Published in final edited form as:
Scand Stat Theory Appl. 2016 March ; 43(1): 103–122. doi:10.1111/sjos.12167.

A Proportional Hazards Regression Model for the Subdistribution with Covariates Adjusted Censoring Weight for Competing Risks Data

PENG HE
Division of Biostatistics, Medical College of Wisconsin, U.S.A.

FRANK ERIKSSON and **THOMAS H. SCHEIKE**
Department of Biostatistics, University of Copenhagen, Denmark

MEI-JIE ZHANG
Division of Biostatistics, Medical College of Wisconsin, U.S.A.

Abstract

With competing risks data, one often needs to assess the treatment and covariate effects on the cumulative incidence function. Fine and Gray proposed a proportional hazards regression model for the subdistribution of a competing risk with the assumption that the censoring distribution and the covariates are independent. Covariate-dependent censoring sometimes occurs in medical studies. In this paper, we study the proportional hazards regression model for the subdistribution of a competing risk with proper adjustments for covariate-dependent censoring. We consider a covariate-adjusted weight function by fitting the Cox model for the censoring distribution and using the predictive probability for each individual. Our simulation study shows that the covariate-adjusted weight estimator is basically unbiased when the censoring time depends on the covariates, and the covariate-adjusted weight approach works well for the variance estimator as well. We illustrate our methods with bone marrow transplant data from the Center for International Blood and Marrow Transplant Research (CIBMTR). Here cancer relapse and death in complete remission are two competing risks.

Keywords

competing risks; cumulative incidence function; proportional hazards model; subdistribution; inverse probability of censoring weight

1 Introduction

Biomedical research often involves competing risks in which each subject is at risk of failure from K different causes. For competing risks data, one only observes the first event to occur and this precludes the occurrence of another event. Also, one often wishes to estimate and model the cumulative incidence function (CIF), which is the marginal probability of

Mei-Jie Zhang, Division of Biostatistics, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226, U.S.A. meijie@mcw.edu.

failure of a specific cause. The standard approach of modeling CIF is to model the cause-specific hazard functions for all causes. Let $\lambda_k(t; \boldsymbol{Z})$ be the kth conditional cause-specific hazard (k = 1, 2 for simplicity), where $\boldsymbol{Z}$ is given set of covariates. The CIF of cause 1 given by $\boldsymbol{Z}$ is

$$F_1(t;Z) = P\left(\tilde{T} \le t, \epsilon = 1 | Z\right) = \int_0^t \lambda_1(s;Z)\ exp\left[-\int_0^{s^-}\{\lambda_1(u;Z) + \lambda_2(u;Z)\}\ du\right] ds,$$

where $\tilde{T}$ is the failure time and $\in$ indicates the type of failure. Here, all cause-specific hazards need to be modeled adequately and correctly. Note that the cumulative incidence function $F_1(t; \boldsymbol{Z})$ is a subdistribution function since $F_1(\infty; \boldsymbol{Z}) < 1$. Prentice et al. (1978) and Cheng et al. (1998) proposed using Cox (1972) proportional hazards model for all causes. Alternatively, Shen and Cheng (1999) considered a special additive model, and Scheike and Zhang (2002, 2003) proposed and studied a flexible Cox-Aalen model, which allows some of the covariates to have time-varying effects. Since the cumulative incidence function of a specific cause is a function of cause-specific hazards for all causes, it is difficult to summarize the covariate effect (Zhang and Fine, 2008) and to identify the covariate effect on the cumulative incidence function. However, regression methods have been developed to directly model the cumulative incidence function. Fine and Gray (1999) (FG) developed a regression method to directly model the CIF by modeling the subdistribution hazard function through a Cox type regression model, $\lambda_k^*(t, \boldsymbol{Z}) = -d\,log\{1 - F_k(t;\boldsymbol{Z})\}/dt = \lambda_{k0}^*(t)\ exp(\boldsymbol{\beta}_k^\tau \boldsymbol{Z})$ based on early work by Gray (1988) and Pepe (1991). FG proposed using an inverse probability of the censoring weighting (IPCW) technique to estimate the regression parameter β_k and cumulative baseline subdistribution hazard function $\Lambda_{k0}^*(t) = \int_0^t \lambda_{k0}^*(s)\,ds$. This approach has been implemented in an R-package, cmprsk. FG's model has been considered and used extensively in cancer studies, epidemiological studies, and many other biomedical studies (Scrucca et al., 2007; Wolbers et al., 2009; Kim, 2007; Lau et al., 2009). Let $r(t) = I\left\{C \ge \left(\tilde{T} \wedge t\right)\right\}$ and $G_C(t) = P(C > t)$, where C is the censoring time. Fine and Gray's approach is based on the fact that $E\left[r(t)/G_C\left\{\left(\tilde{T} \wedge C\right) \wedge t\right\}|Data\right] = 1$ provided that censoring time is independent of the covariates, and FG proposed using the Kaplan-Meier estimator to estimate the unknown censoring distribution G_C. However, in biomedical research studies, the censoring time may depend on some of the covariates and the treatment group. In a clinical trial, patients may be more likely to drop out with some specific value of covariate characteristics, and one treatment group may have a higher dropout rate than the others (Mai, 2008). DiRienzo and Lagakos (2001a,b) showed that when the distribution of censoring depends on both the treatment group and the covariates, in general the null asymptotic distribution of the score test is not centered at zero when the model is misspecified, the tests of treatment group effect can be severely biased. Heinze et al. (2003) showed that if the censoring distributions are not similar in the two comparison groups, the log-rank test and fitting a regression model, such as fitting a proportional hazards model, may not be valid. For the competing risks data, one can show that $E\left[r(t)/G_C\left\{\left(\tilde{T} \wedge C\right) \wedge t|Data\right\}|Data\right] = 1$, where $G_C\left\{\left(\tilde{T} \wedge C\right) \wedge t|Data\right\}$ is the conditional censoring distribution given by *Data*. Thus, parameter estimates using the inverse probability of censoring weighting approach with the Kaplan-Meier estimator may be biased when the censoring distribution depends on some of the covariates. To adjust the IPCW when censoring distribution depends on some of the covariates, Fine and Gray (1999) suggested using a stratified Kaplan-Meier estimator for the discrete covariates and assuming the Cox model for the continuous covariates. In this study, we considered a regression model for the censoring distribution, such as a Cox proportional hazards model, and using the predicted censoring probability for each individual subject for the weight function. With the Cox model adjusted weight, we derived an efficient variance estimator which includes variation contributed from estimated censoring distribution, and we performed a simulation study to examine the bias that would arise without adjusting covariates for estimating the censoring distribution, potential bias reduction and robustness of using the Cox model for the censoring distribution. Furthermore, Fine and Gray proposed using a stabilized factor $\hat{G}_{\mathrm{C}}(t)$ with inverse weight $r(t)\hat{G}_{\mathrm{C}}(t)/\hat{G}_{\mathrm{C}}\left\{\left(\widetilde{T}\wedge C\right)\wedge t\right\}$. Our simulation indicates that this stabilized weight improves the efficiency and reduces the bias, but not enough. With the Cox model adjusted weight function, we also considered using a stabilized weight $r(t)\hat{G}_{\mathrm{C}}(t|\boldsymbol{X}=x)/\hat{G}_{\mathrm{C}}\left\{\left(\widetilde{T}\wedge C\right)\wedge t|\boldsymbol{X}=x\right\}$ to improve efficiency and to reduce bias, where $\boldsymbol{X}$ is the covariates, which is associated with the censoring distribution and could be a subset covariates of $\boldsymbol{Z}$.

The outline of the remainder of the paper is as follows. In Section 2 we describe the competing risks data structure. We introduce a regression-adjusted inverse weighted estimation for the proportional subdistribution hazards model and present the asymptotic results that can be used for inference. Simulation studies are provided in Section 3. In Section 4 we analyze two real data sets, which were originally studied by Kumar et al. (2012) and by Ringdén et al. (2012) using data from the Center for International Blood and Marrow Transplant Research (CIBMTR). Concluding remarks are provided in Section 5. The proof of the main asymptotic result and the simulation procedure are given in Appendix A and B, respectively.

2 Data and covariate adjusted censoring weight

Let $\widetilde{T}_i$ and C_i be the event time and right censoring time for ith individual, respectively. $\epsilon_i \in \{1,\ldots,K\}$ indicates the cause of failure. For simplicity, we assume $K = 2$ in this study. Let $T_i = min\left(\widetilde{T}_i, C_i\right)$ and $\Delta_i = \mathscr{I}\left(\widetilde{T}_i \le C_i\right)$. We observe n independent and identically distributed (*i.i.d.*) data $\{T_i, \Delta_i, \Delta_i\epsilon_i, \boldsymbol{Z}_i\}$ for $i = 1, \ldots, n$, where $\boldsymbol{Z}_i = (Z_{i1},\ldots, Z_{iq})^{\mathrm{T}}$ are associated covariates. We assume that $(\widetilde{T}_i, \epsilon_i)$ are independent of C_i given covariates of $\boldsymbol{Z}_i$. We are interested in modeling the cumulative incidence function of cause 1, $F_1(t; \boldsymbol{Z})$. Based on Gray (1988) subdistribution hazard technique, Fine and Gray (1999) proposed a proportional subdistribution hazards model

$$\lambda_1^*(t;Z) = \frac{-dlog\{1-F_1(t;Z)\}}{dt} = \lambda_{10}^*(t)\ exp\ \{\beta_0^{\mathrm{T}} Z\}. \quad (2.1)$$

There is a direct relationship between the CIF and subdistribution hazard function:

$$F_1(t;Z) = 1 - exp\left\{-\left(\int_0^t \lambda_{10}^*(u)\,du\right) e^{\beta_0^\intercal \boldsymbol{Z}}\right\}.$$

Let $N_i^1(t) = I\left(\widetilde{T}_i \le t, \epsilon_i = 1\right)$ be the underlying counting process associated with cause 1. For right censored competing risks data, $N_i^1(t)$ and $Y_i^1(t) = 1 - N_i^1(t^-)$ are not fully observed. For a censored individual, they are only observed up to the censoring time C_i. Define $r_i(t) = I\left\{C_i \ge \left(\widetilde{T}_i \wedge t\right)\right\}$. Then $r_i(t)\,N_i^1(t)$ and $r_i(t)\,Y_i^1(t)$ are computable for all times t. Let $G_C(t; \boldsymbol{Z}) = P(C \ge t|\boldsymbol{Z})$ be the conditional censoring distribution. Based on the assumption that given covariates $\boldsymbol{Z}$ the event time and censoring time are independent and the models are formulated as standard regression models conditional on $\boldsymbol{Z}$, it then follows that given $\boldsymbol{Z}$

$$\begin{aligned} E\left\{\frac{r(t)N^1(t)}{G_C(T\wedge t;\boldsymbol{Z})}\middle|\boldsymbol{Z}\right\} &= E\left[E\left\{\frac{r(t)N^1(t)}{G_C(T\wedge t;\boldsymbol{Z})}\middle|T,\epsilon,\boldsymbol{Z}\right\}\middle|\boldsymbol{Z}\right] \\ &= E\left[E\left\{N^1(t)\,|T,\epsilon,\boldsymbol{Z}\right\}\frac{E\{r(t)|T,\epsilon,\boldsymbol{Z}\}}{G_C(T\wedge t;\boldsymbol{Z})}\middle|\boldsymbol{Z}\right] \\ &= F_1(t;\boldsymbol{Z}). \end{aligned}$$

FG proposed using an inverse probability of the censoring weighting (IPCW) approach to fit the model (2.1) and proposed an IPCW weight function $\hat{w}_i^{KM}(t) = r_i(t)\,\hat{G}_C^{KM}(t)/\hat{G}_C^{KM}(T_i \wedge t)$, where $\hat{G}_C^{KM}(t)$ is the Kaplan-Meier estimator for the unknown censoring distribution. FG proposed estimating the unknown regression coefficient β by solving the score equation

$$\boldsymbol{U}_{KM}(\beta) = \sum_i \int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\Sigma_j w_j^{KM}(u)\,Y_j^1(u)\,\boldsymbol{Z}_j\,exp\,\{\beta^\intercal \boldsymbol{Z}_j\}}{\Sigma_j w_j^{KM}(u)\,Y_j^1(u)\;exp\,\{\beta^\intercal \boldsymbol{Z}_j\}}\right\} w_i^{KM}(u)\,dN_i^1(u) = 0,$$

where τ is is end of the study time point, and denote the estimate as $\hat{\beta}_{KM}$. FG showed that under regularity conditions and the condition that the censoring distribution is independent of covariates, $\hat{\beta}_{KM}$ is consistent for β_0 and derived large sample properties for $\sqrt{n}\left(\hat{\beta}_{KM} - \beta_0\right)$ and $\sqrt{n}\left\{\hat{\Lambda}_{10}^{KM}(t) - \Lambda_{10}^*(t)\right\}$, where the cumulative baseline subdistribution hazard $\Lambda_{10}^*(t) = \int_0^t \lambda_{10}^*(u)$ is estimated by

$$\hat{\Lambda}_{10}^{KM}(t) = \sum_i \int_0^t \frac{w_i^{KM}(u)\,dN_i^1(u)}{\Sigma_j w_j^{KM}(u)\,Y_j^1(u)\;exp\left\{\hat{\beta}_{KM}^\intercal \boldsymbol{Z}_j\right\}}.$$

It has been shown that in biomedical research studies the censoring time may depend on some of the covariates and the treatment group. To make asymptotically unbiased inference, we needed to model the censoring distribution and to estimate the censoring survival probability, $G_C(T \wedge t; \boldsymbol{Z}_i)$, for each individual. In this study, as suggested by Fine and Gray (1999), we considered the commonly used Cox proportional hazards model for the censoring distribution,

$$\lambda_C(t;\boldsymbol{X}_i) = \lambda_{C0}(t)\ exp\ \{\gamma_0^\intercal \boldsymbol{X}_i\},$$

where $\boldsymbol{X}_i$ is the covariates associated with the censoring distribution and can be a function or subset of $\boldsymbol{Z}_i$. In practice one can use the standard model checking procedure to check the Cox model assumption for the censoring distribution and use the standard model building procedure to identify the risk factors which are associated with the censoring time. Let x_i be the fixed observed value for the ith individual's covariates, we estimate the predicted censoring survival probability $G_{COX}^{C}(t;\boldsymbol{x}_i) = P(C>t|\boldsymbol{X}_i = \boldsymbol{x}_i) = exp\ \{-\Lambda_{C0}(t)\ exp\ (\gamma_0^\intercal \boldsymbol{x}_i)\}$ by

$$\hat{G}_C^{COX}(t;\boldsymbol{x}_i) = exp\ \left\{-\hat{\Lambda}_{C0}(t)\ exp\ (\hat{\gamma}^\intercal \boldsymbol{x}_i)\right\}, \quad (2.2)$$

where $\hat{\gamma}$ is a maximum partial likelihood estimate for γ_0 and $\hat{\Lambda}_{C0}(t)$ is the standard Breslow estimator for the cumulative baseline censoring hazard $\Lambda_{C0}(t) = \int_0^t \lambda_{C0}(u)\,du$. Note that, any administrative censoring events at time τ are not considered as events in the regression estimation. In this study, we considered a covariate-adjusted IPCW weight function

$$\hat{w}_i^{COX}(t) = r_i(t)\,\hat{G}_C^{COX}(t;\boldsymbol{x}_i)\,/\,\hat{G}_C^{COX}(T_i \wedge t;\boldsymbol{x}_i).$$

We estimated β in model (2.1) by solving the score equation

$$\boldsymbol{U}_{COX}(\beta) = \sum_i \int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\Sigma_j \hat{w}_j^{COX}(u)\,Y_j^1(u)\,\boldsymbol{Z}_j\ exp\ \{\beta^\intercal \boldsymbol{Z}_j\}}{\Sigma_j \hat{w}_j^{COX}(u)\,Y_j^1(u)\ exp\ \{\beta^\intercal \boldsymbol{Z}_j\}}\right\} \hat{w}_i^{COX}(u)\,dN_i^1(u) = 0, \quad (2.3)$$

and denoted the estimate as $\hat{\beta}_{COX}$. Then we estimated $\Lambda_{10}^*(t)$ by

$$\hat{\Lambda}_{10}^{COX}(t) = \sum_i \int_0^t \frac{\hat{w}_i^{COX}(u)\,dN_i^1(u)}{\Sigma_j \hat{w}_j^{COX}(u)\,Y_j^1(u)\ exp\ \left\{\hat{\beta}_{COX}^\intercal \boldsymbol{Z}_j\right\}}.$$

Under regularity conditions (given in Section 6.1)), it can be shown that (see (6.6))

$$\sqrt{n}\left(\hat{\beta}_{COX} - \beta_0\right) = \Omega^{-1}\left\{\frac{1}{\sqrt{n}}\boldsymbol{U}_{COX}(\beta_0)\right\} + o_P(1),$$

where $\Omega = \lim_{n\to\infty} n^{-1}\boldsymbol{I}_{COX}(\beta_0)$, $\boldsymbol{I}_{COX}(\beta) = -\partial\{\boldsymbol{U}_{COX}(\beta)\}/\partial\beta$, and Ω can be estimated by $\hat{\boldsymbol{\Omega}} = n^{-1}\boldsymbol{I}_{COX}\left(\hat{\beta}\right)$ Furthermore (see(6.4)),
$n^{-1/2}\boldsymbol{U}_{COX}(\boldsymbol{\beta}_0) = n^{-1/2}\sum_i \left(\boldsymbol{\xi}_i^{COX} + \boldsymbol{\psi}_i^{COX}\right) + o_P(1)$, where explicit expressions for $\boldsymbol{\xi}_i^{COX}$

and $\boldsymbol{\psi}_i^{\text{COX}}$ are given in the Appendix A. The quantities can be estimated by plug-in estimators denoted by $\boldsymbol{\xi}_i^{\text{COX}}$ and $\boldsymbol{\psi}_i^{\text{COX}}$, respectively. It follows that $\sqrt{n}\left(\hat{\boldsymbol{\beta}}_{\text{COX}}-\boldsymbol{\beta}_0\right)$ converges in distribution to a mean zero Gaussian distribution with an asymptotic variance that can be estimated by

$$\hat{\Sigma}_{\beta}^{\text{COX}}=n\sum_i\left(\hat{W}_{\beta,i}^{\text{COX}}\right)^{\otimes 2}=n\left\{\boldsymbol{I}_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right)\right\}^{-1}\left\{\sum_i\left(\hat{\xi}_i^{\text{COX}}+\hat{\psi}_i^{\text{COX}}\right)^{\otimes 2}\right\}\left\{\boldsymbol{I}_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right)\right\}^{-1},$$

where $\boldsymbol{a}^{\otimes 2}=\boldsymbol{a}\boldsymbol{a}^{\intercal}$ for a column vector $\boldsymbol{a}$.

Similarly (see (6.10)), $\sqrt{n}\left\{\hat{\Lambda}_{10}^{\text{COX}}(t)-\Lambda_{10}^{*}(t)\right\}=n^{-1/2}\sum_i W_{\Lambda,i}^{\text{COX}}(t)+o_P(1)$, which converges weakly to a mean zero Gaussian process with asymptotic variances, which can be estimated by

$$\hat{\Sigma}_{\Lambda_{10}}^{\text{COX}}(t)=\frac{1}{n}\sum_i\left\{\hat{W}_{\Lambda,i}^{\text{COX}}(t)\right\}^2.$$

Explicit expressions for $W_{\Lambda,i}^{\text{COX}}(t)$ and $\hat{W}_{\Lambda,i}^{\text{COX}}(t)$ can be found in the Appendix A.

For a given set value of covariates, $\boldsymbol{z}_0$, the predicted CIF of cause 1 can be estimated by $\hat{F}_1^{KM}(t;\boldsymbol{z}_0)=1-exp\left\{-\hat{\Lambda}_{10}^{KM}(t)\ exp\left(\hat{\boldsymbol{\beta}}_{KM}^{\intercal}\boldsymbol{z}_0\right)\right\}$ or $\hat{F}_1^{\text{COX}}(t;\boldsymbol{z}_0)=1-exp\left\{-\hat{\Lambda}_{10}^{\text{COX}}(t)\ exp\left(\hat{\boldsymbol{\beta}}_{\text{COX}}^{\intercal}\boldsymbol{z}_0\right)\right\}$ respectively. Fine and Gray (1999) derived the large sample property for $\sqrt{n}\left\{\hat{F}_1^{KM}(t;\boldsymbol{z}_0)-F_1(t;\boldsymbol{z}_0)\right\}$ when the censoring distribution is independent of the covariates. When the censoring distribution depends on the covariates through a Cox model, by the functional Delta method it follows that $\sqrt{n}\left\{\hat{F}_1^{\text{COX}}(t;\boldsymbol{z}_0)-F_1(t;\boldsymbol{z}_0)\right\}$ converges in distribution to a Gaussian process with mean zero and asymptotic variances, which can be estimated by

$$n\left\{1-\hat{F}_1^{\text{COX}}(t;z_0)\right\}^2\sum_i\left\{\hat{W}_{F_1,i}^{\text{COX}}(t;z_0)\right\}^2,$$

where

$$\hat{W}_{F_{1,i}}^{\text{COX}}(t;z_0)=exp\left(\hat{\beta}_{\text{COX}}^{\intercal}z_0\right)\left\{\hat{\Lambda}_{10}^{\text{COX}}(t)\left(\hat{W}_{\beta,i}^{\text{COX}}\right)^{\intercal}z_0+\hat{W}_{\Lambda,i}^{\text{COX}}(t)\right\}.$$

Resampling techniques can be used to construct confidence bands for $\Lambda_{10}^{*}(t)$ and $F_1(t;\boldsymbol{z}_0)$ (Lin et al., 1994; Scheike et al., 2008).

3 Simulations

We compared the finite-sample performance of the estimator using the covariate-adjusted censoring weight to the unadjusted estimator using the Kaplan-Meier estimator for the censoring distribution. Two simulation studies were considered to examine the potential bias reduction with the covariate-adjusted censoring weight estimator. For the first study, we had one binary covariate. For the second study, we considered one binary covariate and one continuous covariate. In both studies, we compared the performance of estimators using two weights, $\hat{w}_i^{KM}(t)$ and $\hat{w}_i^{\text{COX}}(t)$, respectively.

3.1 Study 1

The regression model below has one binary covariate *Z*. Given *Z*, the cumulative incidence functions are given by

$$F_1(t;Z) = 1 - \left\{1 - p\left(1 - e^{-t}\right)\right\}^{exp(\beta Z)}$$

and

$$F_2(t;Z) = (1-p)^{exp(\beta Z)}\left\{1 - e^{-texp(\beta Z)}\right\},$$

where $p = F_1(\infty|Z = 0)$. We let $p = 0.66$ and *Z* be a Bernoulli random variable, with value 1 for half of the sample and 0 for the other half. For each setting, we simulated 10,000 replicates with sample size of $n = 100$ and 300, respectively (detailed simulating procedure is given in Appendix B). We set $\beta = 1$ and considered the following three simulation scenarios.

Scenario 1	Censoring times are independent of *Z*:
	Generate censoring times from an exponential distribution ~ $\exp(\lambda_C)$
	Set $\lambda_C = 0.556$ for 30% censoring, $\lambda_C = 1.342$ for 50% censoring
Scenario 2	Censoring times depend on *Z* by a Cox model:
	Generate censoring times from a Cox model, $\lambda_C(t|Z) = \lambda_C \exp(\beta_C Z)$
	Set $\beta_C = 2.5$ and $\lambda_C = 0.137$ for 30% censoring
	Set $\beta_C = 2.5$ and $\lambda_C = 0.391$ for 50% censoring
Scenario 3	Censoring times depend on *Z*, not by a Cox model:
	$C \sim U(0.25, 4.00)$, if $Z = 0$, $C \sim U(0.07, 1.12)$, if $Z = 1$ for 30% censoring
	$C \sim U(0.25, 2.00)$, if $Z = 0$, $C \sim U(0.06, 0.46)$, if $Z = 1$ for 50% censoring

The regression coefficient β was estimated by the methods described in Section 2. We report the average of bias (Bias), the sample standard deviation of $\hat{\beta}$ (SD), the average of estimated standard error $\left(\bar{\hat{\sigma}}\right)$ using formula given in Appendix A, average of standardized bias ($\text{Std- B} = E\left\{|\hat{\beta} - \beta|/\hat{\sigma}\right.$), the coverage probability of β, and the mean squared error (MSE).

Table 1 shows the simulation results. We also examined the potential bias of estimating the cumulative baseline subdistribution hazard, $\Lambda_{10}^{*}(t)$, using both weights at a set of time points, $t = (0.25, 0.5, 0.75, 1.00)^{\top}$. Figure 1 shows the simulation results.

The simulation results show that when the censoring time depends on the covariate (scenario 2 and 3), the unadjusted estimator produces significantly biased results, and the estimator using the covariate-adjusted censoring weight provides satisfactory results where the biases are all close to zero. Both estimators give satisfactory variance estimates and have almost identical sample standard deviations (see scenario 2 and 3 in Table 1). Regarding the cumulative subdistribution hazard estimators, estimates using the Cox model adjusted weights have smaller biases compared to those using the unadjusted Kaplan-Meier weight at almost all time points (see Figure 1). Simulation results also indicated that the estimator using the Cox model adjusted weight provides satisfactory results when the Cox model is not the true model for the censoring distribution (see scenario 3 in Table 1 and Figure 1). In scenario 1, where the censoring distribution is independent of the covariate *Z*, both estimators provide satisfactory results in estimating the covariate effect and cumulative baseline subdistribution hazard function. Both estimators also have almost identical sample standard deviation and similar MSE, which indicate that the potential efficiency losses from modeling the censoring distribution are minimal when using covariate-adjusted censoring weights.

3.2 Study 2

The regression models below have one binary covariate Z_1 and one continuous covariate Z_2. Given Z_1 and Z_2, the cumulative incidence functions are given by

$$F_1(t;Z_1;Z_2) = 1 - \left\{1 - p\left(1 - e^{-t}\right)\right\}^{exp(\beta_1 Z_1 + \beta_2 Z_2)}$$

and

$$F_2(t;Z_1, Z_2) = (1-p)^{exp(\beta_1 Z_1 + \beta_2 Z_2)}\left\{1 - e^{-texp(\beta_1 Z_1 + \beta_2 Z_2)}\right\}.$$

We let $p = 0.66$, and Z_1 is a Bernoulli random variable, with a value 1 for half of the sample and 0 for the other half. Z_2 is a $N(0,1)$ random variable. We set $\beta_1 = 1$, $\beta_2 = 0.5$ and considered the following four scenarios.

Scenario 1	Censoring times are independent of Z_1 and Z_2
	Generate censoring times from an exponential distribution ~ $\exp(\lambda_C)$
	Set $\lambda_C = 0.547$ for 30% censoring, $\lambda_C = 1.352$ for 50% censoring
Scenario 2	Censoring times depend on Z_1 by a Cox model
	Generate censoring times from $\lambda_C(t\|\mathbf{Z}) = \lambda_C \exp(\beta_{C1} Z_1)$
	Set $\beta_{C1} = 2.5$. Set $\lambda_C = 0.137$ for 30% censoring,
	$\lambda_C = 0.397$ for 50% censoring

Scenario 3	Censoring times depend on Z_1 and Z_2 by a Cox model
	Generate censoring times from $\lambda_C(t\|\mathbf{Z}) = \lambda_C \exp(\beta_{C1}Z_1 + \beta_{C2}Z_2)$
	Set $\beta_{C1} = 2.5$, $\beta_{C2} = 2.5$. Set $\lambda_C = 0.082$ for 30% censoring,
	$\lambda_C = 0.389$ for 50% censoring
Scenario 4	Censoring times depend on Z_1, not by a Cox model
	$C \sim U(0.25, 4.00)$, if $Z_1 = 0$, $C \sim U(0.07, 1.14)$, if $Z_1 = 1$ for 30% censoring
	$C \sim U(0.25, 2.00)$, if $Z_1 = 0$, $C \sim U(0.06, 0.438)$, if $Z_1 = 1$ for 50% censoring

For each setting, we simulated 10,000 replicates with $n = 100$ and 300. The regression coefficients β_1 and β_2 were estimated by the methods described in Section 2. Table 2 shows the simulation results. We also examined the potential bias of estimating the cumulative baseline subdistribution hazard, $\Lambda_{10}^{*}(t)$, using both weights at a set of time points $t = (0.25, 0.5, 0.75, 1.00)^T$ for selected scenarios. Figure 2 shows the simulation results.

This simulation study shows similar results as in study 1. The unadjusted estimator produces biased results when the censoring distribution depends on the covariates (scenario 2 to 4), and the estimator using the Cox model adjusted weight provides a good bias reduction. Both estimators give satisfactory variance estimates for both parameters. Regarding the cumulative baseline subdistribution hazard estimates, estimates using the Cox-adjusted weight have smaller biases at almost all points (see Figure 2).

Both simulation studies show that the unadjusted estimator produces significant biased results when the censoring time depends on the covariates and the proposed estimator using covariate adjusted weight works well in bias reduction.

4 Real data examples

4.1 Example 1

We considered data from multiple myeloma patients treated with allogeneic stem cell transplantation from the Center for International Blood and Marrow Transplantat Research (CIBMTR) (Kumar et al., 2012). The CIBMTR is comprised of clinical and basic scientists who share data on their blood and bone marrow transplant patients with the CIBMTR Data Collection Center located at the Medical College of Wisconsin. The CIBMTR has a repository of information regarding the results of transplants at more than 450 transplant centers worldwide. The data used in this paper consist of patients transplanted from 1995 to 2005, and we compared the outcomes between transplant periods: 2001-2005 (N=488) versus 1995-2000 (N=375) (Kumar et al., 2012). Two competing events are multiple myeloma relapse and treatment-related mortality (TRM) defined as death without relapse. The CIBMTR study (Kumar et al., 2012) identified that donor type and prior autologous transplantation were associated with relapse or TRM. The variables considered in this example are transplant time period (GP (main interest of the study): 1 for transplanted in 2001-2005 versus 0 for transplanted in 1995-2000), donor type (DNR: 1 for Unrelated or other related donor (N=280) versus 0 for HLA-identical sibling (N=584)), and prior autologous transplant (PREAUTO: 1 for Auto+Allo transplant (N=399) versus 0 for allogeneic transplant alone (N=465)).

First, we fit a Cox model for the censoring distribution where relapsed or dead individuals are considered as censoring subjects. The hazard ratios (HR) are: HR(GP)=6.42 ($P <$ 0.0001); HR(DNR)=0.48 ($p = 0.0018$); HR(PREAUTO)=1.73 ($p = 0.0013$). These results indicate that the censoring distribution depends on the transplant time period, donor type and prior autologous transplantation. Next, we fit a proportional subdistribution hazards model (2.1) with the Kaplan-Meier estimated unadjusted weight and the Cox model adjusted weight, and we computed the predicted cumulative incidence probability for a patient who received an HLA-identical sibling donor allogeneic transplantation in 1995-2000 or in 2001-2005 (see results in Table 3–4 and Figure 3). Both weights give similar estimates for TRM. However, for cancer relapse, the regression estimate of the main treatment effect are $\hat{\beta}=0.38$ and $\hat{\beta}=0.54$ by unadjusted weight and Cox model adjusted weight, respectively. At three years after transplant, the differences in cumulative incidence of relapse between late and early transplant (TX) patients are 0.09 (CIF=0.34 for the late TX versus CIF=0.25 for the early TX) and 0.13 (CIF=0.35 for the late TX versus CIF=0.22 for the early TX) by unadjusted weight and Cox model adjusted weight, respectively. The unadjusted weight underestimates the effect size of CIF of relapse by 4% compared to the point estimate using the Cox model adjusted weight (Table 4). Underestimated effect size counts about 14% (0.04/((0.22+0.35)/2)) of estimated average CIF, which leads to quite a large relative bias.

4.2 Example 2

We considered another CIBMTR study data set (Ringdén et al., 2012) that consists of 177 myeloma patients who received a reduced-intensity conditioning allogeneic transplantation. Cancer relapse and TRM were two competing risks in this study. 105 patients received prior autologous transplant, and 72 patients received allogeneic transplant alone. We were interested in transplant type effect on relapse and TRM. Let PREAUTO be the indicator of transplant type (1 for Auto+Allo transplant versus 0 for Allogeneic transplant alone). Here the censoring distribution depends on the transplant type ($p = 0.0047$). We fit a proportional subdistribution hazards model (2.1) for PREAUTO with unadjusted weight and Cox model adjusted weight, respectively. For relapse, we have $\hat{\beta}_{\text{cox}}=-0.34\,(\hat{\sigma}=0.25)$; $exp\left(\hat{\beta}_{\text{cox}}\right)=0.71$ and $\hat{\beta}_{KM}=-0.41\,(\hat{\sigma}=0.25)$; $exp\left(\hat{\beta}_{KM}\right)=0.66$. Here the Cox model adjusted weight reduces a relative bias of 21% ((0.41 − 0.34)/0.34).

5 Concluding remarks

We have shown that the competing risks regression based on IPCW techniques may be biased when the censoring distribution depends on the covariates and the biases could be significant for fixed sample sizes. We considered a regression model for the censoring distribution, and used the Cox proportional hazards model to predict the censoring weight for each individual.

Clearly, using the Cox model to estimate the censoring weights rely on this model fitting well. Our methodology may be adapted to deal with other regression models for the censoring distribution, for example additive hazards models that are more flexible. Efficient variance estimator, which includes variation contributed from estimated censoring distribution, needs to be derived for any alternative model-based weight function, and a computing package needs to be further developed as well.

The censoring time could depend on a time-dependent covariate, but in this case the predictions given the fixed covariates of the competing risks regression model may be hard to get, and this is generally not directly feasible. As Kalbfleisch and Prentice (2002) pointed out that the predicted survival probability is no longer feasible for a random (internal) time-dependent covariate. Further study will be needed.

Recently, the inverse probability of censoring weighting (IPCW) technique (Robins and Rotnitzky, 1992) has been used extensively for right-censored survival data and, specifically, for completing risks data. It has been shown that regression modeling of the censoring distribution can be used to improve the efficiency of the IPCW technique (Bickel et al., 1993; Van der Laan and Robins, 2003; Scheike et al., 2008) even if the censoring distribution is independent of the covariates. In this study, we showed that the covariate-adjusted IPCW technique can be used to reduce bias for modeling the subdistribution hazard function when censoring depends on the covariates. In general, the covariate-adjusted IPCW technique should be considered to improve efficiency and reduce bias.

We have developed an R-package, wtcrsk, which is available on CRAN.

6 Appendix A

Here we present regularity conditions and give detailed proofs for the asymptotic properties of $\hat{\beta}_{COX}$ and $\hat{\Lambda}_{10}^{\mathrm{COX}}(t)$. Similar arguments can been seen in Ghosh and Lin (2002). First, assuming the censoring distribution depends on covariates $\boldsymbol{X}$ through a Cox proportional hazards model where $\boldsymbol{X}$ could be a subset covariates of $\boldsymbol{Z}$,

$$\lambda_{\mathrm{C}}(t;\boldsymbol{X})=\lambda_{\mathrm{C}0}(t)\ exp\{\gamma_0^{\intercal}\boldsymbol{X}\}.$$

Let $N_i^{\mathrm{C}}(t)=I(T_i\le t;\Delta_i=0)$, $Y_i(t) = I(T_i \ge t)$, $\boldsymbol{U}_{\mathrm{C}}(\gamma)$ be the partial likelihood for the censoring time, and

$$\boldsymbol{S}_{\mathrm{C}}^{(k)}(\gamma,t)=\frac{1}{n}\sum_i Y_i(t)\,\boldsymbol{X}_i^{\otimes k}\,exp\{\gamma^{\intercal}\boldsymbol{X}_i\},\ \text{for } k=0,1,2$$

$$s_{\mathrm{C}}^{(k)}(\gamma,t)=E\left\{Y_1(t)\,\boldsymbol{X}_1^{\otimes k}e^{\gamma^{\intercal}\boldsymbol{X}_1}\right\},\ \text{for } k=0,1,2.$$

Further, let

$$\boldsymbol{S}_{\mathrm{COX}}^{(k)}(\beta,t)=\frac{1}{n}\sum_i \hat{w}_i^{\mathrm{COX}}(t)\,Y_i^1(t)\,\boldsymbol{Z}_i^{\otimes k}exp\{\beta^{\intercal}\boldsymbol{Z}_i\},\ \text{for } k=0,1,2.$$

$$\boldsymbol{s}_{\mathrm{COX}}^{(k)}(\beta,t)=E\left\{Y_1^1(t)\,G_{\mathrm{C}}^{\mathrm{COX}}(t;\mathbf{X}_1)\,\boldsymbol{Z}_1^{\otimes k}e^{\beta^\intercal \boldsymbol{Z}_1}\right\},\ \text{for } k=0,1,2,$$

where $G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X})=exp\{-\Lambda_{\mathrm{C}0}(t)\ exp(\gamma_0^\intercal \boldsymbol{X})\}$ and $\Lambda_{\mathrm{C}0}(t)=\int_0^t \lambda_{CO}(u)\,du$.

We assume the following regularity conditions to be hold throughout the appendix.

6.1 Assumptions

(A1) $\{\widetilde{T}_i, \boldsymbol{X}_i, \boldsymbol{Z}_i, Y_i^1, \Delta_i, \varepsilon_i\}$, $i = 1, \ldots, n$, are i.i.d. instances of $\{\widetilde{T}, \boldsymbol{X}, \boldsymbol{Z}, Y^1, \Delta, \varepsilon\}$.

(A2) $\widetilde{T}$ and C are independent conditional on $\boldsymbol{X}, \boldsymbol{Z}$.

(A3) There is a maximum follow-up time $\tau < \infty$ such that $P(T > \tau) > 0$.

(B1) The hazard for the right-censoring times is $\lambda_{\mathrm{C}}(t;\boldsymbol{X})=\lambda_{\mathrm{C}0}(t)\ exp\{\gamma_0^\intercal \boldsymbol{X}\}$, where $\int_0^\tau \lambda_{\mathrm{C}0}(t)\,dt<\infty$ and $\gamma_0 \in \mathscr{G}$, for a compact $\mathscr{G}$.

(B2) The covariates $\boldsymbol{X}$ are bounded almost surely.

(B3) The matrix

$$\Omega_{\mathrm{C}}=\int_0^\tau\left[\frac{\boldsymbol{s}_{\mathrm{C}}^{(2)}(\gamma_0,t)}{s_{\mathrm{C}}^{(0)}(\gamma_0,t)}-\left\{\frac{\boldsymbol{s}_{\mathrm{C}}^{(1)}(\gamma_0,t)}{s_{\mathrm{C}}^{(0)}(\gamma_0,t)}\right\}^{\otimes 2}\right]s_{\mathrm{C}}^{(0)}(\gamma_0,t)\,\lambda_{\mathrm{C}0}(t)\,dt$$

is positive definite.

(C1) The subdistribution hazards for cause one is $\lambda_1^*(t|\boldsymbol{Z})=\lambda_{10}^*(t)\ exp(\beta_0^\intercal \boldsymbol{Z})$, where $\int_0^\tau \lambda_{10}^*(t)\,dt<\infty$ and $\beta_0 \in \mathscr{B}$, for a compact $\mathscr{B}$.

(C2) The covariates $\boldsymbol{Z}$ are bounded almost surely.

(C3) The matrix

$$\Omega=\int_0^\tau\left[\frac{\boldsymbol{s}_{\mathrm{COX}}^{(2)}(\beta_0,t)}{s_{\mathrm{COX}}^{(0)}(\beta_0,t)}-\left\{\frac{\boldsymbol{s}_{\mathrm{COX}}^{(1)}(\beta_0,t)}{s_{\mathrm{COX}}^{(0)}(\beta_0,t)}\right\}^{\otimes 2}\right]s_{\mathrm{COX}}^{(0)}(\beta_0,t)\,\lambda_{10}^*(t)\,dt$$

is positive definite.

6.2 Preliminaries

To estimate the censoring distribution we use Cox regression with the roles of $\widetilde{T}$ and C exchanged. Any administrative censoring events at time τ are not considered as events in the regression estimation. From assumption (B1) the censoring process $N^{\mathrm{C}}(t)$, for $0 \le t < \tau$, has intensity on the Cox model form. For this process, conditions (A1), (A2), (A3), (B1), (B2), and (B3), are sufficient to fulfill the conditions for the large sample results on the Cox model of Andersen and Gill (1982) (see their Theorem 4.1).

Let $\hat{\gamma}$ and $\hat{\Lambda}_{C0}$ be the Cox estimates of γ_0 and Λ_{C0}, and let

$$\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)=exp\left\{-\hat{\Lambda}_{\mathrm{C}0}(t)\ exp\left(\hat{\gamma}^{\intercal}\boldsymbol{X}_i\right)\right\}$$

By the arguments given by Andersen and Gill (1982), see also Equation (2.1) of Lin et al. (1994), we have

$$\begin{aligned}&\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)\\&-G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)=-\frac{G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{n}\sum_{j=1}^{n}\int_{u=0}^{\tau}\left[\boldsymbol{h}^{\intercal}(t,0,\boldsymbol{X}_i)\,\Omega_{\mathrm{C}}^{-1}\left\{\boldsymbol{X}_j-\frac{\boldsymbol{s}_{\mathrm{C}}^{(1)}(\gamma_0,u)}{s_{\mathrm{C}}^{(0)}(\gamma_0,u)}\right\}+\frac{e^{\left(\gamma_0^{\intercal}\boldsymbol{X}_i\right)}I(u\le t)}{\boldsymbol{s}_{\mathrm{C}}^{0}(\gamma_0,u)}\right]dM_{\mathrm{COX},j}^{\mathrm{C}}(u)\\&+o_P\left(n^{-1/2}\right),\end{aligned} \quad (6.1)$$

where Ω_C was defined in (B3), and

$$\boldsymbol{h}(t,u,\boldsymbol{X})=e^{\left(\gamma_0^{\intercal}\boldsymbol{X}\right)}\int_{v=u}^{t}\left\{\boldsymbol{X}-\frac{\boldsymbol{s}_{\mathrm{C}}^{(1)}(\gamma_0,v)}{s_{\mathrm{C}}^{(0)}(\gamma_0,v)}\right\}d\Lambda_{\mathrm{C}0}(v),$$

and

$$M_{\mathrm{COX},i}^{\mathrm{C}}(t)=N_i^{\mathrm{C}}(t)-\int_0^t Y_i(u)\ exp\left\{\gamma_0^{\intercal}\boldsymbol{X}_i\right\}\,d\Lambda_{\mathrm{C}0}(u),$$

is a martingale with respect to the censoring filtration. Note that by (6.1),

$$\begin{aligned}&\frac{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(T_i\wedge t;\boldsymbol{X}_i)}-\frac{G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{G_{\mathrm{C}}^{\mathrm{COX}}(T_i\wedge t;\boldsymbol{X}_i)}\\&=I(T_i<t)\,\frac{G_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)\left\{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)-G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)\right\}-G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)\left\{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)-G_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)\right\}}{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)G_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)}\\&=-I(T_i<t)\,\frac{G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{G_{\mathrm{C}}^{\mathrm{COX}}(T_i;\boldsymbol{X}_i)}\\&\quad\times\frac{1}{n}\sum_{j=1}^{n}\int_{u=0}^{\tau}\left[\boldsymbol{h}^{\intercal}(t,T_i,\boldsymbol{X}_i)\,\Omega_{\mathrm{C}}^{-1}\left\{\boldsymbol{X}_j-\frac{\boldsymbol{s}_{\mathrm{C}}^{(1)}(\gamma_0,u)}{s_{\mathrm{C}}^{(0)}(\gamma_0,u)}\right\}+\frac{e^{\gamma_0^{\intercal}\boldsymbol{X}_i}I(T_i<u\le t)}{s_{\mathrm{C}}^{(0)}(\gamma_0,u)}\right]dM_{\mathrm{COX},j}^{\mathrm{C}}(u)\\&\quad+o_P\left(n^{-1/2}\right).\end{aligned} \quad (6.2)$$

Let $w_i^{\mathrm{COX}}(t)=r_i(t)\,G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)/G_{\mathrm{C}}^{\mathrm{COX}}(T_i\wedge t;\boldsymbol{X}_i)$. From the i.i.d. assumption (A1) and the boundedness implied by (A3), (B1), (B2), (C1) and (C2)

$$\begin{aligned}&\boldsymbol{S}_{\mathrm{COX}}^{(k)}(\beta,t)\\&=\frac{1}{n}\sum_{i=1}^{n}Y_i^1(t)\,w_i^{\mathrm{COX}}(t)\,\boldsymbol{Z}_i^{\otimes k}e^{\beta^{\intercal}\boldsymbol{Z}_i}+\frac{1}{n}\sum_{i=1}^{n}Y_i^1(t)\,r_i(t)\left\{\frac{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{\hat{G}_{\mathrm{C}}^{\mathrm{COX}}(T_i\wedge t;\boldsymbol{X}_i)}-\frac{G_{\mathrm{C}}^{\mathrm{COX}}(t;\boldsymbol{X}_i)}{G_{\mathrm{C}}^{\mathrm{COX}}(T_i\wedge t;\boldsymbol{X}_i)}\right\}\boldsymbol{Z}_i^{\otimes k}e^{\beta^{\intercal}\boldsymbol{Z}_i}\xrightarrow{P}s_{\mathrm{COX}}^{(k)}(\beta,t),\end{aligned}$$

uniformly in $t \in [0, \tau]$ and $\beta \in \mathscr{B}$, by (6.2) and the uniform weak law of large numbers. Also note that, by the same assumptions, the limit is bounded from above and bounded away from zero, and we may take derivatives by differentiating under the integral sign.

6.3 Asymptotic normality of $n^{-1/2}U_{COX}(\beta_0)$

The IPCW score function (2.3) for β evaluated at β_0 is

$$\begin{aligned} U_{\text{COX}}(\beta_0) &= \sum_{i=1}^{n}\int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\boldsymbol{S}^{(1)}_{\text{COX}}(\beta_0,t)}{S^{(0)}_{\text{COX}}(\beta_0,t)}\right\}\hat{w}_i^{\text{COX}}(t)\,dM_i^1(t) \\ &= \sum_{i=1}^{n}\int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\boldsymbol{s}^{(1)}_{\text{COX}}(\beta_0,t)}{s^{(0)}_{\text{COX}}(\beta_0,t)}\right\}w_i^{\text{COX}}(t)\,dM_i^1(t) \\ &\quad + \sum_{i=1}^{n}\int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\boldsymbol{S}^{(1)}_{\text{COX}}(\beta_0,t)}{S^{(0)}_{\text{COX}}(\beta_0,t)}\right\}\left\{\hat{w}_i^{\text{COX}}(t) - w_i^{\text{COX}}(t)\right\}dM_i^1(t) + o_P\left(n^{1/2}\right). \end{aligned}$$

The first term on the right-hand side above is a sum of mean zero i.i.d. random variables. From (6.2), the second term on the right-hand side is

$$\begin{aligned} &\sum_{i=1}^{n}\int_0^\tau \left\{\boldsymbol{Z}_i - \frac{\boldsymbol{S}^{(1)}_{\text{COX}}(\beta_0,t)}{S^{(0)}_{\text{COX}}(\beta_0,t)}\right\} r_i(t)\left\{\frac{\hat{G}_{\text{C}}^{\text{COX}}(t;\boldsymbol{X}_i)}{\hat{G}_{\text{C}}^{\text{COX}}(T_i;\boldsymbol{X}_i)} - \frac{G_{\text{C}}^{\text{COX}}(t;\boldsymbol{X}_i)}{G_{\text{C}}^{\text{COX}}(T_i;\boldsymbol{X}_i)}\right\}dM_i^1(t) \\ &= \sum_{i=1}^{n}\int_{u=0}^{\tau}\boldsymbol{q}_i^{(1)}(u)\,dM^{\text{C}}_{\text{COX},i}(u) \\ &\quad + o_P\left(n^{1/2}\right), \end{aligned} \tag{6.3}$$

where

$$\begin{aligned} &\boldsymbol{q}_i^{(1)}(u) \\ &= -\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\int_{t=T_j}^{\tau}\left\{\boldsymbol{Z}_j - \frac{\boldsymbol{s}^{(1)}_{\text{COX}}(\beta_0,t)}{s^{(0)}_{\text{COX}}(\beta_0,t)}\right\}w_j^{\text{COX}}(t) \\ &\times\left[\boldsymbol{h}^\intercal(t,T_j,\mathbf{X}_j)\,\Omega_{\text{C}}^{-1}\left\{\boldsymbol{X}_i - \frac{\boldsymbol{s}^{(1)}_{\text{C}}(\gamma_0,u)}{s^{(0)}_{\text{C}}(\gamma_0,u)}\right\} + \frac{e^{\left(\gamma_0^\intercal\boldsymbol{X}_j\right)}I(u\le t)}{s^{(0)}_{\text{C}}(\gamma_0,u)}\right]dM_j^1(t). \end{aligned}$$

Thus,

$$\frac{1}{\sqrt{n}}\boldsymbol{U}_{\text{COX}}(\beta_0) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\eta_i^{\text{COX}} + \psi_i^{\text{COX}}\right) + o_P(1), \tag{6.4}$$

where

$$\eta_i^{\text{COX}} = \int_0^\tau\left\{\boldsymbol{Z}_i - \frac{\boldsymbol{s}^{(1)}_{\text{COX}}(\beta_0,t)}{s^{(0)}_{\text{COX}}(\beta_0,t)}\right\}w_i^{\text{COX}}(t)\,dM_i^1(t),$$

$$\psi_i^{\text{COX}} = \int_0^\tau \boldsymbol{q}_i^{(1)}(t)\, dM_{\text{COX},i}^C(t).$$

and $\left(\eta_i^{\text{COX}} + \psi_i^{\text{COX}}\right)$ are independent and identically distributed zero-mean variables.

6.4 Consistency and asymptotic normality $\hat{\beta}_{\text{COX}}$

Consider the derivative of $\boldsymbol{U}_{\text{COX}}(\beta)$ with respect to β. Let

$$\boldsymbol{I}_{\text{COX}}\left(\widetilde{\beta}\right) = -\frac{\partial}{\partial\beta}\boldsymbol{U}_{\text{COX}}(\beta)\Big|_{\beta=\widetilde{\beta}} = \int_0^\tau \left[\frac{\boldsymbol{S}_{\text{COX}}^{(2)}\left(\widetilde{\beta}, t\right)}{S_{\text{COX}}^{(0)}\left(\widetilde{\beta}, t\right)} - \left\{\frac{\boldsymbol{S}_{\text{COX}}^{(1)}\left(\widetilde{\beta}, t\right)}{S_{\text{COX}}^{(0)}\left(\widetilde{\beta}, t\right)}\right\}^{\otimes 2}\right] \sum_{i=1}^n \hat{w}_i^{\text{COX}}(t)\, dN_i^1(t).$$

By the same boundedness arguments as used in connection with (6.2), $n^{-}\boldsymbol{I}(\beta)$ converges in probability uniformly in β to a continuous limit such that $\lim_{n\to\infty} n^{-1}\boldsymbol{I}_{\text{COX}}(\beta_0 = \Omega)$. From Assumption (C3), Ω is positive definite, and from the previous section we know that $n^{-1}\boldsymbol{U}_{\text{COX}}(\beta_0) = o_P(1)$. Thus, by the argument in the proof of Theorem 2 of Foutz (1977), $\hat{\beta}_{\text{COX}}$ converges in probability to β_0.

Because $U_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right) = 0$, it follows that

$$\begin{aligned} 0 &= \tfrac{1}{\sqrt{n}}\boldsymbol{U}_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right) \\ &= \tfrac{1}{\sqrt{n}}\boldsymbol{U}_{\text{COX}}(\beta_0) - \tfrac{1}{n}\int_0^1 \boldsymbol{I}_{\text{COX}}\left\{\beta_0 + v\left(\hat{\beta}_{\text{COX}} - \beta_0\right)\right\} dv\, \sqrt{n}\left(\hat{\beta}_{\text{COX}} - \beta_0\right), \end{aligned} \quad (6.5)$$

Then by the consistency of $\hat{\beta}_{\text{COX}}$, $n^{-1}\int_0^1 \boldsymbol{I}_{\text{COX}}\left\{\beta_0 + v\left(\hat{\beta}_{\text{COX}} - \beta_0\right)\right\} dv \xrightarrow{P} \Omega$. By (C3), Ω is invertible and (6.5) gives that

$$\begin{aligned} \sqrt{n}\left(\hat{\beta}_{\text{COX}} - \beta_0\right) &= \Omega^{-1}\left\{\tfrac{1}{\sqrt{n}}\boldsymbol{U}_{\text{COX}}(\beta_0)\right\} + 0_P(1) \\ &= \Omega^{-1}\left\{\tfrac{1}{\sqrt{n}}\sum_{i=1}^n \left(\eta_i^{\text{COX}} + \psi_i^{\text{COX}}\right)\right\} + 0_P(1), \end{aligned} \quad (6.6)$$

which is essentially a sum of bounded independent and identically distributed variables and thus by the central limit theorem it is asymptotically normally distributed with mean zero and variance matrix

$$\Sigma_\beta^{\text{COX}} = \Omega^{-1} E\left\{\left(\xi_1^{\text{COX}} + \psi_1^{\text{COX}}\right)\left(\xi_1^{\text{COX}} + \psi_1^{\text{COX}}\right)^{\intercal}\right\}\Omega^{-1}.$$

The variance matrix $\sum_\beta^{\text{COX}}$ can be estimated by a plug-in estimator

$$\hat{\Sigma}_\beta^{\text{COX}} = n\sum_i \left(\hat{\boldsymbol{W}}_{\beta,i}^{\text{COX}}\right)^{\otimes 2} = n\left\{\boldsymbol{I}_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right)\right\}^{-1}\left\{\sum_i \left(\hat{\xi}_i^{\text{COX}} + \hat{\psi}_i^{\text{COX}}\right)^{\otimes 2}\right\}\left\{\boldsymbol{I}_{\text{COX}}\left(\hat{\beta}_{\text{COX}}\right)\right\}^{-1},$$

where

$$\hat{\xi}_i^{\mathrm{COX}} = \int_0^\tau \left\{ \boldsymbol{Z}_i - \frac{\boldsymbol{S}_{\mathrm{COX}}^{(1)}\left(\hat{\beta}_{\mathrm{COX}}, t\right)}{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}}, t\right)} \right\} \hat{w}_i^{\mathrm{COX}}(t)\, d\hat{M}_{\mathrm{COX},i}^{1}(t)$$

$$\hat{\psi}_i^{\mathrm{COX}} = \int_{u=0}^\tau \hat{\boldsymbol{q}}_i^{(1)}(u)\, d\hat{M}_{\mathrm{COX},i}^{\mathrm{C}}(u)$$

$$\begin{aligned}
&\hat{q}_i^{(1)}(u) \\
&= -\sum_{j=1}^{n} \int_{t=T_j}^{\tau} \left\{ \boldsymbol{Z}_j - \frac{\boldsymbol{S}_{\mathrm{COX}}^{(1)}\left(\beta_{\mathrm{COX}}, t\right)}{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}}, t\right)} \right\} \hat{w}_j^{\mathrm{COX}}(t) \\
&\times \left[\hat{\boldsymbol{h}}^\intercal (t, T_j, \boldsymbol{X}_j) \left\{ \boldsymbol{I}_{\mathrm{C}}(\hat{\gamma}) \right\}^{-1} \left\{ \boldsymbol{X}_i - \frac{\boldsymbol{S}_{\mathrm{C}}^{(1)}(\hat{\gamma}, u)}{S_{\mathrm{C}}^{(0)}(\hat{\gamma}, u)} \right\} + \frac{e^{(\hat{\gamma}^\intercal \boldsymbol{X}_j)} I(u \le t)}{S_{\mathrm{C}}^{(0)}(\hat{\gamma}, u)} \right] d\hat{M}_{\mathrm{COX},j}^{1}(t)
\end{aligned}$$

$$\hat{\boldsymbol{h}}(t, T_i, \boldsymbol{X}_i) = e^{(\hat{\gamma}^\intercal \boldsymbol{X}_i)} \int_{u=T_i}^{t} \left\{ \boldsymbol{X}_i - \frac{\boldsymbol{S}_{\mathrm{C}}^{(1)}(\hat{\gamma}, u)}{S_{\mathrm{C}}^{(0)}(\hat{\gamma}, u)} \right\} d\hat{\Lambda}_{\mathrm{C}0}(u)$$

$$\boldsymbol{I}_{\mathrm{C}}(\gamma) = -\frac{\partial \boldsymbol{U}_{\mathrm{C}}(\gamma)}{\partial \gamma} = \sum_i \int_0^\tau \left\{ \frac{\boldsymbol{S}_{\mathrm{C}}^{(2)}(\gamma, t)}{S_{\mathrm{C}}^{(0)}(\gamma, t)} - \left(\frac{\boldsymbol{S}_{\mathrm{C}}^{(2)}(\gamma, t)}{S_{\mathrm{C}}^{(0)}(\gamma, t)} \right)^{\otimes 2} \right\} dN_i^{\mathrm{C}}(t)$$

$$\hat{w}_i^{\mathrm{COX}}(t)\, d\hat{M}_{\mathrm{COX},i}^{t}(t) = \hat{w}_i^{\mathrm{COX}}(t)\, dN_i^1(t) - \hat{w}_i^{\mathrm{COX}}(t)\, Y_i^1(t)\ exp\left(\hat{\beta}_{\mathrm{COX}}^\intercal \boldsymbol{Z}_i\right) d\hat{\Lambda}_{10}^{\mathrm{COX}}(t)$$

$$d\hat{M}_{\mathrm{COX},i}^{\mathrm{C}}(t) = dN_i^{\mathrm{C}}(t) - Y_i(t)\ exp\left(\hat{\gamma}^\intercal \boldsymbol{X}_i\right) d\hat{\Lambda}_{\mathrm{C}0}(t).$$

6.5 Consistency and weak convergence of $\hat{\Lambda}_{10}^{\mathrm{COX}}$

We first note that $\hat{\Lambda}_{10}^{\mathrm{COX}}(t)$ is uniformly consistent for $\Lambda_{10}^*(t) = \int_0^t \lambda^{10}(u)\, du$. To see this, write

$$\hat{\Lambda}_{10}^{\mathrm{COX}}(t) - \Lambda_{10}^*(t) = \frac{1}{n}\sum_{i=1}^{n} \int_0^t \frac{\hat{w}_i^{\mathrm{COX}}(u)\, dM_i^1(u)}{S_{\mathrm{COX}}^{(0)}(\beta_0, u)} + \frac{1}{n}\sum_{i=1}^{n} \int_0^t \left\{ \frac{1}{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}}, u\right)} - \frac{1}{S_{\mathrm{COX}}^{(0)}(\beta_0, u)} \right\} \hat{w}_i^{\mathrm{COX}}(u)\, dN_i^1(u). \quad (6.7)$$

By the arguments in Section 6.2, the consistency of $\hat{\beta}_{\text{COX}}$ and the boundedness away from zero and smoothness of $S^{(0)}_{\text{COX}}$, the right-hand side of (6.7) converges to zero in probability, uniformly in $t \in [0, \tau]$.

Consider the first term on the right-hand side of (6.7). By the same argument as used for establishing (6.3),

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{\hat{w}_i^{\text{COX}}(u)dM_i^1(u)}{S^{(0)}_{\text{COX}}(\beta_0,u)} &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{w_i^{\text{COX}}(u)dM_i^1(u)}{S^{(0)}_{\text{COX}}(\beta_0,u)} + \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{\hat{w}_i^{\text{COX}}(u)-w_i^{\text{COX}}(u)}{S^{(0)}_{\text{COX}}(\beta_0,u)} dM_i^1(s) \\ &= \frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{w_i^{\text{COX}}(s)sM_i^1(s)}{s^{(0)}_{\text{COX}}(\beta_0,u)} + \frac{1}{n}\sum_{i=1}^{n}\int_0^t q_i^{(2)}(u,t)\, dM^{\text{C}}_{\text{COX},i}(u) + o_P\left(n^{-1/2}\right),\end{aligned} \tag{6.8}$$

where

$$\begin{aligned} q_i^{(2)}(u,t) &= -\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}\int_{v=T_j}^{t}\frac{1}{s^{(0)}_{\text{COX}}(\beta_0,v)} \\ &\times\left[\boldsymbol{h}^{\intercal}(v,T_j,\boldsymbol{X}_j)\,\Omega_{\text{C}}^{-1}\left\{\boldsymbol{X}_i - \frac{\boldsymbol{s}^{(1)}_{\text{C}}(\gamma_0,u)}{\boldsymbol{s}^{(0)}_{\text{C}}(\gamma_0,u)}\right\} + \frac{e^{\left(\gamma_0^{\intercal}\boldsymbol{X}_j\right)}\boldsymbol{I}(u\le v)}{\boldsymbol{s}^{(0)}_{\text{C}}(\gamma_0,u)}\right] w_j^{\text{COX}}(v)\, dM_j^1(v).\end{aligned}$$

Similar to (6.5), by a first-order expansion, the last term on the right-hand side of (6.7), is

$$\begin{aligned}&-\int_0^t\left\{\frac{\boldsymbol{S}^{(1)}_{\text{COX}}(\beta_0,u)}{S^{(0)}_{\text{COX}}(\beta_0,u)}\right\}^{\intercal}\frac{n^{-1}\Sigma_{i=1}^{n}\hat{w}_i^{\text{COX}}(u)dN_i^1(u)}{S^{(0)}_{\text{COX}}(\beta_0,u)}\left(\hat{\beta}_{\text{COX}}-\beta_0\right)+o_P\left(n^{-1/2}\right) \\ &= \boldsymbol{h}_{\Lambda}^{\intercal}(t)\,\Omega^{-1}\frac{1}{n}\sum_{i=1}^{n}\left(\eta_i^{\text{COX}}+\psi_i^{\text{COX}}\right)+o_P\left(n^{-1/2}\right),\end{aligned} \tag{6.9}$$

where

$$\boldsymbol{h}_{\Lambda}(t) = -\int_0^t \frac{\boldsymbol{s}^{(1)}_{\text{COX}}(\beta_0,u)}{\boldsymbol{s}^{(0)}_{\text{COX}}(\beta_0,u)}\lambda_{10}^{*}(u)\,du.$$

Combining (6.5), (6.8) and (6.9),

$$\sqrt{n}\left\{\hat{\Lambda}_{10}^{\text{COX}}(t)-\Lambda_{10}^{*}(t)\right\} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}W_{\Lambda,i}^{\text{COX}}(t)+o_P(1), \tag{6.10}$$

where

$$W_{\Lambda,i}^{\text{COX}}(t) = \int_0^t\frac{w_i^{\text{COX}}(u)\,dM_i^1(u)}{s^{(0)}_{\text{COX}}(\beta_0,u)} + \int_0^t q_i^{(2)}(u,t)\,dM^{\text{C}}_{\text{COX},i}(u) + \boldsymbol{h}_{\Lambda}^{\intercal}(t)\,\Omega^{-1}\left(\eta_i^{\text{COX}}+\psi_i^{\text{COX}}\right).$$

The $W_{\Lambda,i}^{\text{COX}}(t)$ terms are i.i.d. stochastic processes forming a Donsker class indexed by t. To see this, note that M_i^1, $M^{\text{C}}_{\text{COX},i}$ and w_i^{COX} are of bounded variation and thus Donsker and that sums of bounded Donsker classes are Donsker (van der Vaart, 1998, Example 19.20, Example 19.11). Thus, $\sqrt{n}\left\{\hat{\Lambda}_{10}^{\mathrm{COX}}(t)-\Lambda_{10}^{*}(t)\right\}$ converges weakly to a mean-zero Gaussian process with variance $\Sigma_{\Lambda_{10}}^{\mathrm{COX}}(t)=E\left\{W_{\Lambda,i}^{\mathrm{COX}}(t)\right\}$. The variance can be estimated by a plug in estimator, $\hat{\Sigma}_{\Lambda_{10}}^{\mathrm{COX}}(t)=n^{-1}\sum_i\left\{\hat{W}_{\Lambda,i}^{\mathrm{COX}}(t)\right\}^2$, where

$$\hat{W}_{\Lambda,i}^{\mathrm{COX}}(t)=\int_0^t\frac{\hat{w}_i^{\mathrm{COX}}(u)\,d\hat{M}_i^1(u)}{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}},u\right)}+\int_0^\tau\hat{q}_i^{(2)}(u,t)\,d\hat{M}_{\mathrm{COX},i}^{\mathrm{C}}(u)+\hat{\boldsymbol{h}}_\Lambda^{\intercal}(t)\,\hat{\Omega}^{-1}\left(\hat{\eta}_i^{\mathrm{COX}}+\hat{\psi}_i^{\mathrm{COX}}\right)$$

$$\begin{aligned}\hat{q}_i^{(2)}(u,t)&\\&=-\frac{1}{n}\sum_{j=1}^{n}\int_{j=1}^{t}\frac{1}{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}},v\right)}\\&\times\left[\hat{\boldsymbol{h}}(v,T_j,\boldsymbol{X}_j)\,\hat{\Omega}_{\mathrm{C}}^{-1}\left\{\boldsymbol{X}_i-\frac{\boldsymbol{S}_{\mathrm{C}}^{(1)}(\hat{\gamma},u)}{S_{\mathrm{C}}^{(0)}(\hat{\gamma},u)}\right\}+\frac{e^{(\hat{\gamma}\boldsymbol{X}_j)}\boldsymbol{I}(u\le v)}{S_{\mathrm{C}}^{(0)}(\hat{\gamma},u)}\right]\hat{w}_j^{\mathrm{COX}}(v)\,d\hat{M}_j^1(v)\end{aligned}$$

$$\hat{\boldsymbol{h}}_\Lambda(t)=-\frac{1}{n}\sum_{i=1}^{n}\int_0^t\frac{\boldsymbol{S}_{\mathrm{COX}}^{(1)}\left(\hat{\beta}_{\mathrm{COX}},u\right)}{\left\{S_{\mathrm{COX}}^{(0)}\left(\hat{\beta}_{\mathrm{COX}},u\right)\right\}^2}\hat{w}_i^{\mathrm{COX}}(u)\,dN_i^1(u)\,.$$

7 Appendix B

Cumulative incidence functions for cause 1 and 2 are generated from

$$F_1(t)=1-\left\{1-p\left(1-e^{-1}\right)\right\}^{exp(betaZ)}$$

$$F_2(t)=(1-p)^{exp(\beta Z)}\left\{1-e^{-texp(\beta Z)}\right\}$$

where $0 < p < 1$. Generating data step (need to set parameters p and β first):

1. Generate covariate Z
2. Based on p, β and Z, compute P_1 and P_2 (probability of type 1 and type 2 failures): $P_1 = F_1(\infty) = 1 - (1 - p)^{\exp(\beta Z)}$ and $P_2 = F_2(\infty) = (1 - p)^{\exp(\beta Z)}$
3. Generate an uniform r.v. $U_1 \sim U(0, 1)$. Generate cause of death indicator, δ, by

$$\delta=\begin{cases}1, & \text{if } U_1\le P_1\\ 2, & \text{Otherwise}\end{cases}$$

4. Based on $\delta = k$, $k = 1,2$, compute the conditional probability

$$\widetilde{F}_k(t) = P(T_k \le t | \delta = k) = \frac{F_k(t)}{P(\delta = k)} = \frac{F_k(t)}{F_k(\infty)} = \frac{F_k(t)}{P_k}$$

5. Generate second uniform r.v. $U_2 \sim U(0,1)$. Then use inverse distribution method to generate T_k based on $\widetilde{F}_k(t)$.

References

Andersen PK, Gill RD. Cox's regression model for counting processes: A large sample study. The Annals of Statistics. 1982; 10:1100–1120.

Bickel, P.; Klassen, C.; Ritov, Y.; Wellner, J. Efficient and Adaptive Estimation for Semiparametric Models. Springer-Verlag; New York: 1993.

Cheng SC, Fine JP, Wei LJ. Prediction of cumulative incidence function under the proportional hazards model. Biometrics. 1998; 54:219–228. [PubMed: 9544517]

Cox DR. Regression models and life tables (with discussion). J. Roy. Statist. Soc. Ser. B. 1972; 34:187–220.

DiRienzo A, Lagakos S. Bias correction for score tests arising from misspecified proportional hazards regression models. Biometrika. 2001a; 88:421–434.

DiRienzo A, Lagakos S. Effects of model misspecification on tests of no randomized treatment effect arising from Cox's proportional hazards model. J. R. Statist. Soc. B. 2001b; 63:745–757.

Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. Journal of the American Statistical Association. 1999; 94:496–509.

Foutz R. On the unique consistent solution to the likelihood equations. Journal of the American Statistical Association. 1977; 72:147–148.

Ghosh D, Lin DY. Marginal regression models for recurrent and terminal events. Statistica Sinica. 2002; 12:663–688.

Gray RJ. A class of *k*-sample tests for comparing the cumulative incidence of a competing risk. The Annals of Statistics. 1988; 16:1141–1154.

Heinze G, Gnant M, Schemper M. Exact log-rank tests for unequal follow-up. Biometrics. 2003; 59:1151–1157. [PubMed: 14969496]

Kalbfleisch, JD.; Prentice, RL. The Statistical Analysis of Failure Time Data. Wiley; New York: 2002.

Kim H. Cumulative incidence in competing risks data and competing risks regression analysis. Clinical Cancer Research. 2007; 13:559–565. [PubMed: 17255278]

Kumar S, Zhang M, Li P, Dispenzieri A, Milone G, Lonial S, Krishnan A, Maiolino A, Wirk B, Weiss B, Freytes C, Vogl D, Vesole D, Lazarus H, Meehan K, Hamadani M, Lill M, Callander N, Majhail N, Wiernik P, Nath R, Kamble R, Vij R, Kyle R, Gale R, Hari P. Trends in allogeneic stem cell transplantation for multiple myeloma: a CIBMTR analysis. Blood. 2012; 118:1979–1988. [PubMed: 21690560]

Lau B, Cole S, Gange S. Competing risk regression models for epidemiologic data. American Journal of Epidemiology. 2009; 170:244–256. [PubMed: 19494242]

Lin DY, Fleming T, Wei L. Confidence bands for survival curves under the proportional hazards model. Biometrika. 1994; 81:73–81.

Mai, Y. PhD thesis. University of Maryland; 2008. Comparing Survival Distributions in the Presence of Dependent Censoring: Asymptotic Validity and Bias Corrections of the Logrank Test.

Pepe MS. Inference for events with dependent risks in multiple endpoint studies. J. Amer. Statist. Assoc. 1991; 86:770–778.

Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow N. The analysis of failure time data in the presence of competing risks. Biometrics. 1978; 34:541–554. [PubMed: 373811]

Ringdén O, Shrestha S, Tunes da Silva G, Zhang M, Dispenzieri A, Remberger M, Kamble R, Freytes C, O. Gale R, Gibson J, Gupta V, Holmberg L, Lazarus H, McCarthy P, Meehan K, Schouten H, Milone G, Lonial S, Hari P. Effect of acute and chronic GVHD on relapse and survival after reduced-intensity conditioning allogeneic transplantation for myeloma. Bone Marrow Transplantation. 2012; 47:831–837. [PubMed: 21946381]

Robins, JM.; Rotnitzky, A. AIDS Epidemiology-Methodological Issues. Birkhäuser; Boston: 1992. Recovery of information and adjustment of dependent censoring using surrogate markers; p. 24-33.

Scheike TH, Zhang M-J. An additive-multiplicative Cox-Aalen model. Scandinavian Journal of Statistics. 2002; 28:75–88.

Scheike TH, Zhang M-J. Extensions and applications of the Cox-Aalen survival model. Biometrics. 2003; 59:1033–1045.

Scheike TH, Zhang M-J, Gerds T. Predicting cumulative incidence probability by direct binomial regression. Biometrika. 2008; 95:205–20.

Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. Bone Marrow Transplant. 2007; 40(4):381–387. [PubMed: 17563735]

Shen Y, Cheng SC. Confidence bands for cumulative incidence curves under the additive risk model. Biometrics. 1999; 55:1093–1100. [PubMed: 11315053]

Van der Laan, MJ.; Robins, JM. Unified Methods for Censored Longitudinal Data and Causality. Springer; 2003.

van der Vaart, A. Asymptotic Statistics. Cambridge University Press; 1998.

Wolbers M, Koller M, Witteman J, Steyerberg E. Prognostic models with competing risks: methods and application to coronary risk prediction. Epidemiology. 2009; 20(4):555–561. [PubMed: 19367167]

Zhang M-J, Fine J. Summarizing differences in cumulative incidence functions. Statist. Med. 2008; 27:4939–49.

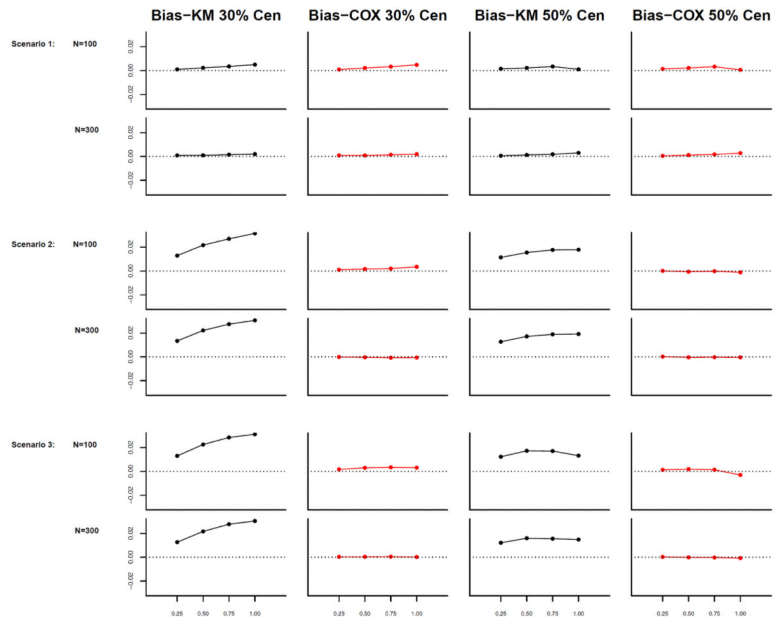

Figure 1.

Simulation results (1 covariate) for biases of cumulative baseline subdistribution hazards at $t = (0.25, 0.5, 0.75, 1)^T$ with 30% and 50% censoring, respectively.

Author Manuscript
Author Manuscript
Author Manuscript

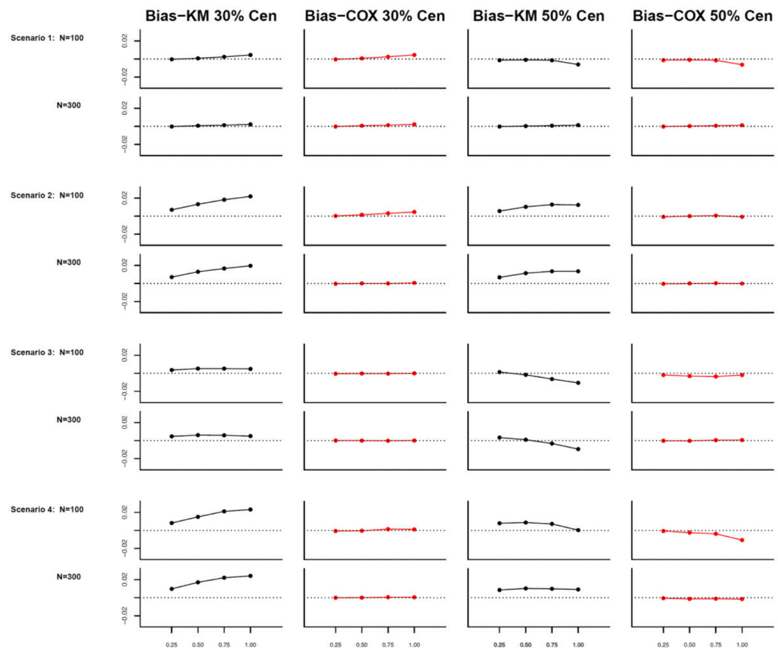

Figure 2.

Simulation results (2 covariates) for biases of cumulative baseline subdistribution hazards at $t = (0.25, 0.5, 0.75, 1.00)^T$ with 30% and 50% censoring, respectively.

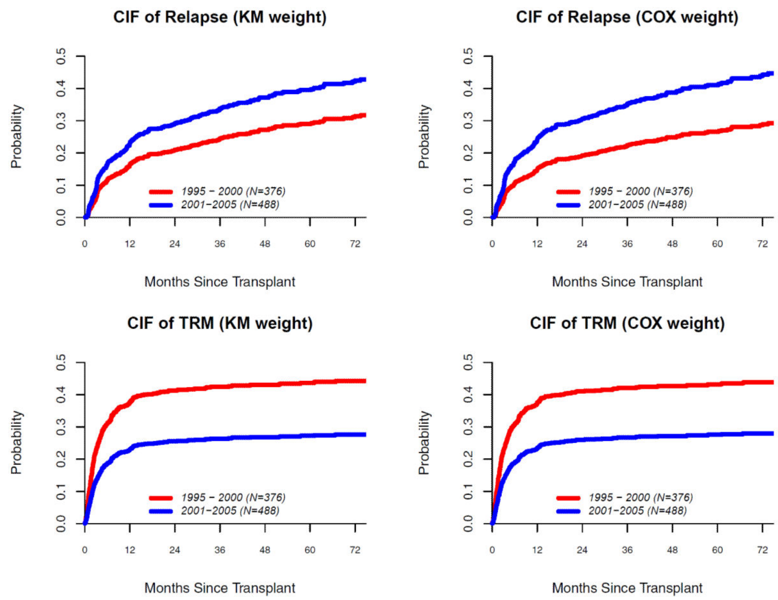

Figure 3.

Predicted cumulative incidence probability of relapse and TRM for a patient who received an HLA-identical sibling donor allogeneic transplantation.

Table 1

Simulation results for bias reduction with a single binary covariate (β = 1).

				Unadjusted weight				Cox model adjusted weight			
Scenario	**N**	**Cens.**	**β**	**Bias (Std-B)**	$\overline{\hat{\sigma}}$**(SD)**	**Coverage**	**MSE**	**Bias (Std-B)**	$\overline{\hat{\sigma}}$**(SD)**	**Coverage**	**MSE**
1	100	30%	β	0.0081 (0.0291)	0.2759 (0.2798)	0.9494	0.0783	0.0089 (0.0319)	0.2757 (0.2797)	0.9495	0.0783
		50%	β	0.0145 (0.0430)	0.3329 (0.3384)	0.9488	0.1147	0.0153 (0.0450)	0.3331 (0.3387)	0.9489	0.1149
	300	30%	β	0.0010 (0.0061)	0.1584 (0.1586)	0.9499	0.0252	0.0013 (0.0079)	0.1580 (0.1583)	0.9512	0.0251
		50%	β	0.0030 (0.0155)	0.1899 (0.1909)	0.9498	0.0365	0.0033 (0.0174)	0.1898 (0.1908)	0.9503	0.0364
2	100	30%	β	−.1119 (0.3592)	0.3041 (0.3116)	0.9278	0.1096	0.0050 (0.0160)	0.3080 (0.3123)	0.9487	0.0976
		50%	β	−.1162 (0.2719)	0.4075 (0.4271)	0.9359	0.1959	0.0175 (0.0398)	0.4217 (0.4388)	0.9501	0.1929
	300	30%	β	−.1244 (0.7109)	0.1741 (0.1750)	0.8865	0.0461	0.0042 (0.0236)	0.1765 (0.1762)	0.9503	0.0311
		50%	β	−.1336 (0.5596)	0.2306 (0.2346)	0.9063	0.0729	0.0055 (0.0226)	0.2393 (0.2409)	0.9511	0.0581
3	100	30%	β	−.1020 (0.3383)	0.2937 (0.3016)	0.9312	0.1013	0.0045 (0.0145)	0.3036 (0.3096)	0.9503	0.0958
		50%	β	−.0988 (0.2507)	0.3797 (0.3941)	0.9350	0.1650	0.0168 (0.0401)	0.4057 (0.4198)	0.9487	0.1765
	300	30%	β	−.1099 (0.6491)	0.1685 (0.1692)	0.8927	0.0407	0.0026 (0.0147)	0.1748 (0.1749)	0.9535	0.0306
		50%	β	−.1088 (0.5009)	0.2152 (0.2173)	0.9152	0.0591	0.0072 (0.0310)	0.2309 (0.2332)	0.9508	0.0544

Bias=Average of bias of $\hat{\beta}$; SD=Sample standard deviation of $\hat{\beta}$; $\overline{\hat{\sigma}}$=Average of estimated standard error; Std-B = $E\left\{|\hat{\beta}-\beta|/\hat{\sigma}\right\}$; MSE=Mean squared error.

Table 2

Simulation results for biases using 1 binary and 1 continuous covariate ($\beta_1 = 1, \beta_2 = 0.5$).

				Unadjusted weight				Cox model adjusted weight			
Scenario	N	Cens.	β	Bias (Std-B)	$\hat{\sigma}(SD)$	Coverage	MSE	Bias (Std-B)	$\hat{\sigma}(SD)$	Coverage	MSE
1	100	30%	β_1	0.0148 (0.0521)	0.2793 (0.2849)	0.9462	0.0814	0.0154 (0.0542)	0.2793 (0.2847)	0.9463	0.0813
			β_2	0.0130 (0.0873)	0.1424 (0.1483)	0.9381	0.0222	0.0130 (0.0876)	0.1424 (0.1482)	0.9394	0.0221
		50%	β_1	0.0293 (0.0847)	0.3353 (0.3463)	0.9474	0.1207	0.0299 (0.0863)	0.3355 (0.3464)	0.9468	0.1209
			β_2	0.0143 (0.0812)	0.1676 (0.1763)	0.9362	0.0313	0.0146 (0.0827)	0.1676 (0.1764)	0.9362	0.0313
	300	30%	β_1	0.0039 (0.0242)	0.1595 (0.1608)	0.9498	0.0259	0.0041 (0.0254)	0.1592 (0.1602)	0.9495	0.0257
			β_2	0.0042 (0.0521)	0.0807 (0.0813)	0.9494	0.0066	0.0043 (0.0535)	0.0806 (0.0811)	0.9485	0.0066
		5%0	β_1	0.0062 (0.0327)	0.1898 (0.1883)	0.9525	0.0355	0.0063 (0.0333)	0.1897 (0.1881)	0.9529	0.0354
			β_2	0.0063 (0.0661)	0.0941 (0.0958)	0.9448	0.0092	0.0065 (0.0674)	0.0941 (0.0957)	0.9438	0.0092
2	100	30%	β_1	−.0938 (0.2949)	0.3036 (0.3180)	0.9243	0.1099	0.0119 (0.0370)	0.3101 (0.3220)	0.9421	0.1038
			β_2	0.0273 (0.1778)	0.1462 (0.1536)	0.9355	0.0243	0.0160 (0.1037)	0.1468 (0.1540)	0.9357	0.0240
		50%	β_1	−.0955 (0.2241)	0.4031 (0.4262)	0.9325	0.1908	0.0179 (0.0409)	0.4181 (0.4385)	0.9414	0.1926
			β_2	0.0247 (0.1349)	0.1740 (0.1831)	0.9333	0.0341	0.0142 (0.0771)	0.1748 (0.1839)	0.9354	0.0340
	300	30%	β_1	−.1089 (0.6286)	0.1734 (0.1733)	0.9002	0.0419	0.0049 (0.0277)	0.1771 (0.1769)	0.9514	0.0313
			β_2	0.0174 (0.2056)	0.0827 (0.0845)	0.9409	0.0074	0.0050 (0.0589)	0.0831 (0.0848)	0.9428	0.0072
		50%	β_1	−.1136 (0.4956)	0.2271 (0.2291)	0.9153	0.0654	0.0049 (0.0205)	0.2363 (0.2380)	0.9516	0.0566
			β_2	0.0142 (0.1418)	0.0979 (0.1002)	0.9407	0.0102	0.0027 (0.0273)	0.0984 (0.1004)	0.9434	0.0101
3	100	30%	β_1	−.0299 (0.0945)	0.3061 (0.3168)	0.9418	0.1012	0.0145 (0.0459)	0.3070 (0.3149)	0.9466	0.0994
			β_2	−.0527 (0.2665)	0.1898 (0.1979)	0.9224	0.0419	0.0014 (0.0069)	0.1934 (0.1992)	0.9422	0.0397
		50%	β_1	−.0692 (0.1711)	0.3872 (0.4042)	0.9366	0.1682	0.0163 (0.0398)	0.3935 (0.4090)	0.9441	0.1676
			β_2	−.0963 (0.3695)	0.2501 (0.2605)	0.9087	0.0771	−.0007 (0.0025)	0.2630 (0.2728)	0.9380	0.0744
	300	30%	β_1	−.0473 (0.2687)	0.1750 (0.1762)	0.9363	0.0333	0.0018 (0.0102)	0.1749 (0.1758)	0.9486	0.0309
			β_2	−.0631 (0.5767)	0.1079 (0.1094)	0.8987	0.0159	−.0016 (0.0148)	0.1097 (0.1101)	0.9494	0.0121
		50%	β_1	−.0874 (0.3920)	0.2196 (0.2231)	0.9266	0.0574	0.0046 (0.0206)	0.2230 (0.2251)	0.9489	0.0507
			β_2	−.1035 (0.7300)	0.1404 (0.1418)	0.8698	0.0308	0.0022 (0.0146)	0.1488 (0.1504)	0.9451	0.0226

				Unadjusted weight				Cox model adjusted weight			
Scenario	**N**	**Cens.**	**β**	**Bias (Std-B)**	$\hat{\sigma}(SD)$	**Coverage**	**MSE**	**Bias (Std-B)**	$\hat{\sigma}(SD)$	**Coverage**	**MSE**
4	100	30%	β_1	−.0773 (0.2578)	0.2934 (0.2999)	0.9358	0.0959	0.0149 (0.0481)	0.3035 (0.3089)	0.9476	0.0956
			β_2	0.0251 (0.1642)	0.1441 (0.1526)	0.9346	0.0239	0.0146 (0.0952)	0.1444 (0.1529)	0.9361	0.0236
		50%	β_1	−.0731 (0.1868)	0.3799 (0.3913)	0.9402	0.1584	0.0265 (0.0636)	0.4040 (0.4163)	0.9494	0.1740
			β_2	0.0248 (0.1390)	0.1708 (0.1783)	0.9361	0.0324	0.0150 (0.0840)	0.1713 (0.1787)	0.9366	0.0322
	300	30%	β_1	−.0924 (0.5450)	0.1672 (0.1695)	0.9076	0.0372	0.0055 (0.0314)	0.1735 (0.1750)	0.9485	0.0306
			β_2	0.0142 (0.1703)	0.0816 (0.0833)	0.9439	0.0071	0.0028 (0.0330)	0.0818 (0.0834)	0.9462	0.0070
		50%	β_1	−.0877 (0.4128)	0.2149 (0.2124)	0.9327	0.0528	0.0101 (0.0446)	0.2290 (0.2269)	0.9544	0.0516
			β_2	0.0123 (0.1255)	0.0960 (0.0977)	0.9434	0.0097	0.0023 (0.0232)	0.0963 (0.0982)	0.9455	0.0096

Bias=Average of bias of $\hat{\beta}$; SD=Sample standard deviation of $\hat{\beta}$; $\overline{\hat{\sigma}}$=Average of estimated standard error; Std-B = $E\left\{|\hat{\beta}-\beta|/\hat{\sigma}\right\}$; MSE=Mean squared error.

Table 3

Fit a proportional subdistribution hazards model.

	Unadjusted weight	**Cox model adjusted weight**
Variable	$\hat{\beta}$; exp(β) (95% CI); *P*	$\hat{\beta}$; exp(β) (95% CI); *P*
RELAPSE		
GP	0.38; 1.47(1.16–1.86); 0.0017	0.54; 1.71(1.34–2.20); < 0.0001
DNR	0.39; 1.48(1.18–1.86); 0.0007	0.35; 1.42(1.13–1.78); 0.0027
PREAUTO	0.41; 1.51(1.19–1.91); 0.0007	0.42; 1.53(1.21–1.93); 0.0004
TRM		
GP	−0.59; 0.55(0.42–0.73); < 0.0001	−0.56; 0.57(0.43–0.75); < 0.0001
DNR	0.57; 1.76(1.38–2.25); < 0.0001	0.55; 1.73(1.35–2.20); < 0.0001
PREAUTO	−0.38; 0.68(0.51–0.91); 0.0099	−0.37; 0.69(0.52–0.92); 0.0117

Table 4

Predicted CIF of relapse and TRM for a patient who received an HLA-identical sibling donor and allogeneic along transplantation

	Unadjusted Weight			**Cox model adjusted Weight**		
	1995–2000	**2001–2005**		**1995–2000**	**2001–2005**	
Time	$\hat{F}_1$ **(95% CI)**	$\hat{F}_2$ **(95% CI)**	$\lvert\hat{F}_1 - \hat{F}_2\rvert$	$\hat{F}_1$ **(95% CI)**	$\hat{F}_2$ **(95% CI)**	$\lvert\hat{F}_1 - \hat{F}_2\rvert$
RELAPSE						
1 Year	0.16 (0.13–0.19)	0.23 (0.18–0.27)	0.07	0.15 (0.13–0.17)	0.24 (0.18–0.30)	0.09
3 Year	0.25 (0.20–0.29)	0.34 (0.28–0.40)	0.09	0.22 (0.20–0.25)	0.35 (0.28–0.42)	0.13
5 Year	0.29 (0.24–0.34)	0.40 (0.33–0.46)	0.11	0.26 (0.24–0.30)	0.41 (0.33–0.49)	0.15
TRM						
1 Year	0.38 (0.32–0.43)	0.23 (0.18–0.28)	0.15	0.37 (0.34–0.41)	0.23 (0.17–0.29)	0.14
3 Year	0.42 (0.37–0.48)	0.26 (0.20–0.32)	0.16	0.42 (0.38–0.46)	0.27 (0.20–0.33)	0.15
5 Year	0.44 (0.38–0.49)	0.27 (0.21–0.33)	0.17	0.43 (0.39–0.47)	0.27 (0.21–0.34)	0.16