Published in final edited form as:

Biometrics. 2008 December ; 64(4): 1090–1099. doi:10.1111/j.1541-0420.2008.00986.x.

Regression Survival Analysis with an Assumed Copula for Dependent Censoring: A Sensitivity Analysis Approach

Xuelin Huang^{*} and

Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Unit 447, Houston, TX 77030, U.S.A

Nan Zhang

Department of Statistics, Rice University, Houston, TX 77005, U.S.A

Summary

In clinical studies, when censoring is caused by competing risks or patient withdrawal, there is always a concern about the validity of treatment effect estimates that are obtained under the assumption of independent censoring. Since dependent censoring is non-identifiable without additional information, the best we can do is a sensitivity analysis to assess the changes of parameter estimates under different degrees of assumed dependent censoring. Such an analysis is especially useful when knowledge about the degree of dependent censoring is available through literature review or expert opinions. In a regression analysis setting, the consequences of falsely assuming independent censoring on parameter estimates are not clear. Neither the direction nor the magnitude of the potential bias can be easily predicted. We provide an approach to do sensitivity analysis for the Cox proportional hazards models. The joint distribution of the failure and censoring times is assumed to be a function of their marginal distributions. This function is called a copula. Under this constraint, we propose an iteration algorithm to estimate the regression parameters and marginal survival functions. Simulation studies show that this algorithm works well. We apply the proposed sensitivity analysis approach to data from an AIDS clinical trial in which 27% of the patients withdrew due to toxicity or at the request of the patient or investigator.

Keywords

Competing risks; Dependent censoring; Frank copula; Kendall's tau; Survival analysis

1. INTRODUCTION

In ACTG 175, a trial conducted by the AIDS Clinical Trial Group, 23% of the total 2467 patients failed, while 27% of the patients went off treatment due to toxicity or at the request of the patient or investigator (Hammer et al., 1996). In such a study, the validity of the data analysis results obtained under the assumption of independent censoring is of concern.

Generally speaking, independent censoring is a key assumption that is usually made in the analysis of time-to-event data. Although this assumption may hold for administrative

^{*}xlhuang@mdanderson.org.

censoring, in many cases it does not hold for censoring due to competing risks or patient drop-out. The dilemma is that, without additional data or assumptions, dependent censoring is not identifiable (Tsiatis, 1975). In such a case, a sensitivity analysis would be helpful. A sensitivity analysis can assess the effect of dependent censoring on the parameter estimates, and can answer questions such as those which follow. Suppose the estimated treatment effect is significantly beneficial under the assumption of independent censoring. Will this effect disappear if dependent censoring is taken into account? Assuming a reasonable range of degrees of dependent censoring, what will be the corresponding range of the treatment effect estimate?

Under a falsely assumed assumption of independent censoring, the effect of dependent censoring on parameter estimates depends on many factors. First, clearly it depends on the degree and direction of correlation between failure and censoring. A positive correlation between failure and censoring will result in overestimation of the survival function, while a negative correlation will result in its underestimation. Second, the magnitude of the resulting bias depends on the fraction of subjects being censored. Third, in a regression analysis, the imbalance of censoring percentages between different covariate groups (e.g., treatment vs. control) has a great influence on the regression parameter estimates. When censoring percentages are balanced between the treated and control groups, the effects of dependent censoring on them may be canceled. In such a scenario, even with dependent censoring, regression parameters may still be roughly unbiased. That is to say, the existence of dependent censoring does not necessarily invalidate all the parameter estimates. On the other hand, however, if there is a severe imbalance between the censoring fractions between the treatment and control groups, and censoring is correlated with failure, then falsely assuming independent censoring can lead to severely biased estimates for the treatment effects.

Many approaches have been proposed to deal with dependent censoring. For example, Peterson (1976) established bounds for a joint distribution function with fixed sub-distribution functions. Slud and Rubinstein (1983) proposed a conditional hazard ratio function to model dependent competing risks. Tawn (1988) studied bivariate extreme value models. Emoto and Matthews (1990) used bivariate Weibull distributions. Slud (1992) showed the nonparametric identifiability of marginal survival distributions in the presence of dependent competing risks and a prognostic covariate. Zheng and Klein (1995) assumed a copula function for the joint distribution of dependent competing risks, and estimated their marginal distributions non-parametrically under the constraint of the assumed copula. Fine, Jiang and Chappell (2001) proposed a copula approach for jointly modeling the first disease recurrence and death. Since death censors disease recurrence, but recurrence does not censor death, they called them semi-competing risks. Andersen et al. (2005) proposed a class of goodness-of-fit tests for a copula based on bivariate right-censored data. Park, Tian and Wei (2006) proposed a nonparametric inference procedure to identify the collection of all attainable values of the parameter of interest. Many authors used auxiliary information to adjust informative censoring in the estimation of the failure time distributions (e.g., Scharfstein et al, 2001; Scharfstein and Robins, 2002).

Although stimulating and useful, the methods listed above did not consider regression models for the primary outcomes. In this study, we extended the copula approach of Zheng

and Klein (1995) to the setting of regression models. We chose to work on the most widely used regression model in survival analysis, namely the Cox (1972) proportional hazards model. Zheng and Klein's method is a good sensitivity analysis tool for the Kaplan-Meier (1958) estimator. It is easy to recognize the importance of extending their method to the regression setting, particularly to the Cox proportional hazard models. However, before completing this study, it was not clear how such an extension could be accomplished. The latest work in this area was by Peng and Fine (2007), who commented that the identifiability of bivariate proportional hazards model had been established for competing risks, but had not been translated into a practical estimation procedure. Their proposal was in the regression setting, but focused on semi-competing risks data. Such data are easier to handle since one of the risks is not subject to censoring by the other, and thus can be analyzed by any of the current methods for univariate survival data. It is not clear whether the estimation method of Peng and Fine could be extended to the competing risks framework. Another recent study on sensitivity analysis in survival analysis by Wei, Tian and Park (2006), was also not for the regression setting. These authors commented that it would be challenging to extend their proposals to general problems because of difficulties in identifying values of the regression parameters with dependent censoring under a nonparametric setting.

In this article, we propose an estimation method for the bivariate proportional hazards model for competing risks. In this model, marginally, each one of the dependent competing risks under study is modeled by a Cox proportional hazards model. Their dependence structure is modeled by an assumed copula function. Using Efron's (1967) ideas of the "redistribution of mass to the right" and self-consistent estimating equations, we propose an extension of the partial likelihood function of Cox (1975). Then we estimate the regression parameters and marginal survival functions by maximizing this extended partial likelihood function. This is done through an iteration method that is similar to the EM algorithm (Dempster, Laird, and Rubin, 1977). The parameter in the copula function that specifies the degree of dependence is assumed to be known. By varying this parameter over a reasonable range, we can perform a sensitivity analysis for the Cox proportional hazards model with respect to the independent censoring assumption. When knowledge is available from experts or the literature about the degree and direction of the correlation between failure and censoring, our approach can be used to obtain less biased parameter estimates. When this knowledge is not available, our approach would still be helpful for assessing the boundary of the parameter estimates by assuming the minimum or maximum of correlation, as was done by Peterson (1976) and Park, Tian and Wei (2006).

While many approaches dealing with dependent censoring can be used to do sensitivity analysis, the copula approach is particularly appealing. By this approach, the marginal distributions of failure and censoring need not be specified, and can be estimated nonparametrically. Such flexibility is desirable. In addition, many copula functions have been proposed and their properties have been well-studied (Nelsen, 1999); therefore, another advantage of the copula approach is its convenience. When we look for a joint distribution function with some specified properties, it is relatively easy to find copula functions that satisfy our needs. Moreover, there is a unified framework for comparing different copula functions. That is, the degrees of association between failure and censoring specified by

them can be directly compared through some commonly used measures, such as Kendall's tau (details provided in Section 2.1).

There is rich literature on sensitivity analysis. Rosenbaum has done a lot of constructive work in this area (e.g., Rosenbaum and Krieger, 1999), although most of his work was not in the setting of survival analysis. Troxel, Ma and Heitjan (2004) proposed an index of local sensitivity to nonignorability. Zhang and Heitjan (2006), Siannis, Copas and Lu (2005), and Siannis (2004) used parametric survival models to do sensitivity analyses. Rotnitzky, Robins and Scharfstein (1998), and Scharfstein, Rotnitzky, and Robins (1999) proposed models for non-ignorable nonresponse and drop-out to do sensitivity analyses for the estimation of the mean of a longitudinal outcome. None of these methods can be used to do sensitivity analysis for the most widely used Cox proportional hazards models; however, our proposed method does just that.

We present the proposed method in section 2, and its evaluation through simulation studies in Section 3. In Section 4, we apply our method to data from an AIDS clinical trial, and close with a summary and discussion in Section 5.

2. METHOD

2.1 Examples of copula functions

A two-dimensional copula is a bivariate distribution function on the square $[0, 1] \times [0, 1]$ with uniform one-dimensional marginal distributions. The degree of association specified by a copula can be conveniently measured by Kendall's τ . Here are some examples of commonly used copulas. Let $u, v \in [0, 1]$ and α be a parameter.

1. Clayton (1978) copula

$$H(u, v; \alpha) = u + v - 1 + \left\{ (1-u)^{-\frac{1}{\alpha}} + (1-v)^{-\frac{1}{\alpha}} - 1 \right\}^{-\alpha}, \quad \alpha > 0. \quad (1)$$

$$\tau = \frac{1}{1+2\alpha}.$$

2. Gumbel-Hougaard copula (Gumbel, 1961; Hougaard, 1986)

$$H(u, v; \alpha) = \exp \left[-\{(-\log u)^\alpha + (\log v)^\alpha\}^{1/\alpha} \right], \quad \alpha \geq 1. \quad (2)$$

$$\tau = 1 - \frac{1}{\alpha}.$$

3. Frank (1979) copula

$$H(u, v; \alpha) = \log_\alpha \left\{ 1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1} \right\}, \quad \alpha > 0, \quad \alpha \neq 1. \quad (3)$$

$$\tau = 1 + 4\gamma^{-1} \{D_1(\gamma) - 1\}, \quad \text{where } \gamma = -\log \alpha, \text{ and,}$$

$$D_1(\gamma) = \frac{1}{\gamma} \int_0^\gamma \frac{t}{e^t - 1} dt.$$

A good introduction to copula functions was given by Nelsen (1999). The simplest copula function not mentioned above is the independence copula $H(u, v) = uv$.

2.2 Failure and censoring time models

Denote the failure time by T , informative censoring time by C , and administrative censoring time by A . We will refer to the informative censoring time as simply the censoring time, while reference to the administrative censoring time will always be explicit. The observed survival time is denoted as $X = \min(T, C, A)$. We also observe indicators $\delta = I(X = T)$ and $\eta = I(X = C)$. The covariates Z and W are associated with failure and censoring, with dimensions $p \times 1$ and $q \times 1$ respectively. They may be identical, overlapped, or completely distinct. Suppose there are subjects, $i = 1, \dots, n$.

The marginal hazard functions for T_i and C_i are respectively assumed to be

$$\lambda_i(t|Z_i, W_i) = \lambda_0(t) \exp(Z_i' \beta), \quad (4)$$

$$\psi_i(t|Z_i, W_i) = \psi_0(t) \exp(W_i' \gamma), \quad (5)$$

where $\lambda_0(t)$ and $\psi_0(t)$ are unspecified baseline hazard functions, and β and γ are unknown parameters with respective dimensions $p \times 1$ and $q \times 1$. Denote their cumulative hazard functions by $\Lambda_0(t)$ and $\Psi_0(t)$ respectively. Then the marginal cumulative distribution functions for T_i and C_i are respectively listed below.

$$\begin{aligned} F_i(t|Z_i, W_i) &= 1 - S_i(t|Z_i, W_i) \\ &= 1 - \exp\{-\Lambda_0(t) \exp(Z_i' \beta)\}, \end{aligned} \quad (6)$$

$$\begin{aligned} G_i(t|Z_i, W_i) &= 1 - R_i(t|Z_i, W_i) \\ &= 1 - \exp\{-\Psi_0(t) \exp(W_i' \gamma)\} \end{aligned} \quad (7)$$

Hereafter, we suppress the covariates in the notation and use $F_i(t)$, $S_i(t)$, $G_i(t)$, and $R_i(t)$ to denote the above functions. Suppose $H(u, v; \alpha)$ is a copula, i.e., a cumulative distribution function defined on the square $[0, 1] \times [0, 1]$, where α is a known parameter. Then the joint cumulative distribution function of T_i and C_i is assumed to be,

$$\begin{aligned} J_i(t, c) &= \Pr(T_i \leq t, C_i \leq c) \\ &= H\{F_i(t), G_i(c); \alpha\} \end{aligned} \quad (8)$$

Heckman and Honore (1989) have shown the identifiability of the parameters in the above competing risks model. However, although estimation methods for some special cases have been proposed, a general estimation method is not yet available in the literature.

2.3 How to fit the joint model

Our method of dealing with the dependent censoring problem extends from the idea of “redistribution of mass” by Efron (1967). Assume each subject has unit mass before its failure, and zero mass afterward. Sort all the event time points (failure or censoring) in ascending order, with the smallest on the far left, and the largest on the far right. Under the

assumption of independent censoring, the mass of a censored subject is redistributed uniformly to all the event time points on its right. After this procedure is done for all censored subjects from left to right, the mass is distributed on failure time points only. The result is the Kaplan-Meier estimator. Zheng and Klein (1994) assumed a copula function to model dependent competing risks. They applied this idea of “redistribution of mass” to obtain self-consistent estimators for the marginal distribution functions. Under the constraint of a copula function for the joint distribution, the redistribution of mass to the right is no longer uniform. Here we further extend Zheng and Klein’s method to the Cox proportional hazards model. For a censored subject, we show below how its mass is redistributed to the right.

Assume that $x_i, i = 1, \dots, n$, are sorted time points in ascending order without ties. If the subject i is censored at time x_i , then for each time point $x_j > x_i$, we would like to compute the probability that this subject i fails at time x_j . Denote this probability by $D_i(x_j)$. Under the joint distribution assumption specified by the copula H , the joint density of T and C is as follows,

$$f_{TC}(t, c) = \frac{\partial^2 H(u, v)}{\partial u \partial v} \Big|_{u=F(t), v=G(c)} \cdot f(t)g(c).$$

Then, with some calculus (Zheng and Klein, 1994), we have,

$$\Pr\{T \geq x, C \in [x, x+dx]\} = g(x) dx \left\{ 1 - \frac{\partial H(u, v)}{\partial v} \Big|_{u=F(x), v=G(x)} \right\}$$

Denote $H_v(a, b) = \frac{\partial H(u, v)}{\partial v} \Big|_{(u, v)=(a, b)}$. Then we have the following conditional survival function, for $t > x_i$,

$$\begin{aligned} \Pr(T_i \geq t | T_i \geq x_i, C_i = x_i) &= \frac{\Pr(T_i \geq t, C_i = x_i)}{\Pr(T_i \geq x_i, C_i = x_i)} \\ &= \frac{1 - H_v\{F_i(t), G_i(x_i)\}}{1 - H_v\{F_i(x_i), G_i(x_i)\}}, \end{aligned} \quad (9)$$

Denote this probability by $P_i(t)$. Then, the mass that subject i loses at time x_j is,

$$D_i(x_j) = P_i(x_{j-1}) - P_i(x_j). \quad (10)$$

Similarly, all the other subjects censored before time x_j lose some mass at the time point x_j . Mimicking the Cox partial likelihood, we define an extended partial likelihood function as follows,

$$L_j^{(T)}(\beta) = \prod_{i=1}^j \left\{ \frac{P_i(x_j) \exp(Z_i' \beta)}{\sum_{k=1}^n P_k(x_j) \exp(Z_k' \beta)} \right\}^{D_i(x_j)}, \quad (11)$$

$$\begin{aligned} L^{(T)}(\beta) &= \prod_{j=1}^n L_j^{(T)}(\beta) \\ &= \prod_{j=1}^n \prod_{i=1}^j \left\{ \frac{P_i(x_j) \exp(Z_i' \beta)}{\sum_{k=1}^n P_k(x_j) \exp(Z_k' \beta)} \right\}^{D_i(x_j)}. \end{aligned} \quad (12)$$

Here $L_j^{(T)}(\beta)$ is the likelihood function for the time point x_j . To make the above equation well-defined, we need to set $P_k(x_j) = 1$ for $k \leq j$. For a failed subject i , set $P_i(x_j) = 1$ for $j \leq x_i$, and $P_i(x_j) = 0$ for $j > x_i$. However, instead of using (10), we set $D_i(x_i) = 1$, $D_i(x_j) = 0$ for $j > x_i$. Then a failed subject contributes only one term in this extended partial likelihood function. When there are tied failure events, the above equations naturally handle them by Breslow's (1974) method. Actually, if we view the pieces of mass $D_i(x_j)$, $i = 1, \dots, j$, as ties at time x_j , then the formation of $L_j^{(T)}(\beta)$ in (11) is obtained from Breslow's method.

In addition to the above extended partial likelihood for the failure events, we also need to give the counterpart for the censoring events. We can best understand the manipulation of failure and censoring by putting them in the setting of competing risks. If subject i fails at time x_i , similarly, for $c > x_i$, we have

$$\Pr(C_i \geq c | C_i \geq x_i, T_i = x_i) = \frac{1 - H_u\{F_i(x_i), G_i(c)\}}{1 - H_u\{F_i(x_i), G_i(x_i)\}}, \quad (13)$$

where $H_u(a, b) = \frac{\partial H(u, v)}{\partial u} \big|_{(u, v) = (a, b)}$. Denote this probability by $Q_i(c)$, and also set $Q_i(c) = 1$ for $c \leq x_i$. Similarly we define that

$$U_i(x_j) = Q_i(x_{j-1}) - Q_i(x_j). \quad (14)$$

The extended partial likelihood function for censoring events is as shown below,

$$L^{(C)}(\gamma) = \prod_{j=1}^n \prod_{i=1}^j \left\{ \frac{Q_i(x_j) \exp(W_i' \gamma)}{\sum_{k=1}^n Q_k(x_j) \exp(W_k' \gamma)} \right\}^{U_i(x_j)}, \quad (15)$$

Then the parameters β and γ can be estimated by maximizing the following extended joint partial likelihood,

$$L(\beta, \gamma) = L^{(T)}(\beta) L^{(C)}(\gamma). \quad (16)$$

The likelihood $L^{(T)}(\beta)$ in (12) depends on not only parameter β , it also implicitly depends on parameter γ through the functions $P_k(x_j)$ and $D_i(x_j)$, so it is better written as $L^{(T)}(\beta, \gamma)$. Similarly, the likelihood $L^{(C)}(\gamma)$ in (15) is better written as $L^{(C)}(\beta, \gamma)$.

We need some additional set-ups for the above extended likelihood functions to be well defined. First, set $Q_k(x_j) = 1$ for $k = j$. For a censored subject i , set $U_i(x_i) = 1$ and $U_i(x_j) = 0$ for $j > i$. For an administratively censored subject i , set $P_i(x_j) = Q_i(x_j) = 1$ for $j \leq i$, $P_i(x_j) = Q_i(x_j) = 0$ for $j > i$, $D_i(x_j) = U_i(x_j) = 0$ for all j .

Since the above functions involve unknown quantities such as $P_i(x_j)$, etc, we iterate, as described below.

1. Assuming independent censoring, fit two Cox proportional hazards models to get initial estimators $\hat{\beta}^{(0)}$ and $\hat{\gamma}^{(0)}$ for β and γ , respectively. Then use the Breslow (1972) method to obtain estimators $\hat{\Lambda}_0^{(0)}(\cdot)$ for $\Lambda_0(\cdot)$, and $\hat{\Psi}_0^{(0)}(\cdot)$ for $\Psi_0(\cdot)$. let $m = 0$.
2. For $i = 1, \dots, n$, compute

$$\begin{aligned}\hat{S}_i^{(m)}(t) &= \exp\{-\hat{\Lambda}_0^{(m)}(t) \exp(Z_i' \hat{\beta}^{(m)})\}, \\ \hat{R}_i^{(m)}(t) &= \exp\{-\hat{\Psi}_0^{(m)}(t) \exp(W_i' \hat{\gamma}^{(m)})\}.\end{aligned}$$

Then for each time point x_j such that $x_j > x_i$, compute

$$\hat{P}_i^{(m)}(x_j) = \frac{1 - H_v\{\hat{F}_i^{(m)}(x_j), \hat{G}_i^{(m)}(x_j); \alpha\}}{1 - H_v\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_i); \alpha\}}, \quad (17)$$

$$\hat{Q}_i^{(m)}(x_j) = \frac{1 - H_u\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_j); \alpha\}}{1 - H_u\{\hat{F}_i^{(m)}(x_i), \hat{G}_i^{(m)}(x_i); \alpha\}}. \quad (18)$$

Then we can get $\hat{D}_i^{(m)}(x_j)$ and $\hat{U}_i^{(m)}(x_j)$. All these computations are done for all $1 \leq i < j \leq n$. For $1 \leq j \leq i \leq n$, set $\hat{P}_i^{(m)}(x_j) = 1$, and $\hat{Q}_i^{(m)}(x_j) = 1$.

3. Replace unknown functions P_i , Q_i , D_i , U_i in $L^{(T)}(\beta)$ and $L^{(C)}(\gamma)$ by their estimates at step m , and then maximize the likelihood functions in (12) and (15) with respect to β and γ . Denote the resulting estimators for β and γ by $\hat{\beta}^{(m+1)}$ and $\hat{\gamma}^{(m+1)}$.
4. Using $\hat{\beta}^{(m+1)}$, $\hat{\gamma}^{(m+1)}$, $\hat{P}_i^{(m)}(\cdot)$, $\hat{Q}_i^{(m)}(\cdot)$, $\hat{D}_i^{(m)}(\cdot)$, and $\hat{U}_i^{(m)}(\cdot)$ obtain Breslow estimators $\hat{\Lambda}_0^{(m+1)}(\cdot)$ for $\Lambda_0(\cdot)$, and $\hat{\Psi}_0^{(m+1)}(\cdot)$ for $\Psi_0(\cdot)$, as shown below.

$$\hat{\Lambda}_0^{(m+1)}(t) = \sum_{j: x_j \leq t} \frac{\sum_{i: x_i \leq x_j} \hat{D}_i^{(m)}(x_j)}{\sum_{k=1}^n \hat{P}_k^{(m)}(x_j) \exp(Z_k' \hat{\beta}^{(m+1)})}, \quad (19)$$

$$\hat{\Psi}_0^{(m+1)}(t) = \sum_{j: x_j \leq t} \frac{\sum_{i: x_i \leq x_j} \hat{U}_i^{(m)}(x_j)}{\sum_{k=1}^n \hat{Q}_k^{(m)}(x_j) \exp(W_k' \hat{\gamma}^{(m+1)})}. \quad (20)$$

5. Let $m = m + 1$, return to step 2, and iterate until convergence.

After convergence, we get estimators $\hat{\beta}$, $\hat{\gamma}$, $\hat{\Lambda}_0(\cdot)$ and $\hat{\Psi}_0(\cdot)$, respectively, for β , γ , $\Lambda_0(\cdot)$ and $\Psi_0(\cdot)$. Denote these sets of estimators and parameters by $\hat{\theta}$ and θ , respectively. Note that $D_i(x_j)$, $U_i(x_j)$, $P_i(x_j)$ and $Q_i(x_j)$ are all functions of θ . Denote the estimator of θ at the m^{th} iteration by $\hat{\theta}^{(m)}$. Then at convergence, $\hat{\theta}$ satisfies the following equations:

$$\sum_{j=1}^n \sum_{i=1}^j \hat{D}_i(x_j; \hat{\theta}) \left\{ Z_i - \frac{\sum_{k=1}^n \hat{P}_k(x_j; \hat{\theta}) Z_k \exp(Z_k' \hat{\beta})}{\sum_{k=1}^n \hat{P}_k(x_j; \hat{\theta}) \exp(Z_k' \hat{\beta})} \right\} = 0 \quad (21)$$

$$\sum_{j=1}^n \sum_{i=1}^j \hat{U}_i(x_j; \hat{\theta}) \left\{ W_i - \frac{\sum_{k=1}^n \hat{Q}_k(x_j; \hat{\theta}) W_k \exp(W_k' \hat{\gamma})}{\sum_{k=1}^n \hat{Q}_k(x_j; \hat{\theta}) \exp(W_k' \hat{\gamma})} \right\} = 0 \quad (22)$$

$$\hat{\Lambda}_0(t) = \sum_{j: x_j \leq t} \frac{\sum_{i: x_i \leq x_j} \hat{D}_i(x_j; \hat{\theta})}{\sum_{k=1}^n \hat{P}_k(x_j; \hat{\theta}) \exp(Z_k' \hat{\beta})}, \quad (23)$$

$$\hat{\Psi}_0(t) = \sum_{j: x_j \leq t} \frac{\sum_{i: x_i \leq x_j} \hat{U}_i(x_j; \hat{\theta})}{\sum_{k=1}^n \hat{Q}_k(x_j; \hat{\theta}) \exp(W_k' \hat{\gamma})}. \quad (24)$$

These equations for θ are self-consistent as proposed by Efron (1967). There are many applications of his idea in different scenarios. For example, Turnbull (1974, 1976) used this idea to estimate the survival distribution from interval-censored data and other complicated types of data. Laird (1978) applied this idea to obtain a nonparametric maximum likelihood estimator of a mixed distribution. Tsai and Crowley (1985) discussed the theoretical properties of self-consistent estimators in general non-regression settings. They showed: (1) the guaranteed convergence of the above iteration algorithm and its connection with the EM algorithm (Dempster, Laird and Rubin, 1977), (2) that such a self-consistent estimator is actually a generalized maximum likelihood estimator in the sense of Kiefer and Wolfowitz (1956), (3) the strong consistency of the self-consistent estimators, and (4) its weak convergence to a Gaussian process. Although these results and the simulation studies in the next section indicate potentially good large sample properties of our estimator in the regression setting, further theoretical investigation is still needed.

The covariance matrices of the above estimators can be obtained by the bootstrap method (Efron, 1979). By the above algorithm, it can be seen that the final survival estimator $\hat{S}_0(t)$ has jumps at failure time points only. Starting at $m = 0$, since $\hat{\Lambda}_0^{(m)}(t)$ is flat at a censoring time point x , so in turn are $\hat{S}_i^{(m)}(t)$, $\hat{P}_i^{(m)}(t)$ and $\hat{D}_i^{(m)}(t)$ for all $i = 1, \dots, n$. This results in

$\hat{\Lambda}_0^{(m+1)}(t)$ also being flat at the censoring time point x . Consequently, by iteration, the final survival estimator is flat at censoring time points.

One may try to maximize the nonparametric likelihood function to estimate parameters. To do that, we need to assume the failure and censoring distributions are discrete and have masses on their respective event points only. One way to parameterize it is to assume discrete failure hazards $\lambda_1, \dots, \lambda_n$, and censoring hazards $\psi_1, \dots, \psi_n$, with them being positive only on those failure and censoring event time points, respectively (otherwise being zero). With the copula function assumption, these parameters completely specify the joint distribution (if not considering covariates). Considering covariates, for a failure event at time x_i , its contribution to the joint likelihood is shown below.

$$\lambda_i \exp(Z_i' \beta) \exp \left\{ - \sum_{j=1}^i \lambda_j \exp(Z_j' \beta) \right\} [1 - H_u\{F_i(x_i), G_i(x_i)\}],$$

where,

$$F_i(x_i) = 1 - \Pr(T_i \geq x_i) = 1 - \exp \left\{ - \sum_{j=i+1}^n \lambda_j \exp(Z_j' \beta) \right\},$$

$$G_i(x_i) = 1 - \Pr(C_i > x_i) = 1 - \exp \left\{ - \sum_{j=i+1}^n \psi_j \exp(W_j' \gamma) \right\}.$$

Similarly, we can determine the contribution of a censoring event to the joint likelihood function. However, choosing any copula function described in this article, and using it in the above likelihood, it can be seen that the copula function itself is already very complicated, and its partial derivative functions are even more complicated. For example, for the Frank copula,

$$\frac{\partial H(u, v; \alpha)}{\partial u} = \frac{\alpha^u (\alpha^v - 1)}{\alpha - 1 + (\alpha^u - 1)(\alpha^v - 1)}.$$

After substituting u and v by $F_i(x)$ and $G_i(x)$, respectively, in the above equation, we can see that it will be extremely difficult, if possible, to directly maximize the joint likelihood function.

3. SIMULATION STUDY

We conducted simulation studies to evaluate the proposed estimation method. We generated 500 data sets by using each of the Frank, Clayton and Gumbel-Hougaard copulas, all with Kendall's $\tau = 0.8$. Then we analyzed the data sets by assuming, respectively, (1) a Frank copula with Kendall's $\tau = 0.8$, and (2) independent censoring. For the 500 data sets generated using the Frank copula, we also analyzed them by assuming a Frank copula with

Kendall's $\tau = 0.5$ and 0.2 . That is to say, we used simulation studies to assess the proposed method in different situations: (1) both the form of the copula and the degree of association (measured by Kendall's τ) are correctly specified, (2) the copula function is mis-specified, and (3) the degree of association is mis-specified.

We use the following model to generate data. There are two covariates: $Tr = 0$ or 1 with equal probability and centered $Age \sim \text{Uniform}(-10, 10)$. We specified the marginal distributions for failure and censoring times T and C , respectively, by the following hazard functions:

$$\lambda(t) = 0.5 t \exp(-0.5 Tr + 0.1 \text{ Age}), \quad (25)$$

$$\psi(c) = 0.2 \exp(0.3 Tr + 0.2 \text{ Age}). \quad (26)$$

We used different copulas to specify the joint distribution of T and C . To achieve the same Kendall's $\tau = 0.8$, we needed to use $\alpha = 1.258 \times 10^{-8}$ for the Frank copula, $\alpha = 0.125$ for the Clayton copula, and $\alpha = 5.0$ for the Gumbel-Hougaard copula.

Nelsen (1986) gave the conditional distribution of the Frank copula. It leads to the following algorithm for implementing the above data generation.

1. Generate two independent $\text{Uniform}(0, 1)$ random variables u and v' .
2. Let

$$v = \log_{\alpha} \left\{ 1 + \frac{v'(\alpha - 1)}{v' + (1 - v')\alpha^u} \right\}$$

Then (u, v) has a joint distribution specified by the Frank copula in (3). The generation of data for the Clayton and Gumbel-Hougaard copulas were given by Clayton (1978) and Marshall and Olkin (1988), respectively. Details are omitted here. After generating (u, v) , we let $u = F(t) = 1 - \exp\{-\Lambda(t)\}$ and $v = G(c) = 1 - \exp\{-\Psi(c)\}$, and solved them for t and c . Then we assigned the solutions to failure time T and C , respectively. In addition, we generated another random time variable A from the $\text{Uniform}(0, 10)$ distribution independent of everything else. This is the administrative censoring time. The observed survival time is $X = \min(T, C, A)$. We also have indicators $\delta = I(X = T)$ and $\eta = I(X = C)$. We generated 500 data sets in this way for each copula, with 200 observations in each data set. On average, about 45% of the observations were failures, 35% were withdrawals, and the rest were administratively censored.

We fitted the generated data sets according to the proposed method described in Section 2. We used 200 bootstrap replicates for each estimation of standard errors. The results are summarized in Table 1. The parameter α , which specifies the association parameter Kendall's τ , is not estimated. Instead, it is assumed to be known. From Table 1, first, it can be seen that, when using the correctly specified copula and Kendall's τ , all the estimates for the regression parameters are very close to the true values. The means of the estimated

standard errors match well with their empirical counterparts. The coverage probabilities of the estimated 95% confidence intervals are close to the nominal level. These results show that the proposed method works well for the estimation of the regression parameters. Second, when the dependence structure (i.e., the copula) is mis-specified, but the degree of association (Kendall's τ) is correctly specified, the estimation has some slight bias. Third, when the degree of association is mis-specified, the estimates for regression parameters are severely biased. The degree of bias increases with the deviation of the assumed Kendall's τ from the truth. From these simulation studies, it can be seen that falsely assuming independent censoring can lead to much more severe bias than choosing a wrong copula.

Conditional on covariate values $Tr = 0$ and centered $Age = 0$, the true survival functions for failure time T is $S_0(t) = \exp(-0.25t^2)$. The mean of its estimates is shown in Figure 1(a). By the proposed method, the estimated survival curve is almost indistinguishable from the truth. However, under the falsely assumed independent censoring, the estimated survival curve is far from the truth. From Figure 1(b), we can see that the mean of the estimated standard errors of the proposed estimator is very close to the empirical standard deviation obtained from the 500 simulations. The estimation for the distribution of censoring time C gives similar results (not shown).

To check the effect of sample size and censoring fraction on the performance of the proposed method, we changed the failure time hazard function in (25) to $\lambda(t) = 2t \exp(-0.5 Tr + 0.1 Age)$ and changed the distribution of the administrative censoring time A to $Uniform(0, 3)$. The hazard function for the censoring time C was unchanged. We used a Frank copula with $\tau = 0.8$ in the data generation. The resulting failure and censoring percentages are 50% and 15%, respectively. The results for sample sizes $n = 200$ and $n = 400$ are reported in Table 2. Together with the results in the top panel of Table 1, we can see that (1) the proposed estimation procedure works well for these different sample sizes and censoring fractions; (2) When falsely assuming independent censoring, the higher percentage of censoring, the more severe bias; and (3) a larger sample size gives more accurate estimation results, as reflected by smaller standard errors.

When the independence copula is used, the extended partial likelihood does not algebraically reduce to the Cox partial likelihood. However, simulations showed that they give very similar results (see Table 2).

The proposed estimation procedure is computationally intensive. One replicate of the simulation in Table 1 (with sample size $n = 200$ and 200 bootstrap samples in the estimation of standard errors) takes about 15 minutes on a Sun Solaris server.

4. Application

We applied the proposed method to data from a double-blind clinical trial (ACTG 175) conducted by the AIDS Clinical Trials Group. In that trial, 2467 HIV-1-infected patients with CD4 counts between 200 and 500/mm³ were randomized to one of four daily antiretroviral regimens (Hammer et al., 1996). For simplicity, we describe the subset of 1132 patients who received either AZT monotherapy or AZT plus didanosine combination therapy. The primary end point was a 50% decline in the CD4 count, development of

AIDS, or death, which we call failure. In the chosen subset of patients, 309 patients experienced failure. Another 327 patients went off treatment before experiencing failure due to toxicity or at the request of the patient or investigator. The primary outcomes for those patients were censored by withdrawal. The primary outcomes of the remaining 796 patients were administratively censored. Viewing all types of censoring as independent of failure, Hammer et al. (1996) concluded that the AZT monotherapy was inferior to each of the other three treatments. While administrative censoring can be plausibly assumed to be independent of failure, withdrawal is most likely to be positively associated with failure. It is of interest to perform a sensitivity analysis for this important data set.

We first analyzed this subset of data with Cox proportional hazard models, and then used the method proposed in Section 2 to do sensitivity analysis. We selected covariates related to failure or withdrawal using a stepwise selection algorithm assuming independent censoring. We set a p -value of 0.15 as the threshold for both variable entry and remaining significant. That choice was based on a consideration of the potential change in p -values for covariates after accounting for dependent censoring. A total of nine variables were selected. They are listed in Table 3. The distributions of the numbers of failures, withdrawals, and administrative censoring among different covariate groups are also listed in Table 3. By dividing patient age into groups of 10 years, it can be seen that groups with ages younger than 50 years had very similar failure risks, whereas groups older than 50 years had significantly higher failure risks. Thus, we divided patients into two age groups: whether older than 50 years or not.

To perform sensitivity analysis, we included all nine selected variables in both the marginal models for failure and withdrawal. A Frank copula was assumed for the association between failure and withdrawal with Kendall's τ being 0, 0.2, 0.5 or 0.8 respectively. The results under $\tau = 0$ is actually obtained from the Cox models assuming independence between failure and withdrawal. Figure 2 shows changes in the regression coefficients for covariate effects on failure under these different degrees of assumed association. The effects of the following covariates remain significant ($p < 0.05$) as Kendall's τ changes from 0 to 0.8: AZT monotherapy vs. the combination therapy, symptomatic vs. asymptomatic disease status, and higher vs. lower baseline CD4 count. The associated p -values were all less than 0.01 when $\tau = 0, 0.2, 0.5$, or 0.8 (Figure 2(1, 3, 8)). The effect of prior antiretroviral therapy changed from significant ($p = 0.01$ when $\tau = 0$) to non-significant ($p = 0.73$ when $\tau = 0.8$) (Figure 2(7)). The effects of hemophilia and race changed from non-significant ($p = 0.9$ and $p = 0.17$, respectively, when $\tau = 0$) to significant ($p = 0.04$ for both when $\tau = 0.8$) (Figure 2(2,6)). The effects of sex, IV drug use and age remained non-significant. The changes in p -values could be slight or substantial (Figure 2(4, 5, 9)). While it is useful to show all these changes on covariate effects, it is inappropriate to use these figures to "look for" significant results.

From the above example, we can see that it is hard to predict the changes in regression coefficient estimates. After accounting for dependent censoring, their point estimates and the corresponding p -values can change either way. This is different from the estimation of survival functions, for which the direction of bias is easily predictable. Therefore, our method is useful in the sense that it provides a tool for assessing the bias in the estimation of

regression parameters. It can assess the reliability of the results obtained under the assumption of independent censoring. For example, by the sensitivity analysis shown in Figure 2(1), we can see that the superiority of the combination therapy remains true under different degrees of association between failure and withdrawal. Therefore, it is fairly safe to conclude that the combination therapy is superior to the AZT monotherapy. That conclusion was the primary research interest of the ACTG 175 clinical trial. The above sensitivity analysis helps us gain more confidence in the conclusion of Hammer et al. (1996).

We also show the effect of changing α on regression parameter estimates of the model for patient withdrawal, treating death as “censoring” (Figure 3). It can be seen that, under different degrees of assumed association between failure and withdrawal, some trends remain significant (or marginally so) in the same direction. The patient characteristics significantly related to withdrawal from the study were the following: hemophilia, symptomatic disease, female, and lack of prior experience with antiretroviral therapy. The effects of other patient characteristics on withdrawal may change from non-significant to significant, or the other way around. However, none changes its direction (i.e., from positive to negative or vice versa).

We conducted further sensitivity analyses by assuming the Clayton and Gumbel-Hougaard copulas. The results were similar to those obtained from assuming a Frank copula, as shown in Figures 2 and 3, and are not reported in this article. Interested readers can check the Biometrics web site for supplementary materials.

5. Summary and Discussion

We have proposed a method to do sensitivity analysis with respect to the assumption of independent censoring for the Cox proportional hazards models. In many scenarios, we have a common sense or some prior knowledge about the direction of association between censoring and the event of interest. For example, patient withdrawal due to toxicity is most likely to be positively associated with failure. In such a case, we can predict the direction of bias of the Kaplan-Meier estimators. However, it is much more difficult to predict the direction of the bias of the regression parameter estimates by the Cox proportional hazards models. We have provided a tool for such a task. By varying the degree of association between failure and censoring, we can evaluate the robustness of the results of the Cox models. While some of the parameter estimates might be very sensitive to the assumption of independent censoring, others might not. Our method can help describe the whole picture, as illustrated through the example in Section 4.

An implicit assumption used in our method is that, the association between failure and censoring is the same among different subgroups defined by covariate values. This assumption is plausible in most situations. This assumption would fail if, for example, those patients who withdraw are those sicker patients in group A, and healthier patients in group B. However, such phenomenon is unlikely, especially in a blinded randomized trial.

The copula parameter α is assumed to be known in the application of our method. In reality, it is unknown, but knowledge about its range can usually be available from experts or the literature. If there is not any knowledge available for the association between failure and

censoring, we can still use the proposed method to do a sensitivity analysis in a very conservative way, i.e. letting the Kendall's τ change from -1 to 1 to perform the sensitivity analysis. In some situations, for instance, when censoring percentages are small and/or balanced across covariate groups, the estimated regression parameters and their confidence intervals may not change much. Consequently, the conclusion made under independent censoring assumption may remain valid. In such situations, the sensitivity analyses can help us gain a great deal of confidence in the data analysis results.

Acknowledgments

The research of Xuelin Huang was supported by the National Institute of Health of U.S.A. grant CA100632 and CA108631. The ACTG 175 data set was obtained from the National Technical Information Service, an agency of the Department of Commerce of U.S.A..

References

- Andersen PK, Ekstrom CT, Klein JP, Shu YY, Zhang MJ. A class of goodness of fit tests for a copula based on bivariate right-censored data. *Biometrical Journal*. 2005; 47:815–824. [PubMed: 16450854]
- Breslow N. Discussion of Dr. Cox's paper "Regression models and life tables". *Journal of the Royal Statistical Society, Series B*. 1972; 34:216–217.
- Breslow N. Covariance analysis of censored survival data. *Biometrics*. 1974; 30:89–99. [PubMed: 4813387]
- Clayton DG. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*. 1978; 65:141–151.
- Cox DR. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*. 1972; 34:187–220.
- Cox DR. Partial likelihood. *Biometrika*. 1975; 62:269–276.
- Dempster AP, Laird NM, Rubin DB. Maximum likelihood estimation for incomplete data via the EM algorithms (with discussion). *Journal of the Royal Statistical Society, Series B*. 1977; 39:1–38.
- Efron, B. Proceedings of the 5th Berkeley Symposium. Vol. 4. Berkeley: University of California Press; 1967. The two sample problem with censored data; p. 831–853.
- Efron B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*. 1979; 7:1–26.
- Emoto SE, Matthews PC. A Weibull Model for Dependent Censoring. *The Annals of Statistics*. 1990; 18:1556–1577.
- Fine JP, Jiang H, Chappell R. On semi-competing risks data. *Biometrika*. 2001; 88:907–919.
- Frank MJ. On the simultaneous association of $F(x, y)$ and $x+y - F(x, y)$. *Aequationes Mathematicae*. 1979; 21:37–38.
- Gumbel EJ. Distributions des valeurs extrêmes en plusieurs dimensions. *Publ Inst Statist Univ Paris*. 1961; 9:171–173.
- Hammer SM, Katzenstein DA, Hughes MD, Gundacker H, Schooley RT, Haubrich RH, Henry WK, Lederman MM, Phair JP, Niu M, Hirsch MS, Merigan TC. A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with CD4 cell counts from 200 to 500 per cubic millimeter. *New England Journal of Medicine*. 1996; 335:1081–1090. [PubMed: 8813038]
- Heckman JJ, Honore BE. The identifiability of the competing risks model. *Biometrika*. 1989; 76:325–330.
- Hougaard P. A class of multivariate failure time distributions. *Biometrika*. 1986; 73:671–678.
- Kaplan EL, Meier P. Nonparametric Estimation From Incomplete Observations. *Journal of the American Statistical Association*. 1958; 53:457–481.
- Kiefer J, Wolfowitz J. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of mathematical Statistics*. 1956; 27:887–906.

- Laird N. Nonparametric maximum likelihood estimation of a mixing distribution. *Journal of the American Statistical Association*. 1978; 73:805–811.
- Marshall AW, Olkin I. Families of multivariate distributions. *Journal of the American Statistical Association*. 1988; 83:834–841.
- Nelsen RB. Properties of a one-parameter family of bivariate distributions with specified marginals. *Communications in Statistics, Part A: Theory and Methods*. 1986; 15:3277–3285.
- Nelsen, RB. *An Introduction to Copulas*. New York: Springer; 1999.
- Park Y, Tian L, Wei LJ. One-and two-sample nonparametric inference procedures in the presence of a mixture of independent and dependent censoring. *Biostatistics*. 2006; 7:252–267. [PubMed: 16282531]
- Peng L, Fine JP. Regression Modeling of semicompeting risks data. *Biometrics*. 2007; 63:96–108. [PubMed: 17447934]
- Peterson AV. Bounds for a joint distribution with fixed sub-distribution functions: application to competing risks. *Proceedings of the national academy of Sciences of the United States of America*. 1976; 73:11–13. [PubMed: 16592299]
- Rosenbaum PR, Krieger AM. Sensitivity of two-sample permutation inferences in observational studies. *Journal of the American Statistical Association*. 1990; 85:493–498.
- Rotnitzky A, Robins JM, Scharfstein DO. Semiparametric regression for repeated outcomes with nonignorable nonresponse. *Journal of the American Statistical Association*. 1998; 93:1321–1339.
- Scharfstein DO, Robins JM. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika*. 2002; 89:617–634.
- Scharfstein DO, Robins JM, Eddings W, Rotnitzky A. Inference in randomized studies with informative censoring and discrete time-to-event endpoints. *Biometrics*. 2001; 57:404–413. [PubMed: 11414563]
- Scharfstein DO, Rotnitzky A, Robins JM. Adjusting for nonignorable dropout using semiparametric nonresponse models (with discussion). *Journal of the American Statistical Association*. 1999; 94:1096–1146.
- Siannis F, Copas J, Lu G. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*. 2005; 6:77–91. [PubMed: 15618529]
- Siannis F. Applications of a parametric model for informative censoring. *Biometrics*. 2004; 60:704–714. [PubMed: 15339293]
- Slud, E. Nonparametric identifiability of marginal survival distributions in the presence of dependent competing risks and a prognostic covariate. In: Klein, JP.; Goel, PK., editors. *Survival Analysis: State of the Art*. Kluwer Academic Publishers; Netherland: 1992. p. 355–368.
- Slud EV, Rubinstein LV. Dependent Competing Risks and Summary Survival Curves. *Biometrika*. 1983; 70:643–649.
- Tawn JA. Bivariate Extreme Value Theory: Models and Estimation. *Biometrika*. 1988; 75:397–415.
- Troxel AB, Ma G, Heitjan DF. An index of local sensitivity to nonignorability. *Statistica Sinica*. 2004; 14:1221–1237.
- Tsai WY, Crowley J. A large sample study of the generalized maximum likelihood estimators from incomplete data via self-consistency. *The Annals of Statistics*. 1985; 13:1317–1334.
- Tsiatis A. A Nonidentifiability Aspect of the Problem of Competing Risks. *Proceedings of the National Academy of Science, USA*. 1975; 72:20–22.
- Turnbull BW. Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association*. 1974; 69:169–173.
- Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society, Series B: Methodological*. 1976; 38:290–295.
- Zhang J, Heitjan DF. A simple local sensitivity analysis tool for nonignorable coarsening: Application to dependent censoring. *Biometrics*. 2006; 62:1260–1268. [PubMed: 17156301]
- Zheng M, Klein JP. A Self-Consistent Estimator of Marginal Survival Functions Based on Dependent Competing Risk Data and an Assumed Copula. *Communications in Statistics, Part A—Theory and Methods*. 1994; 23:2299–2311.

Zheng M, Klein JP. Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. *Biometrika*. 1995; 82:127–138.

Figure 1.

(a) Estimators for the survival function of failure time (for subjects with all covariate values being zero). Solid line, the true distribution; dotted line, average over 500 simulations of the proposed estimator for dependent censoring with correctly specified copula and parameter α ; dashed line, average over 500 simulations of the estimator by the Cox proportional model assuming independent censoring. (b) Comparison of the mean of standard error estimates of the proposed estimator as specified in (a) with the empirical standard deviation. Solid line, the empirical sample standard deviation of the proposed estimator over 500 simulations; dotted line, average over 500 simulations of bootstrap standard error estimates (200 bootstrap data sets generated for each estimate of standard error).

Figure 2.

Sensitivity analysis for the Cox model for failure: Assuming a Frank copula for the ACTG 175 trial data, the changes in regression parameter estimates under different degrees of assumed association between patient failure and withdrawal (from independence to Kendall's $\tau=0.8$).

Figure 3. Sensitivity analysis for the Cox model for withdrawal: Assuming a Frank copula for the ACTG 175 trial data, the changes on regression parameter estimates under different degrees of assumed association between patient failure and withdrawal (from independence to Kendall's $\tau = 0.8$).

Table 1

Parameter Estimates by Proposed Method

Truth	Est	Bias	SD	SE	CP	Est	Bias	SD	SE	CP	
Copula used in data generation: Frank with $\tau = 0.8$											
Copula assumed in data analyses: Frank with $\tau = 0.8$											
Independent censoring											
β_{Tr}	-0.5	-0.487	-3%	0.191	0.187	0.942	-0.857	71%	0.241	0.227	0.652
β_{Age}	0.1	0.103	3%	0.0189	0.0176	0.932	0.0454	-55%	0.0244	0.0230	0.336
γ_{Tr}	0.3	0.311	4%	0.285	0.286	0.938	0.534	78%	0.252	0.251	0.850
γ_{Age}	0.2	0.196	-2%	0.0378	0.0358	0.924	0.245	23%	0.0303	0.0302	0.710
Copula used in data generation: Frank with $\tau = 0.8$											
Copula assumed in data analyses: Frank with $\tau = 0.5$											
Frank with $\tau = 0.2$											
β_{Tr}	-0.5	-0.631	-26%	0.201	0.195	0.896	-0.764	53%	0.217	0.211	0.762
β_{Age}	0.1	0.0853	-15%	0.0201	0.0187	0.838	0.0647	-35%	0.0215	0.0200	0.548
γ_{Tr}	0.3	0.406	35%	0.257	0.255	0.926	0.495	65%	0.252	0.251	0.874
γ_{Age}	0.2	0.221	11%	0.0313	0.0317	0.922	0.238	19%	0.0298	0.0313	0.826
Copula used in data generation: Clayton with $\tau = 0.8$											
Copula assumed in data analyses: Frank with $\tau = 0.8$											
Independent censoring											
β_{Tr}	-0.5	-0.487	-3%	0.195	0.185	0.942	-0.836	67%	0.237	0.225	0.662
β_{Age}	0.1	0.099	-1%	0.0156	0.0171	0.962	0.0457	-54%	0.0209	0.0224	0.304
γ_{Tr}	0.3	0.346	15%	0.287	0.286	0.952	0.554	85%	0.261	0.254	0.852
γ_{Age}	0.2	0.190	-5%	0.0319	0.0327	0.932	0.236	18%	0.0298	0.0296	0.822
Copula used in data generation: Gumbel-Hougaard with $\tau = 0.8$											
Copula assumed in data analyses: Frank with $\tau = 0.8$											
Independent censoring											
β_{Tr}	-0.5	-0.526	5%	0.182	0.187	0.950	-0.864	73%	0.225	0.224	0.644
β_{Age}	0.1	0.105	5%	0.0179	0.0179	0.924	0.0523	-48%	0.0228	0.0225	0.416
γ_{Tr}	0.3	0.360	20%	0.298	0.298	0.952	0.576	92%	0.264	0.260	0.820
γ_{Age}	0.2	0.219	10%	0.0374	0.0371	0.914	0.262	31%	0.0316	0.0320	0.548

Est = mean of estimates; SD = empirical standard deviation of estimates; SE = mean of estimated standard errors; CP = coverage probability of 95% confidence interval.

Table 2
Effects of Sample Size and Censoring Fraction and The Analyses using Independence Copula

Data generation and analysis	Sample size	Para-meter	Truth	Est	Bias	SD	SE	CP
Both by Frank copula with $\tau = 0.8$	$n = 200$, 50% T , 15% C , 35% A^*	β_{Tr}	-0.5	-0.488	-2%	0.192	0.188	0.938
		β_{Age}	0.1	0.102	2%	0.0180	0.0176	0.943
		γ_{Tr}	0.3	0.313	4%	0.486	0.518	0.957
		γ_{Age}	0.2	0.190	-5%	0.0602	0.0584	0.920
Both by Frank copula with $\tau = 0.8$	$n = 400$, 50% T , 15% C , 35% A	β_{Tr}	-0.5	-0.477	-5%	0.133	0.130	0.939
		β_{Age}	0.1	0.102	2%	0.0119	0.0118	0.952
		γ_{Tr}	0.3	0.288	4%	0.376	0.358	0.932
		γ_{Age}	0.2	0.193	-4%	0.0513	0.0479	0.934
Both by Independence copula (proposed method)	$n = 200$, 50% T , 35% C , 15% A	β_{Tr}	-0.5	-0.504	-1%	0.215	0.220	0.965
		β_{Age}	0.1	0.099	-1%	0.0232	0.0241	0.963
		γ_{Tr}	0.3	0.306	2%	0.242	0.253	0.964
		γ_{Age}	0.2	0.190	-5%	0.0328	0.0325	0.934
Generated by Independence copula, analyzed by Cox model	$n = 200$, 50% T , 35% C , 15% A	β_{Tr}	-0.5	-0.497	-1%	0.210	0.215	0.952
		β_{Age}	0.1	0.103	3%	0.0237	0.0226	0.938
		γ_{Tr}	0.3	0.304	1%	0.239	0.245	0.958
		γ_{Age}	0.2	0.202	1%	0.0284	0.0270	0.940

* Note: A=administrative censoring time.

Est = mean of estimates; SD = empirical standard deviation of estimates; SE = mean of estimated standard errors; CP = coverage probability of 95% confidence interval.

Table 3

Failure and withdrawal events distributions

Covariate	Category	Code	Total	Failure	Withdrawal
Treatment	AZT monotherapy	1	619	196 (32%)	155 (25%)
	Combination therapy	0	613	113 (18%)	172 (28%)
Hemophiliac	Yes	1	100	23 (23%)	42 (42%)
	No	0	1132	286 (25%)	285 (25%)
Disease status	Symptomatic	1	218	75 (34%)	58 (27%)
	Asymptomatic	0	1014	234 (23%)	269 (27%)
Sex	Male	1	1005	267 (27%)	255 (25%)
	Female	0	227	42 (19%)	72 (32%)
IV drug use	Yes	1	174	33 (19%)	52 (30%)
	No	0	1058	276 (26%)	275 (26%)
Race	White	1	876	238 (27%)	241 (28%)
	Others	0	356	71 (20%)	86 (24%)
Prior antiretroviral therapy	Yes	1	700	209 (27%)	173 (25%)
	No	0	532	100 (19%)	154 (29%)
Baseline CD4 count	> 350/mm ³	1	593	113 (19%)	155 (26%)
	350/mm ³	0	639	196 (31%)	172 (27%)
Age > 50 years	Yes	1	56	21 (38%)	9 (16%)
	No	0	1176	288 (24%)	318 (27%)