, we define two strata based on the treatment assignment and calculate the CMC defined in the Models and Inference Section. The proposed model fit the data the best with a CMC value of 0.037, compared to the CMC values of 0.141 and 0.106 for the CoxPH$_t$ and the TimeCox models, respectively. The standard Cox proportional model without time-dependent covariates fit the data very poorly, with a CMC value of 0.542.

5. Discussion

We have extended the two-sample semiparametric hazards rate model of Yang and Pren-tice (2005) to incorporate short-term and long-term effects of potentially time-dependent covariates. We have studied the nonparametric maximum likelihood estimation for the proposed model (4) and established the asymptotic properties for the NPMLEs. Unlike existing varying-coefficient Cox model, the estimation and inference procedures are likelihood-based and statistically efficient. Numerical studies and the applications to the Gastrointestinal tumor study and the COGA study demonstrate that the proposed inference procedures perform well in practical situations.

There are two inexplicit assumptions under the proposed model. First, the log-hazard ratios are assumed to be monotonic over time and it can be shown that the log-hazard ratios at any

time point are are between β and γ. The proposed method may not work well when the covariate effects are not monotonic. In such situations, a variety of alternative approaches described in the Introduction section can be explored. Secondly, under the proposed model, all subjects will eventually experience the event of interest given sufficiently long follow-up. This assumption, however, may not be true in several applications. In particular, a stable plateau at the tail of each curve in Figure 1 suggested that a fraction of individuals may never develop alcoholism, i.e., they were cured. The proposed model can evaluate the long-term hazard ratios between two sets of covariates but is not applicable to evaluating the covariate effects on the long-term survival probability in the presence of a cured subpopulation. To evaluate the covariate effects on the cured probability, cure rate models developed by Taylor (1995), Sy and Taylor (2000), Tsodikov (1998, 2002), Tsodikov et al. (2003), Mao and Wang (2010), and Diao and Yin (2012), among others, are recommended. To account for the presence of a cured subpopulation, we describe an extension of the proposed model such that population survival function takes the form

$$S_{pop}(t|\boldsymbol{Z},\overline{X}(t))=1-\pi(\boldsymbol{Z})+\pi(\boldsymbol{Z})S(t|\overline{X}(t)),$$

where $\mathbf{Z}$ is a set of covariates, $\pi(\mathbf{Z})$ is the uncured fraction, and $S(t|\bar{X}(t))$ is the survival function for the uncured subpopulation under the proposed model in this manuscript. The uncured fraction can be modelled through a logistic regression model

$$\pi(\boldsymbol{Z})=\frac{\exp(\boldsymbol{\xi}^T\boldsymbol{Z})}{1+\exp(\boldsymbol{\xi}^T\boldsymbol{Z})}.$$

Under this cure rate model, we allow for time-varying covariate effects in the uncured population. We are currently investigating the properties of this new cure rate model and comparing its performance with existing cure rate models.

We have implemented the new method in C language using the quasi-Newton algorithm described in Press et al. (1992). The convergence of the quasi-Newton algorithm is very fast and it takes less than 0.2 second to analyze one data set with 400 subjects on a Dell PowerEdge 2900 server. The efficiency of our computer program makes it feasible to apply our method to gene expression data and genome-wide association studies. Our user-friendly computer program is freely available on the website: http://mason.gmu.edu/~gdiao/software/. While the quasi-Newton algorithm works well for the proposed model, alternative algorithms, such as the EM-based algorithm (Tsodikov et al., 2003), the weighted Breslow-type algorithm (Chen, 2009), and the majorize-minimize algorithm (Mao and Wang, 2010), are available for nonparametric maximum likelihood estimation. It would be interesting to explore these algorithms in the proposed model.

For the purpose of illustration, we assume that observations in the COGA study are independent. Although the failure times within the same family tend to be correlated, the NPMLEs $\hat{\theta_n}$ can be shown to be consistent for θ and asymptotically normally distributed provided that the marginal model is corrected specified. However, the naive covariance matrix estimator for $\hat{\theta_n}$ using the inverse of the observed Fisher information matrix, is no longer valid in the presence of within-family dependence. To account for within-family correlations, one option is to fit marginal models and then use the robust sandwich estimators of covariance matrix. For the COGA data, the naive and robust covariance estimates were very close suggesting weak within-family correlations. Currently we are

investigating the extensions of the semiparametric hazards rate model (4) to correlated failure time data by using random effects.

We also note that the COGA study sampled families with alcoholism cases and involved retrospective ascertainment of family members of the probands. Consequently, the results from the proposed approach may only be applicable to a severely alcohol-dependent population and under the assumption that no latent dependence exists within families. More robust approaches to take into consideration the study design can be based on the conditional likelihood approach (de Andrade and Amos, 2000) or the retrospective likelihood approach (Li and Zhong, 2002; Zhong and Li, 2004). It would be interesting to develop conditional likelihood and/or retrospective likelihood-based approaches using the proposed model. Future research is warranted.

We have described a Cramér-von Mises type criterion to assess the goodness-of-fit of the proposed model. Alternatively, following the idea of Lin et al. (1993) for the Cox model, we can develop a goodness-of-fit procedure for assessing the adequacy of the proposed model (4) based on martingale residuals. Details are given in Web Appendix G. The theoretical justification of this procedure, however, is challenging since the partial likelihood function is not available under model (4). We are currently investigating this type of goodness-of-fit procedures for general semiparametric survival models including model (4).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the editor, an associate editor and a referee for helpful comments. The authors are grateful to the COGA investigators and Jean W. MacCluer for providing the COGA data from GAW14, which was supported in part by the NIH grant GM31575. The authors thank Dr. William Rosenberger for making valuable comments and suggestions, which lead to a considerable improvement in the presentation of this manuscript. The work of the first author was supported by the NIH grant R15CA150698.

References

Bennett S. Analysis of survival data by the proportional odds model. Statistics in Medicine. 1983; 2:273–277. [PubMed: 6648142]

Bickel, PJ.; Klaassen, CAJ.; Ritov, Y.; Wellner, JA. Efficient and Adaptive Estimation for Semiparametric Models. Vol. Ch 3. Johns Hopkins University Press; Baltimore: 1993.

Broyden CG. The convergence of a class of double rank minimization algorithms: 2: The new algorithm. IMA Journal of Applied Mathematics. 1970; 6:222–231.

Cai Z, Sun Y. Local linear estimation for time-dependent coefficients in coxs regression models. Scandinavian Journal of Statistics. 2003; 30:93–111.

Chen YH. Weighted breslow-type and maximum likelihood estimation in semi-parametric transformation models. Biometrika. 2009; 96:591–600.

Cox DR. Regression model and life-tables (with Discussion). Journal of the Royal Statistical Society, Series B. 1972; 34:187–220.

de Andrade M, Amos CI. Ascertainment issues in variance components models. Genetic epidemiology. 2000; 19:333–344. [PubMed: 11108643]

Diao G, Lin DY. Variance-componens methods for linkage and association analysis of ordinal traits in general pedigrees. Genetic Epidemiology. 2010; 34:232–237. [PubMed: 19918762]

Diao G, Yin G. A general transformation class of semiparametric cure rate frailty models. Annals of the Institute of Statistical Mathematics. 2012; 64:959–989.

Fletcher R. A new approach to variable metric algorithms. The Computer Journal. 1970; 13:317–322.

Gastrointestinal Tumor Study Group. A comparison of combination chemotherapy and combined modality therapy for locally advanced gastric carcinoma. Cancer. 1982; 49:1771–1777. [PubMed: 6176313]

Goldfarb D. A family of variable metric methods derived by variational means. Mathematics of Computation. 1970; 24:23–26.

Hasin D. Classification of alcohol use disorders. Alcohol Research & Health. 2003; 27:5–17. [PubMed: 15301396]

Hess KR. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. Statistics in Medicine. 1994; 13:1045–1062. [PubMed: 8073200]

Hougaard, P. Analysis of multivariate survival data. Springer Verlag; 2000.

Kosorok MR, Lee BL, Fine JP. Robust inference for univariate proportional hazards frailty regression models. The Annals of Statistics. 2004; 32:1448–1491.

Li H, Zhong X. Multivariate survival models induced by genetic frailties, with application to linkage analysis. Biostatistics. 2002; 3:57–75. [PubMed: 12933624]

Lin DY, Wei LJ, Ying Z. Checking the Cox model with cumulative sums of martingale-based residuals. Biometrika. 1993; 80:557–572.

Mao M, Wang JL. Semiparametric efficient estimation for a class of generalized proportional odds cure models. Journal of the American Statistical Association. 2010; 105:302–311. [PubMed: 22865944]

Martinussen, T.; Scheike, TH. Dynamic Regression Models for Survival Data. Springer; New York: 2006.

Martinussen T, Scheike TH, Skovgaard IM. Efficient estimation of fixed and time-varying covariate effects in multiplicative intensity models. Scandinavian Journal of Statistics. 2002; 29:57–74.

Murphy SA. Testing for a time dependent coefficient in coxs regression model. Scandinavian Journal of Statistics. 1993; 20:35–50.

Murphy SA, Rossini AJ, van der Vaart AW. Maximal likelihood estimation in the proportional odds model. Journal of the American Statistical Association. 1997; 92:968–976.

Murphy SA, Sen PK. Time-dependent coefficients in a cox-type regression model. Stochastic Processes and their Applications. 1991; 39:153–180.

Murphy SA, van der Vaart AW. On the profile likelihood. Journal of the American Statistical Association. 2000; 95:449–465.

Palmer LJ, Katrina JT, Burton PR. Genome-wide linkage analysis using genetic variance components of alcohol dependency-associated censored and continuous traits. Genetic Epidemiology. 1999; 17(Suppl 1):S283–S288. [PubMed: 10597450]

Peng L, Huang Y. Survival analysis with temporal covariate effects. Biometrika. 2007; 94:719–733.

Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical Recipes in C: The Art of Scientific Computing. 2. Cambridge University Press; Cambridge: 1992.

Shanno DF. Conditioning of quasi-newton methods for function minimization. Mathematics of Computation. 1970; 24:647–650.

Sy JP, Taylor JMG. Estimation in a cox proportional hazards cure model. Biometrics. 2000; 56:227–236. [PubMed: 10783800]

Taylor J. Semiparametric estimation in failure time mixture-models. Biometrics. 1995; 51:899–907. [PubMed: 7548707]

Therneau, TM.; Grambsch, PM. Modeling Survival Data: Extending the Cox Model. Springer; New York: 2000.

Tian L, Zucker D, Wei LJ. On the cox model with time-varying regression coefficients. Journal of the American Statistical Association. 2005; 100:172–183.

Tsodikov A. A proportional hazards model taking account of long-term survivors. Biometrics. 1998; 54:1508–1516. [PubMed: 9883549]

Tsodikov A. Semi-parametric models of long-and short-term survival: an application to the analysis of breast cancer survival in utah by age and stage. Statistics in medicine. 2002; 21:895–920. [PubMed: 11870824]

Tsodikov AD, Ibrahim JG, Yakovlev AY. Estimating cure rates from survival data: An alternative to two-component mixture models. Journal of the American Statistical Association. 2003; 98:1063–1078. [PubMed: 21151838]

Wang X, Ye Y, Zhang H. Family-based association test for ordinal traits adjusting for covariates. Genetic Epidemiology. 2006; 30:728–736. [PubMed: 17086513]

Winnett A, Sasieni P. Iterated residuals and time-varying covariate effect in cox regression. Journal of the Royal Statistical Society: Series B. 2003; 65:473–488.

Yang S, Prentice R. Semiparametric analysis of short-term and long-term hazard ratios with two-sample survival data. Biometrika. 2005; 92:1–17.

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

Zhong X, Li H. Score tests of genetic association in the presence of linkage based on the additive genetic gamma frailty model. Biostatistics. 2004; 5:307–327. [PubMed: 15054033]

Zucker DM, Karr AF. Nonparametric survival analysis with timedependent covariate effects: A penalized partial likelihood approach. Annals of Statistics. 1990; 18:329–353.

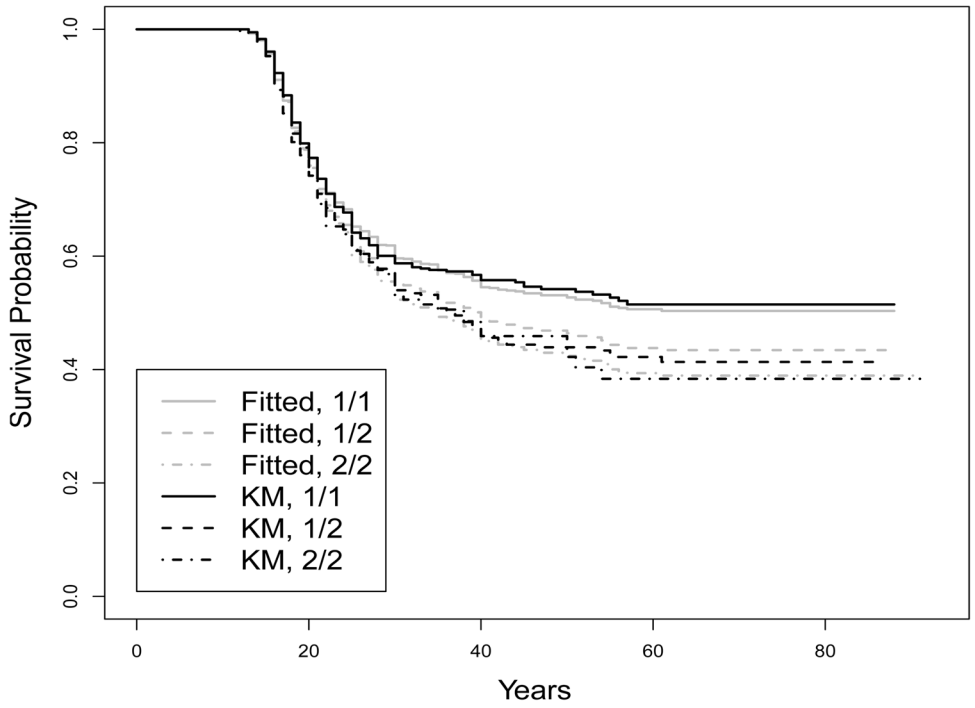

Figure 1.
Kaplan-Meier and model-fitted survival curves from the COGA study.

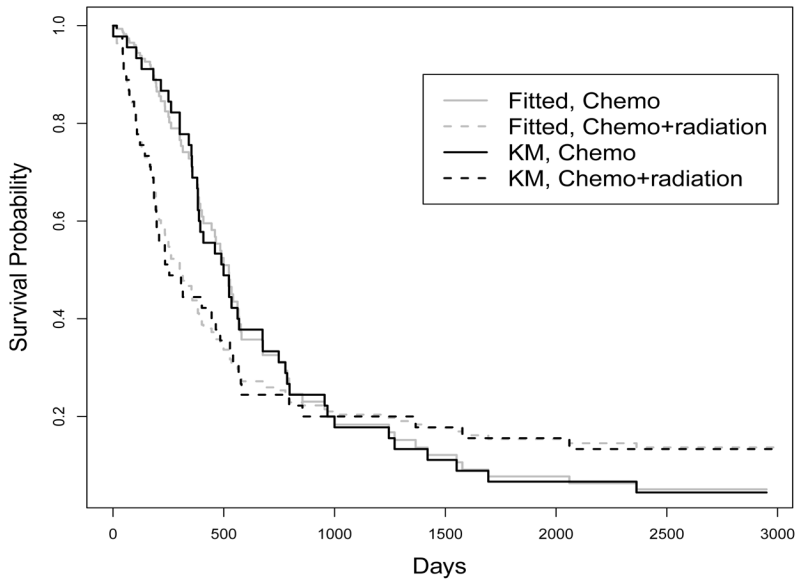

Figure 2.
Kaplan-Meier and model-fitted survival curves from the Gastrointestinal tumor study.

Table 1

Summary statistics for the simulation studies based on 1,000 replications

n	Par	Est	SE	SEE	CP	Est	SE	SEE	CP
		(β, γ) = (−0.5, 0.5)				(β, γ) = (−0.5, 0.0)			
100	β	−0.511	0.413	0.412	0.956	−0.506	0.407	0.409	0.958
	γ	0.465	0.570	0.556	0.938	−0.04	0.564	0.568	0.944
	Λ (0.5)	0.508	0.086	0.085	0.946	0.507	0.085	0.086	0.962
	Λ (1.0)	1.019	0.145	0.146	0.954	1.019	0.146	0.149	0.959
	β_{PH}	−0.116	0.217	0.209	-	−0.317	0.219	0.211	0.926
	β_{PO}	−0.294	0.327	0.320	-	−0.511	0.326	0.322	0.944
200	β	−0.512	0.291	0.288	0.954	−0.507	0.287	0.286	0.955
	γ	0.496	0.400	0.389	0.940	−0.007	0.401	0.399	0.950
	Λ (0.5)	0.504	0.059	0.059	0.954	0.504	0.058	0.060	0.953
	Λ (1.0)	1.012	0.104	0.101	0.947	1.012	0.104	0.104	0.953
	β_{PH}	−0.107	0.153	0.147	-	−0.308	0.154	0.148	-
	β_{PO}	−0.287	0.231	0.225	-	−0.504	0.231	0.226	0.948
		(β, γ) = (0.0, 0.5)				(β, γ) = (0.5, 0.5)			
100	β	−0.012	0.406	0.405	0.954	0.495	0.414	0.409	0.945
	γ	0.490	0.570	0.563	0.934	0.512	0.587	0.585	0.947
	Λ (0.5)	0.510	0.087	0.085	0.952	0.509	0.087	0.087	0.954
	Λ (1.0)	1.023	0.146	0.148	0.958	1.027	0.147	0.151	0.959
	β_{PH}	0.188	0.211	0.210	-	0.499	0.216	0.214	0.952
	β_{PO}	0.202	0.321	0.319	-	0.707	0.327	0.325	-
200	β	−0.009	0.284	0.282	0.956	0.496	0.287	0.285	0.962
	γ	0.501	0.398	0.395	0.947	0.503	0.410	0.411	0.944
	Λ (0.5)	0.506	0.059	0.060	0.952	0.505	0.059	0.061	0.957
	Λ (1.0)	1.014	0.104	0.102	0.944	1.015	0.104	0.105	0.957
	β_{PH}	0.193	0.149	0.147	-	0.498	0.151	0.150	0.946
	β_{PO}	0.207	0.227	0.225	-	0.706	0.228	0.229	-

Par, the parameter to be estimated; Est, the average estimate; SE, the sample standard deviation of the estimates; SEE, the average standard error; CP, the coverage probability of the nominal 95% confidence intervals.

Table 2

Summary statistics for the simulation studies with time-dependent covariate based on 1,000 replications with $n = 100$

Par	Est	SE	SEE	CP	Est	SE	SEE	CP
	$(\beta, \gamma) = (-0.5, 0.5)$				$(\beta, \gamma) = (-0.5, 0.0)$			
β	−0.487	0.254	0.261	0.953	−0.482	0.270	0.265	0.947
γ	0.505	0.180	0.175	0.946	−0.016	0.164	0.160	0.952
Λ(0.5)	0.506	0.082	0.101	0.973	0.506	0.079	0.096	0.975
Λ(1.0)	1.009	0.141	0.174	0.975	1.012	0.136	0.162	0.973
	$(\beta, \gamma) = (0.0, 0.5)$				$(\beta, \gamma) = (0.5, 0.5)$			
β	0.002	0.270	0.273	0.953	0.501	0.294	0.286	0.941
γ	0.520	0.180	0.182	0.961	0.523	0.206	0.206	0.958
Λ(0.5)	0.507	0.080	0.098	0.977	0.509	0.082	0.097	0.972
Λ(1.0)	1.012	0.14	0.167	0.975	1.013	0.141	0.164	0.967

Par, the parameter to be estimated; Est, the average estimate; SE, the sample standard deviation of the estimates; SEE, the average standard error; CP, the coverage probability of the nominal 95% confidence intervals. True values of Λ(0.5) and Λ(1.0) are 0.5 and 1.0, respectively.

Table 3

Empirical size/power of the Wald test at significance level of 0.05 based on 1,000 replications

β	γ	H1	H2	H3	H4	H5
0.0	0.0	0.040	0.052	0.050	0.050	0.059
−0.5	−0.5	0.434	0.246	0.860	0.052	0.917
−0.5	−0.4	0.439	0.184	0.801	0.053	0.874
−0.5	−0.3	0.428	0.142	0.723	0.059	0.815
−0.5	−0.2	0.429	0.086	0.638	0.071	0.741
−0.5	−0.1	0.437	0.051	0.563	0.096	0.656
−0.5	0.0	0.438	0.050	0.499	0.137	0.544
−0.5	0.1	0.431	0.062	0.440	0.166	0.447
−0.5	0.2	0.437	0.089	0.396	0.220	0.345
−0.5	0.3	0.432	0.137	0.372	0.268	0.254
−0.5	0.4	0.428	0.189	0.363	0.328	0.179
−0.5	0.5	0.433	0.262	0.362	0.398	0.129
−0.4	0.5	0.304	0.258	0.287	0.341	0.074
−0.3	0.5	0.195	0.253	0.217	0.264	0.063
−0.2	0.5	0.104	0.266	0.191	0.228	0.075
−0.1	0.5	0.055	0.266	0.212	0.176	0.154
0.0	0.5	0.044	0.263	0.272	0.139	0.265
0.1	0.5	0.056	0.261	0.359	0.109	0.400
0.2	0.5	0.099	0.245	0.476	0.086	0.563
0.3	0.5	0.174	0.248	0.632	0.065	0.718
0.4	0.5	0.308	0.231	0.742	0.057	0.840
0.5	0.5	0.417	0.223	0.851	0.047	0.911

Table 4

Mean squared errors of the proposed NPMLEs and the pseudo-maximum likelihood estimators (PMLEs) of Yang and Prentice (2005) for (β, γ)

		PMLE		NPMLE		PMLE/NPMLE	
n	(β,γ)	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\beta}$	$\hat{\gamma}$
100	(−0.5,0.5)	0.090	0.108	0.073	0.111	1.242	0.978
	(−0.5,0.0)	0.085	0.114	0.061	0.105	1.390	1.084
	(0.0,0.5)	0.069	0.107	0.063	0.110	1.101	0.967
	(0.5,0.5)	0.088	0.144	0.067	0.133	1.314	1.087
200	(−0.5,0.5)	0.048	0.060	0.036	0.054	1.360	1.107
	(−0.5,0.0)	0.041	0.061	0.031	0.0543	1.310	1.119
	(0.0,0.5)	0.030	0.050	0.030	0.0516	1.025	0.974
	(0.5,0.5)	0.035	0.064	0.030	0.0598	1.152	1.068

Table 5

Comparison of mean squared errors of estimators of $\Lambda(t|x)$

			CoxPH_t		TimeCox	
x	***t***	**New MSE**	**MSE**	**RE**	**MSE**	**RE**
Under correct model specification						
−1	0.5	0.0064	0.0089	1.391	0.0173	2.688
	1.0	0.0193	0.0290	1.502	0.0374	1.933
	1.5	0.0448	0.0521	1.163	0.0635	1.417
1	0.5	0.0041	0.0041	1.002	0.0050	1.223
	1.0	0.0131	0.0133	1.016	0.0182	1.387
	1.5	0.0327	0.0456	1.396	0.0500	1.529
Under model mis-specification						
−1	0.5	0.0096	0.0098	1.021	0.0173	1.796
	1.0	0.0315	0.0217	0.691	0.0297	0.944
	1.5	0.0602	0.0408	0.678	0.0500	0.830
1	0.5	0.0037	0.0033	0.891	0.0041	1.091
	1.0	0.0121	0.0115	0.950	0.0119	0.984
	1.5	0.0334	0.0377	1.128	0.0372	1.114

Note: CoxPH_t and TimeCox represent the Cox proportional hazards model (7) incorporating time-dependent covariates and the varying coefficient Cox model of Martinussen and Scheike (2006) respectively; RE represents the relative efficiency of the proposed estimator with respect to the estimators under the CoxPH_t and the TimeCox models.