Statistica Sinica **24** (2014), 1411-1432
doi:http://dx.doi.org/10.5705/ss.2012.303

ANALYSIS OF DEPENDENTLY CENSORED DATA BASED ON QUANTILE REGRESSION

Shuang Ji[1], Limin Peng[2], Ruosha Li[3] and Michael J. Lynn[2]

[1]*University of Texas Health Science Center at Houston,*
[2]*Emory University and* [3]*University of Pittsburgh*

Abstract: Dependent censoring occurs in many biomedical studies and poses considerable methodological challenges for survival analysis. We develop a new approach for analyzing dependently censored data by adopting quantile regression models. We formulate covariate effects on the quantiles of the marginal distribution of the event time of interest. Such a modeling strategy can accommodate a more dynamic relationship between covariates and survival time compared to traditional regression models in survival analysis, which usually assume constant covariate effects. We propose estimation and inference procedures, along with an efficient and stable algorithm. We establish the uniform consistency and weak convergence of the resulting estimators. Extensive simulation studies demonstrate good finite-sample performance of the proposed inferential procedures. We illustrate the practical utility of our method via an application to a multicenter clinical trial that compared warfarin and aspirin in treating symptomatic intracranial arterial stenosis.

Key words and phrases: Copula model, dependent censoring, empirical process, martingale, regression quantile.

1. Introduction

In survival analysis, a commonly adopted assumption is noninformative censoring, the time to censoring is independent of the event time of interest given covariates if any. This assumption may not be valid in many practical situations. A good example comes from the Warfarin-Aspirin Symptomatic Intracranial Disease (WASID) study, the first clinical trial that compared warfarin and aspirin in treating atherosclerotic intracranial arterial stenosis (Chimowitz et al. (2005)). In this trial, the primary endpoint was ischemic stroke, brain hemorrhage, or death from vascular causes other than stroke. Study medications were terminated early for 125 patients out of the total 569 patients due to such disease-related reasons as adverse events and changes in health conditions. It is questionable that these withdrawals were independent of the disease endpoints of interest. In addition, 44 withdrawals occurred in the aspirin arm and 81 in the warfarin arm. Such an unbalanced allocation can amplify the estimation bias for

treatment effect caused by falsely treating withdrawals as independent censoring (Huang and Zhang (2008)). These considerations necessitate properly adjusting for dependent censoring.

By viewing the occurrence of dependent censoring as one type of failure, we can formulate the survival data subject to dependent censoring as competing risks data. As a result, the dependent censoring problem can be tackled by employing techniques for handling competing risks that are generally classified into two categories (Kalbfleisch and Prentice (2002)): approaches based on the *crude quantities*, such as the cause-specific hazard and the cumulative incidence function that reflect the failure process in the presence of the competing risks; methods that focus on the *net quantities*, for example the marginal distribution function, which hypothesize the removal of the competing risks. When dependent censoring is caused by events that preclude the observation, but not the development, of the endpoint of interest, such as the informative withdrawals in the WASID study, the latter approach may be preferred because it produces inference that corresponds to the setting without interruption of the observation process, and hence may be of more scientific relevance.

There is a rich literature on competing risks approaches based on *net quantities*. As a common feature, additional assumptions on the relationship among times to distinct failure types are required because the marginal and joint distributions are not nonparametrically identifiable (Tsiatis (1975)). For example, in the one-sample case, much previous work with dependently censored data restricts the joint distribution using either semiparametric or parametric models (Link (1989) and Emoto and Matthews (1990), among others). Due to lack of sufficient information to verify the assumed dependence structure, performing a sensitivity analysis (Peterson (1976); Slud and Rubinstein (1983); Klein and Moeschberger (1988); Zheng and Klein (1995); Scharfstein et al. (2001); Scharfstein and Robins (2002)) has been advocated to yield bounds for the estimands of interest under various plausible assumptions on the joint distribution of the event time and the censoring time.

The general regression setting is the focus here. Among existing work, Huang and Zhang (2008) extended the Zheng and Klein (1995) approach to a bivariate Cox proportional hazards model, where the joint distributions of competing risks are linked to their marginal distributions through a known copula. Chen (2010) developed a non-parametric maximum likelihood approach for a general class of semiparametric transformation models, similarly assuming a copula model to address the identifiability issue. These regression methods base inference on models that only allow for constant effects, which may not be adequate in many datasets (Kaslow et al. (1987); Dickson (1989); Thorogood et al. (1990); Verweij and Van Houwelingen (1995); Jensen et al. (1997)).

We propose a new regression method for dependently censored data based on quantile regression modeling (Koenker and Bassett (1978)). By its flexibility to accommodate varying covariate effects, quantile regression can often provide useful scientific insight that cannot be uncovered by traditional regression models with constant effects (Peng and Huang (2008); Peng and Fine (2009)). While substantial work has been devoted to develop quantile regression methods for survival data with known censoring times or independent censoring times (Powell (1984, 1986); Ying, Jung, and Wei (1995); Yang (1999); Honoré, Khan, and Powell (2002); Portnoy (2003); Peng and Huang (2008); Wang and Wang (2009); Huang (2010)), to the best of our knowledge, little work has been done to accommodate the survival with dependent censoring. Peng and Fine (2009) proposed a quantile regression method for competing risks data based on the cumulative incidence function that cannot be applied to draw inference on net quantities as desired in the WASID study.

We assume linear quantile regression models for both the event time and the dependent censoring time. For identifiability, we specify the dependence structure between the event time and the dependent censoring time by a copula model, as in Huang and Zhang (2008) and Chen (2010). We utilize martingales associated with the cause-specific hazards to construct unbiased estimating equations for the assumed models, following Peng and Huang (2008). Here the estimation involves more delicate issues with the identifiability of upper tail quantiles due to censoring, as well as technical challenges due to the dependent entanglement of event time and censoring time. In the estimation procedure, we apply a "truncation" technique to avoid the estimation of upper tail quantiles. With the theory in empirical processes and stochastic integral equations, we establish asymptotic properties of the proposed estimators, including uniform consistency and weak convergence. We develop an efficient and stable iterative algorithm to solve the estimating equations. We present the method along with the asymptotic results in Section 2. In Section 3 we report results from simulation studies. An application to the WASID study is presented in Section 4 to illustrate the practical utility of our method. Section 5 concludes the paper with a few remarks.

2. Methods

2.1. Data and model

Let T denote the failure time, D denote time to dependent censoring, and C be an additional independent censoring time. Let $\tilde{\boldsymbol{Z}}$ be a $p \times 1$ covariate vector. Take $\tilde{T} = T \wedge D$, $X = \tilde{T} \wedge C$, and $\boldsymbol{Z} = (1, \tilde{\boldsymbol{Z}}^T)^T$. Let $\tilde{\delta} = I(\tilde{T} \leq C)$. The censoring indicator is $\delta = \tilde{\delta}$ if $T \leq D$, and $\delta = 2\tilde{\delta}$ if $D < T$. The observed data consist of n replicates of $(X, \delta, \boldsymbol{Z})$, denoted by $\{(X_i, \delta_i, \boldsymbol{Z}_i),\ i = 1, \cdots, n\}$.

Take the conditional τth quantile of a random variable Y given $\boldsymbol{Z}$ to be $Q_Y(\tau|\boldsymbol{Z}) = \inf\{t : F_Y(t|\boldsymbol{Z}) \geq \tau\}$, where $F_Y(t|\boldsymbol{Z}) = \Pr(Y \leq t|\boldsymbol{Z})$. We consider the quantile regression model for T and D that takes the forms

$$Q_T(\tau|\boldsymbol{Z}) = g\{\boldsymbol{Z}^T\boldsymbol{\beta}_0(\tau)\}, \quad \tau \in (0,1), \tag{2.1}$$

$$Q_D(\tau|\boldsymbol{Z}) = h\{\boldsymbol{Z}^T\boldsymbol{\alpha_0}(\tau)\}, \quad \tau \in (0,1), \tag{2.2}$$

where $g(\cdot)$ and $h(\cdot)$ are known monotone link functions, and the unknown coefficient vectors, $\boldsymbol{\beta_0}(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$, represent the covariate effects on $Q_T(\tau|\boldsymbol{Z})$ and $Q_D(\tau|\boldsymbol{Z})$, respectively. For simplicity, we assume h and g are increasing functions and $h = g$; the method can be readily extended to allow $h \neq g$ and one (or both) of them is non-increasing monotone. While interest is generally centered on $\boldsymbol{\beta}_0(\tau)$, the estimation of $\boldsymbol{\alpha}_0(\tau)$ can be useful in practice. For example, in the WASID study, inference on $\boldsymbol{\alpha}_0(\tau)$ can help to investigate the factors that contribute to the early withdrawal of patients.

Due to the dependence between T and D given the covariates, models concerning the marginal distribution functions or quantile functions, such as (2.1) and (2.2), cannot be identified without additional assumptions on the dependence structure between T and D. Here we specify the dependence structure between T and D by a copula model that relates the joint survival function of (T, D) to the marginal distributions through

$$\Pr(T > t_1, D > t_2|\boldsymbol{Z}) = H\{\Pr(T > t_1|\boldsymbol{Z}), \Pr(D > t_2|\boldsymbol{Z})\}, \tag{2.3}$$

where $H(\cdot,\cdot)$ is known copula function. For example, $H(\cdot,\cdot)$ can be chosen from a variety of parametric classes, such as the Clayton copula (Clayton (1978)) $H(u,v) = \{u^{-r_c} + v^{-r_c} - 1\}^{-1/r_c}, r_c > 0$, and the Frank copula (Genest (1987)) $H(u,v) = \log_{r_f}\{1 + (r_f^u - 1)(r_f^v - 1)/(r_f - 1)\}, r_f > 0$ and $r_f \neq 1$, where r_c and r_f are known copula parameters. In practice, the copula parameter may be chosen according to prior knowledge on the strength of the association between T and D. Alternatively, one can perform a sensitivity analysis to obtain bounds of $\boldsymbol{\beta}_0(\tau)$ and hence $Q_T(\tau|\boldsymbol{Z})$ by perturbing r in a plausible range.

2.2. Estimation equations

To estimate $\boldsymbol{\beta}_0(\tau)$ at (2.1), we utilize martingales associated with cause-specific hazard functions. Let the counting process for T be $N_1(t) = I(X \leq t, \delta = 1)$, and take $M_1(t) = N_1(t) - \int_0^t Y(u)\lambda_1^*(u|\boldsymbol{Z})\,\mathrm{d}u$, where $Y(u) = I(X \geq u)$ and $\lambda_1^*(t|\boldsymbol{Z}) = \lim_{h\to 0}\Pr\{t \leq T < t + h, T < D|T \geq t, D \geq t; \boldsymbol{Z}\}/h$ as the cause-specific hazard function for T. As shown by Kalbfleisch and Prentice (2002), $M_1(t)$ is a martingale with respect to the filtration $\mathscr{F}_t = \{N_1(u), Y(u+), \boldsymbol{Z}\}$. This implies

$$E\{N_1(t) - \int_0^t Y(s)\lambda_1^*(s|\boldsymbol{Z})\,\mathrm{d}s\} = 0, \forall t \geq 0. \tag{2.4}$$

By using $\lambda_1^*(t|\boldsymbol{Z}) = -\partial \log[H\{\bar{F}_T(t_1|\boldsymbol{Z}), \bar{F}_D(t_2|\boldsymbol{Z}); r\}]/\partial t_1|_{t_1=t_2=t}$ (Kalbfleisch and Prentice (2002)) and employing variable transformation inside the integral, we can show that

$$\int_0^t Y(s)\lambda_1^*(s|\boldsymbol{Z})\,\mathrm{d}s = \int_0^{F_T\{t|\boldsymbol{Z}\}} Y\{Q_T(u|\boldsymbol{Z})\}\phi_1(1-u, \bar{F}_D\{Q_T(u|\boldsymbol{Z})|\boldsymbol{Z}\})\,\mathrm{d}u, \tag{2.5}$$

where $\phi_1(v_1, v_2) = \partial \log\{H(v_1, v_2)\}/\partial v_1$ and $\bar{F}_W(t)$ is the survival function for a random variable W. Under models (2.1) and (2.2),

$$F_D(t|\boldsymbol{Z}) = \int_0^1 I\{v \le F_D(t|\boldsymbol{Z})\}\,\mathrm{d}v = \int_0^1 I[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}_0(v)\} \le t]\,\mathrm{d}v$$

and therefore

$$\bar{F}_D\{Q_T(u|\boldsymbol{Z})|\boldsymbol{Z}\} = 1 - \int_0^1 I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}_0(v) \le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(u)\}\,\mathrm{d}v. \tag{2.6}$$

From (2.1),(2.4), (2.5), and (2.6) we have

$$E\Big[\frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{1i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(\tau)\}] - \int_0^\tau Y[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(u)\}] \times\phi_1\Big(1-u, 1-\int_0^1 I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}_0(v) \le \boldsymbol{Z}_i^T\boldsymbol{\beta}_0(u)\}\,\mathrm{d}v\Big)\,\mathrm{d}u\Big\}\Big] = 0, \tag{2.7}$$

where $N_{1i}(t)$ is the sample analog of $N_1(t)$.

By treating T as the dependent censoring to D, a parallel equality to (2.7) can be derived for $\boldsymbol{\alpha}_0(\cdot)$. Take $N_2(t) = I(X \le t, \delta = 2)$ and let $\{N_{2i}(t)\}_{i=1}^n$ be the sample analogs of $N_2(t)$. With $\phi_2(v_1, v_2) = \partial \log\{H(v_1, v_2)\}/\partial v_2$, we can show that

$$E\Big[\frac{1}{n}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{2i}[g\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(\tau)\}] - \int_0^\tau Y[g\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(u)\}] \times\phi_2\Big(1-\int_0^1 I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}_0(v) \le \boldsymbol{Z}_i^T\boldsymbol{\alpha}_0(u)\}\,\mathrm{d}v, 1-u\Big)\,\mathrm{d}u\Big\}\Big] = 0. \tag{2.8}$$

Motivated by (2.7) and (2.8), we propose to estimate $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$ from the estimating equations

$$n^{1/2}\boldsymbol{S}_n^{(k)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0, \quad k = 1, 2, \tag{2.9}$$

where

$$\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{1i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}] - \int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}]$$

$$\times\phi_1\Big(1-u, 1-\int_0^1 I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(v)\le \boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}\,\mathrm{d}v\Big)\,\mathrm{d}u\Big\},$$

$$\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta},\boldsymbol{\alpha},\tau)=n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{2i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(\tau)\}]-\int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}]$$

$$\times\phi_2\Big(1-\int_0^1 I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(v)\le \boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}\,\mathrm{d}v, 1-u\Big)\,\mathrm{d}u\Big\}.$$

The estimating equation (2.9) requires that $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$ be identifiable for all $\tau\in(0,1)$, which may not be possible due to the censoring to T or D. Here we modify (2.9) by truncating the time scale by an upper bound, $g\{\boldsymbol{Z}^T\boldsymbol{\beta}_0(\tau_{U,1})\}\wedge g\{\boldsymbol{Z}^T\boldsymbol{\alpha}_0(\tau_{U,2})\}$, where $\tau_{U,1}$, $\tau_{U,2}\in(0,1)$. This leads the new estimating equation

$$n^{1/2}\boldsymbol{S}_n^{*(k)}(\boldsymbol{\beta},\boldsymbol{\alpha},\tau)=0,\quad k=1,2, \tag{2.10}$$

where

$$\boldsymbol{S}_n^{*(1)}(\boldsymbol{\beta},\boldsymbol{\alpha},\tau)=n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{1i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau)\}]I\{g^{-1}(X_i)\le \boldsymbol{Z}_i^T\boldsymbol{\alpha}(\tau_{U,2})\}$$

$$-\int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}]I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\le \boldsymbol{Z}_i^T\boldsymbol{\alpha}(\tau_{U,2})\}$$

$$\times\phi_1\Big(1-u, 1-\int_0^{\tau_{U,2}} I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(v)\le \boldsymbol{Z}_i^T\boldsymbol{\beta}(u)\}\,\mathrm{d}v\Big)\,\mathrm{d}u\Big\},$$

$$\boldsymbol{S}_n^{*(2)}(\boldsymbol{\beta},\boldsymbol{\alpha},\tau)=n^{-1}\sum_{i=1}^n \boldsymbol{Z}_i\Big\{N_{2i}[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(\tau)\}]I\{g^{-1}(X_i)\le \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_{U,1})\}$$

$$-\int_0^\tau Y_i[g\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}]I\{\boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\le \boldsymbol{Z}_i^T\boldsymbol{\beta}(\tau_{U,1})\}$$

$$\times\phi_2\Big(1-\int_0^{\tau_{U,1}} I\{\boldsymbol{Z}_i^T\boldsymbol{\beta}(v)\le \boldsymbol{Z}_i^T\boldsymbol{\alpha}(u)\}\,\mathrm{d}v, 1-u\Big)\,\mathrm{d}u\Big\}.$$

It can be shown that the terms in (2.10) still have expectation 0 at the true parameters. A nice feature of the modified estimating equations is that they only involve the estimation of $\{\boldsymbol{\beta}_0(\tau),\ \tau\in(0,\tau_{U,1})\}$ and $\{\boldsymbol{\alpha}_0(\tau),\ \tau\in(0,\tau_{U,2})\}$, and thus do not demand the identifiability of $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$ in the upper tail of τ. The conditions for $\tau_{U,1}$ and $\tau_{U,2}$ are deferred to the statement of asymptotic results. In practice, $\tau_{U,1}$ and $\tau_{U,2}$ may need to be selected adaptively, and some empirical rules for selecting them are to be presented.

2.3. Computing algorithms

We develop an iterative algorithm for finding the solution to (2.10), Algorithm A, described as follows.

Step A0. Set $m=0$. Choose the initial value $\hat{\boldsymbol{\alpha}}^{[m]}(\tau), \tau\in(0,\tau_{U,2}]$.

Step A1. Solve $\boldsymbol{S}_n^{*(1)}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{[m]}, \tau) = 0$ for $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau), \tau \in (0, \tau_{U,1}^{[m+1]}]$. Update $\tau_{U,1}$ with $\tau_{U,1}^{[m+1]}$.

Step A2. Solve $\boldsymbol{S}_n^{*(2)}(\hat{\boldsymbol{\beta}}^{[m+1]}, \boldsymbol{\alpha}, \tau) = 0$ for $\hat{\boldsymbol{\alpha}}^{[m+1]}(\tau), \tau \in (0, \tau_{U,2}^{[m+1]}]$. Update $\tau_{U,2}$ with $\tau_{U,2}^{[m+1]}$.

Step A3. Let $m = m + 1$. Repeat Steps A1 and A2 until convergence criteria are met.

At Step A0, one way to set the initial estimates is to fit model (2.1) for T and model (2.2) for D using existing quantile regression techniques that assume T and D are independent, for example, using Peng and Huang (2008)'s method.

At Step A1, we adopt a grid-based procedure that assumes $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau)$ to be a right-continuous step function jumping only on a prespecified grid, $\mathscr{G}_{L_n} = \{0 = \tau_0 < \tau_1 < \cdots < \tau_{L_n} = \tau_{U,1}^{[m+1]} < 1\}$. Let $\|\mathscr{G}_{L_n}\|$ be the size of the grid $\mathscr{G}_{L_n}$, $\max\{\tau_{j+1} - \tau_j;\, j = 0, \cdots, L_n - 1\}$. The solution can be obtained by sequentially solving the following monotone estimating equation in $\boldsymbol{\beta}(\tau_j)(j = 1, \cdots, L_n)$:

$$
\begin{aligned}
n^{-1/2} \sum_{i=1}^{n} \boldsymbol{Z}_i \Big\{ & I[X_i \le g\{\boldsymbol{Z}_i^T \boldsymbol{\beta}(\tau_j)\}, \delta_i = 1] I\{g^{-1}(X_i) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} \\
& - \sum_{l=0}^{j-1} (\tau_{l+1} - \tau_l) I[X_i \ge g\{\boldsymbol{Z}_i^T \boldsymbol{\beta}(\tau_l)\}] I\{\boldsymbol{Z}_i^T \boldsymbol{\beta}(\tau_l) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} \\
& \times \phi_1 \Big(1 - \tau_l, 1 - \int_0^{\tau_{U,2}} I\{\boldsymbol{Z}_i^T \hat{\boldsymbol{\alpha}}^{[m]}(v) \le \boldsymbol{Z}_i^T \boldsymbol{\beta}(\tau_l)\}\, \mathrm{d}v \Big) \Big\} = 0, \qquad (2.11)
\end{aligned}
$$

with $g\{\boldsymbol{Z}_i^T \boldsymbol{\beta}(0)\}$ set to be 0.

Due to the monotonicity of (2.11), the root-finding problem in (2.11) is equivalent to locating the minimizer of the L_1-type convex function

$$
\begin{aligned}
l_j(\boldsymbol{h}) = \sum_{i=1}^{n} & |I(\delta_i = 1) I\{g^{-1}(X_i) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} g^{-1}(X_i) \\
& - I(\delta_i = 1) I\{g^{-1}(X_i) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} \boldsymbol{h}^T \boldsymbol{Z}_i| \\
& + |R^* - \boldsymbol{h}^T \sum_{l=1}^{n} \{-I(\delta_l = 1)\} I\{g^{-1}(X_l) \le \boldsymbol{Z}_l^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} \boldsymbol{Z}_l| \\
& + \Big| R^* - \boldsymbol{h}^T \sum_{r=1}^{n} \Big(2\boldsymbol{Z}_r \sum_{s=0}^{j-1} I[g^{-1}(X_r) \ge \boldsymbol{Z}_r^T \boldsymbol{\beta}(\tau_s)] I\{\boldsymbol{Z}_r^T \boldsymbol{\beta}(\tau_s) \le \boldsymbol{Z}_r^T \hat{\boldsymbol{\alpha}}^{[m]}(\tau_{U,2})\} \\
& \times \phi_1 \Big(1 - \tau_s, 1 - \int_0^{\tau_{U,2}} I\{\boldsymbol{Z}_r^T \hat{\boldsymbol{\alpha}}^{[m]}(v) \le \boldsymbol{Z}_r^T \boldsymbol{\beta}(\tau_s)\} \mathrm{d}v \Big) \times (\tau_{s+1} - \tau_s) \Big) \Big|, \qquad (2.12)
\end{aligned}
$$

where R^* is suitably large.

Note that $\tau_{U,1}$ is adaptively selected and may vary at each iteration. At the end of the mth iteration, we choose $\tau_{U,1}$ to be $\tau_{U,1}^{[m+1]}$, the largest quantile at which $\hat{\boldsymbol{\beta}}^{[m+1]}(\cdot)$ can be solved. For example, we examine the distance between $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau_j)$ and $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau_{j+1})$ for each j, d_j, and if it exceeds a moderate prespecified threshold, we stop the sequential procedure and set $J = j$ and thus $\tau_{U,1}^{[m+1]} = \tau_J$. In our numerical studies we set the threshold to 10. The underlying rationale is that, given a fine grid $\mathscr{G}_{L_n}$, $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau_j)$ and $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau_{j+1})$ are expected be close in the identifiable τ–region for $\hat{\boldsymbol{\beta}}(\cdot)$ when $j > 0$.

The root-finding procedure at Step A2 can be transformed to minimizing a L_1-type convex function parallel to (2.12) and we omit the exact expressions here. The L_1–minimization problem can be readily solved by using existing packages implemented in standard statistical software, such as $l1fit()$ function in S-PLUS and $rq()$ function in R. A similar adaptive strategy as for selecting $\tau_{U,1}$ can be adopted for $\tau_{U,2}$, which is updated at the mth iteration with $\tau_{U,2}^{[m+1]}$, the largest quantile at which $\hat{\boldsymbol{\alpha}}^{[m+1]}(\cdot)$ can be identified.

Based on our experience, the numerical performance of this algorithm can be unstable when there is heavy censoring on D. For example, in the context of the WASID study, about 80% of the observations on D were censored by either T or C. This is expected in a well-designed study when D represents informative dropouts. In such a case, adopting a more restrictive version of model (2.2) for D can improve the estimation efficiency and thus help increase the numerical stability. One specific remedy is to adopt an AFT model for D that only allows the intercept $\alpha_0^{(0)}(\tau)$ to vary with τ, but imposes constancy on each covariate effect $\alpha_0^{(k)}(\tau)$ for $k = 1, \cdots, p$. Since D is subject to dependent censoring posed by T, classical AFT estimation that requires conditional independent censoring is not applicable to fit a AFT model for D. Taking this into account, we propose a modified version of Algorithm A, Algorithm B, described as follows.

Step B0. Set $m = 0$. Obtain the initial values $\hat{\boldsymbol{\alpha}}^{[m]}(\tau), \tau \in (0, \tau_{U,2}]$ by fitting an AFT model using Jin et al. (2003)'s method.

Step B1. Solve $\boldsymbol{S}_n^{*(1)}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{[m]}, \tau) = 0$ for $\hat{\boldsymbol{\beta}}^{[m+1]}(\tau), \tau \in (0, \tau_{U,1}^{[m+1]}]$. Update $\tau_{U,1}$ with $\tau_{U,1}^{[m+1]}$.

Step B2. Obtain $\hat{\boldsymbol{\alpha}}^{[m+1]}(\tau), \tau \in (0, \tau_{U,2}^{[m+1]}]$ via the following procedure

(a) Solve $\boldsymbol{S}_n^{*(2)}(\hat{\boldsymbol{\beta}}^{[m+1]}, \boldsymbol{\alpha}, \tau) = 0$ for $\tilde{\boldsymbol{\alpha}}^{[m+1]}(\tau), \tau \in (0, \tilde{\tau}_{U,2}]$.

(b) Obtain the constant $\hat{\alpha}^{[m+1]^{(k)}}$ by taking the average of $\tilde{\alpha}^{[m+1]^{(k)}}(\tau)$ over $\tau \in [\tau_a, \tau_b]$ for $k = 1, \cdots, p$, where $\tau_a \in (0, \tilde{\tau}_{U,2})$ and $\tau_b \in (\tau_a, \tilde{\tau}_{U,2})$ are prespecified constants that represent a well-identified region for $\tilde{\boldsymbol{\alpha}}^{[m+1]}(\tau)$.

(c) Compute the residual on the g^{-1} scale, $g^{-1}(\epsilon_i) = g^{-1}(X_i) - Q_i$, where $Q_i = \tilde{\boldsymbol{Z}}_i^T(\hat{\alpha}^{[m+1](1)}, \cdots, \hat{\alpha}^{[m+1](p)})^T$.

(d) Obtain $\hat{\alpha}^{(0)[m+1]}(\tau)$ for $\tau \in (0, \tau_{U,2}^{[m+1]}]$ by solving $\boldsymbol{S}_n^{**(2)}(\hat{\boldsymbol{\beta}}^{[m+1]}, \alpha^{(0)}, \tau) = 0$, where

$$\begin{aligned}
&\boldsymbol{S}_n^{**(2)}(\hat{\boldsymbol{\beta}}^{[m+1]}, \alpha^{(0)}, \tau) \\
&\quad = n^{-1}\sum_{i=1}^{n} \boldsymbol{Z}_i \Big\{ I[g^{-1}(\epsilon_i) \le \alpha^{(0)}(\tau), \delta_i = 2] I\{g^{-1}(X_i) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\beta}}^{[m+1]}(\tau_{U,1})\} \\
&\qquad - \int_0^{\tau} I[g^{-1}(\epsilon_i) \ge \alpha^{(0)}(\tau)] I\{\alpha^{(0)}(u) \le \boldsymbol{Z}_i^T \hat{\boldsymbol{\beta}}^{[m+1]}(\tau_{U,1}) - Q_i\} \\
&\qquad \times \phi_2\Big(1 - \int_0^{\tau_{U,1}} I\{\boldsymbol{Z}_i^T \hat{\boldsymbol{\beta}}^{[m+1]}(v) - Q_i \le \alpha^{(0)}(u)\}\, \mathrm{d}v, 1 - u\Big)\, \mathrm{d}u \Big\}.
\end{aligned}$$

Update $\tau_{U,2}$ with $\tau_{U,2}^{[m+1]}$.

Step B3. Let $m = m + 1$. Repeat Steps B1 and B2 until convergence criteria are met.

Here $\tau_{U,2}$ is selected in a slightly different manner than in Algorithm A. At the mth iteration, we set $\tau_{U,2}$ at $\tau_{U,2}^{[m+1]}$, the largest τ at which the intercept $\hat{\alpha}^{(0)}(\tau)$ can be obtained. We still select $\tau_{U,1}$ based on the identifiability of the $p+1$ vector $\hat{\boldsymbol{\beta}}(\tau)$. As in Steps A1 and A2, equations involved in Steps B1 and B2 can also be treated as L_1 minimization problems and thus conveniently solved. Details of the convergence criteria for Steps A3 and B3 are provided in Appendix D.

In practice, one may consider more general parametric submodeling of $\boldsymbol{\alpha}_0(\tau)$ along the lines of Fine, Yan, and Kosorok (2004) to bring down the dimensionality of regression quantiles for D. The AFT model based remedy can be viewed as a special case of this type of analysis, where constant submodels are assumed for all non-intercept coefficients in $\boldsymbol{\alpha}_0(\tau)$. The algorithm B can be adapted accordingly.

2.4. Asymptotic results

Under regularity conditions C1−C5 (provided in Appendix A), we establish uniform consistency and weak convergence for $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\boldsymbol{\alpha}}(\tau)$.

Theorem 1. *If* C1−C5 *hold and* $\lim_{n\to\infty} \|\mathscr{G}_{L_n}\| = 0$, *then* $\sup_{\tau\in[\nu_1,\tau_{U,1}]} \|\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| \xrightarrow{p} 0$ *and* $\sup_{\tau\in[\nu_2,\tau_{U,2}]} \|\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\| \xrightarrow{p} 0$, *where* $0 < \nu_1 < \tau_{U,1}$ *and* $0 < \nu_2 < \tau_{U,2}$.

Theorem 2. *If* C1−C5 *hold and* $\lim_{n\to\infty} n^{1/2}\|\mathscr{G}_{L_n}\| = 0$, *then* $n^{1/2}\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}$ *converges weakly to a Gaussian process for* $\tau \in [\nu_1, \tau_{U,1}]$ *with* $0 < \nu_1 <$

$\tau_{U,1}$, and $n^{1/2}\{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}$ converges weakly to a Gaussian process for $\tau \in [\nu_2, \tau_{U,2}]$ with $0 < \nu_2 < \tau_{U,2}$.

The proofs of these theorems can be viewed as extensions of those in Peng and Huang (2008) to the bivariate case, but are not straightforward. To prove Theorem 1, we first note that the proposed estimating functions $\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)$ and $\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)$ converge to their expectations $\boldsymbol{s}^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)$ and $\boldsymbol{s}^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau)$ uniformly in τ. Second, with fixed $\boldsymbol{\alpha}$ in equations $\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ and $\boldsymbol{s}^{(1)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$, the solutions for $\boldsymbol{\beta}$ can be viewed as functionals of $\boldsymbol{\alpha}$, namely $\hat{\boldsymbol{\beta}}(\boldsymbol{\alpha}, \tau)$ and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}, \tau)$, respectively. We can then use $\tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}, \tau)$ to bridge $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\alpha}}, \tau)$ and $\boldsymbol{\beta}_0(\tau) = \tilde{\boldsymbol{\beta}}(\boldsymbol{\alpha}_0, \tau)$. Similarly we can use $\tilde{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}}, \tau)$ to bridge $\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\beta}}, \tau)$ and $\boldsymbol{\alpha}_0(\tau) = \tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}_0, \tau)$, where $\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \tau)$ and $\tilde{\boldsymbol{\alpha}}(\boldsymbol{\beta}, \tau)$ are the solutions for $\boldsymbol{\alpha}$ to $\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ and $\boldsymbol{s}^{(2)}(\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau) = 0$ with fixed $\boldsymbol{\beta}$, respectively. To circumvent the difficulty that $\|\hat{\boldsymbol{\beta}}(0)\| = \infty$ and $\|\hat{\boldsymbol{\alpha}}(0)\| = \infty$, which is implied by models (2.1) and (2.2) and our estimating procedure, we consider

$$\boldsymbol{\theta}(\tau) = \boldsymbol{\mu}\begin{pmatrix}\boldsymbol{\beta}(\tau)\\ \boldsymbol{\alpha}(\tau)\end{pmatrix} = \begin{pmatrix} E\Big(\boldsymbol{Z}N_1[g\{\boldsymbol{Z}^T\boldsymbol{\beta}(\tau)\}]\Big)\\ E\Big(\boldsymbol{Z}N_2[g\{\boldsymbol{Z}^T\boldsymbol{\alpha}(\tau)\}]\Big)\end{pmatrix},$$

and prove that $\hat{\boldsymbol{\theta}}(\tau)$ converges in probability to $\boldsymbol{\theta}_0(\tau)$ uniformly for $\tau \in (0, \tau_U]$. This result further leads to the uniform convergency of $\hat{\boldsymbol{\beta}}(\tau)$ for $\tau \in [\nu_1, \tau_{U,1}]$ and $\hat{\boldsymbol{\alpha}}(\tau)$ for $\tau \in [\nu_2, \tau_{U,2}]$.

To prove Theorem 2, we first establish the connection between $n^{1/2}(\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}^T, \{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}^T)^T$ and $n^{1/2}\{-\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)^T, -\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)^T\}^T$ via a stochastic integral equation. This result allows us to express $n^{1/2}(\{\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\}^T, \{\hat{\boldsymbol{\alpha}}(\tau) - \boldsymbol{\alpha}_0(\tau)\}^T)^T$ as a linear map of $n^{1/2}\{-\boldsymbol{S}_n^{(1)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)^T$, $-\boldsymbol{S}_n^{(2)}(\boldsymbol{\beta}_0, \boldsymbol{\alpha}_0, \tau)^T\}^T$. The latter can be shown to have weak convergence, which implies the result in Theorem 2. The detailed proofs of Theorems 1 and 2 are provided in Appendices B and C.

2.5. Inferences

Given the complex limiting distributions of $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\boldsymbol{\alpha}}(\tau)$, seen in the proof of Theorem 2, we employ the bootstrap approach (Efron (1979)) to make inference on $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$. We obtain B bootstrapped samples, each of which is obtained by resampling with replacement n times from the original dataset. For the b-th bootstrapped sample, we conduct the estimation procedure presented in Sections 2.2–2.3 and obtain $\{\boldsymbol{\beta}_b^*(\tau), \tau \in (0, \tau_{U,1,b}^*]\}$ and $\{\boldsymbol{\alpha}_b^*(\tau), \tau \in (0, \tau_{U,2,b}^*]\}$ $(b = 1, \ldots, B)$. For each fixed τ, we estimate the variances of $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\boldsymbol{\alpha}}(\tau)$ by the sample variances of $\{\boldsymbol{\beta}_b^*(\tau)\}_{b=1}^B$ and $\{\boldsymbol{\alpha}_b^*(\tau)\}_{b=1}^B$, respectively, and construct confidence intervals of $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$ using normal approximations.

Hypotheses testing can be conducted to further investigate the patterns of the covariate effects. Let $\beta_0^{(q)}$ be the coefficient corresponding to $\tilde{\boldsymbol{Z}}^{(q)}$, the qth component of $\tilde{\boldsymbol{Z}}$ $(q = 1, \cdots, p)$. One might be especially interested in testing the overall significance of $\beta_0^{(q)}(\tau)$ across a pre-specified range of τ, say $[l, u]$, where $0 < l < u < \tau_{U,1}$, and the constancy of $\beta_0^{(q)}(\tau)$ over $\tau \in [l, u]$. The corresponding null hypotheses can be formulated as $H_0 : \beta_0^{(q)}(\tau) = 0, \tau \in [l, u]$ and $\tilde{H}_0 : \beta_0^{(q)}(\tau) = \rho_0, \tau \in [l, u]$, respectively, where ρ_0 is an unknown constant. To address these problems, we first consider $\eta_{0,q} \equiv \int_l^u \beta_0^{(q)}(v)dv/(u-l)$ for $q = 1, \cdots, p$, which may be interpreted as the average covariate effect of $\tilde{\boldsymbol{Z}}^{(q)}$ across $\tau \in [l, u]$. Following the justification provided in Peng and Huang (2008), it can be shown that $\hat{\eta}_q = \int_l^u \hat{\beta}^{(q)}(v)dv/(u-l)$ is a consistent estimator for $\eta_{0,q}$ and is asymptotically normal. Given the observed data, the limiting distribution of $\hat{\eta}_q$ can be approximated by the sample $\{\eta_{b,q}^*\}_{b=1}^B$, where $\eta_{b,q}^* = \int_l^u \beta_b^{*(q)}(v)dv/(u-l)$. To test H_0, we note that under the null, the limit distribution of $\hat{\eta}_q$ is a mean zero normal distribution, the variance of which can be estimated via the resampling procedure mentioned above. Therefore, a Wald-type test statistic for testing H_0 is given by $\hat{\eta}_q$ divided by its standard error.

Regarding $\tilde{H}_0$, one can adopt the test statistic $\tilde{\Gamma} = \{n^{1/2} \int_l^u \{\hat{\beta}^{(q)}(v) - \hat{\eta}_q\}\Theta(v)\mathrm{d}v\}/(u-l)$, where $\Theta(\cdot)$ is a pre-specified nonconstant and nonnegative weight function. The essential idea of $\tilde{\Gamma}$ is to compare two different weighted averages of $\hat{\beta}^{(q)}(\tau)$ over τ, which is expected to be small if $\beta_0^{(q)}(\tau)$ is constant over τ. When $\tilde{H}_0$ holds, $\beta_0^{(q)}(v) = \eta_{0,q}$ for all $v \in [l, u]$ and thus $\tilde{\Gamma} = n^{1/2} \int_l^u [\{\hat{\beta}^{(q)}(v) - \beta_0^{(q)}(v)\} - (\hat{\eta}_q - \eta_{0,q})]\Theta(v)\,\mathrm{d}v/(u-l)$. Given the functional linearity of $\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)$ implied by the proof of Theorem 2.2, we can show that the limiting distribution of $\tilde{\Gamma}$ under $\tilde{H}_0$ is a normal distribution, which can be approximated by the conditional distribution of $\tilde{\Gamma}_b^* = n^{1/2} \int_l^u [\{\beta_b^{*(q)}(v) - \hat{\beta}^{(q)}(v)\} - (\eta_{b,q}^* - \hat{\eta}_q)]\Theta(v)\,\mathrm{d}v/(u-l)$ given the observed data. Therefore, a percentile based test of size α is to reject $\tilde{H}_0$ when $\tilde{\Gamma} > d_{1-\alpha/2}$ or $\tilde{\Gamma} < d_{\alpha/2}$, where $d_{1-\alpha/2}$ and $d_{\alpha/2}$ are the $100(1-\alpha/2)$th and $100(\alpha/2)$th percentiles of the empirical distribution of $\tilde{\Gamma}_b^*$. Alternatively, one can estimate the variance of $\tilde{\Gamma}$ by $\hat{\sigma}(\tilde{\Gamma})^2$, the empirical variance of $\tilde{\Gamma}_b^*$. The p value for the Wald type test can be obtained by comparing $\tilde{\Gamma}/\hat{\sigma}(\tilde{\Gamma})$ with the distribution $N(0, 1)$.

The hypothesis testing procedures presented follow similar lines as in Fine, Yan, and Kosorok (2004) and Peng and Huang (2008). Detailed justifications are thus omitted.

3. Simulation Studies

We studied the finite-sample performance of the proposed estimators via Monte-Carlo simulations. For the association structure between T and D, we

considered the Clayton copula with association parameter r_c and the Frank copula with association parameter r_f. We set $r_c = \exp(1)$ and $r_f = \exp(-7.325)$ and, correspondingly, the values of Kendall's tau are the same for both settings, 0.576, representing moderate dependency. To achieve the desired dependence structure between T and D, we generated ϵ_1 and ϵ_2 based on a shared frailty model so that (ϵ_1, ϵ_2) follows a Clayton or Frank copula model. We adopted the gamma frailty to generate the Clayton copula model, following the procedure provided in Oakes (1989); we used the Log-series frailty for Frank copula as described in Yan (2007).

We generated T from a log linear model with heteroscedastic errors:

$$\log T = b_1 Z_1 + b_2 Z_2 + \epsilon_1,$$

where $Z_1 \sim Unif(0,1)$, $Z_2 \sim Bernoulli(0.5)$, and ϵ_1 follows $N(0, 0.2^2)$ if $Z_2 = 0$ and $N(0, 0.4^2)$ if $Z_2 = 1$. In addition, D was generated from the AFT model

$$\log D = a_1 Z_1 + a_2 Z_2 + \epsilon_2,$$

where $\epsilon_2 \sim N(\mu_2, 0.3^2)$. The independent censoring time C was $Unif(0, c_u)$. Under this set-up, models (2.1) and (2.2) hold with $g(\cdot) = \exp(\cdot)$. It can be shown that the underlying regression quantile is $\boldsymbol{\beta}_0(\tau) = \{\beta_0^{(0)}(\tau), \beta_0^{(1)}(\tau), \beta_0^{(2)}(\tau)\}^T$, where $\beta_0^{(0)}(\tau) = Q_{N(0,0.2^2)}(\tau)$, $\beta_0^{(1)} = b_1$, and $\beta_0^{(2)} = b_2 + Q_{N(0,0.4^2)}(\tau) - Q_{N(0,0.2^2)}(\tau)$. It can also be seen that $\boldsymbol{\alpha}_0(\tau) = \{Q_{\epsilon_2}(\tau), a_1, a_2\}^T$. Under each copula, we considered two specific configurations: (I) $\mu_2 = 0, b_1 = 0.27, b_2 = 0, a_1 = 0, a_2 = 0.3$, and $c_u = 12$, which results in 10% independent censoring and 45% dependent censoring to T, and thus 45% dependent censoring to D; (II) $\mu_2 = 0.1, b_1 = 0.27, b_2 = 0, a_1 = 0, a_2 = 0.3$, and $c_u = 12$, which results in 10% independent censoring and 30% dependent censoring to T, and thus 60% dependent censoring to D. For case (I) we assumed a general quantile regression model for D. For case (II) we adopted the modified algorithm assuming AFT model for D with $\tau_a = 0.1$ and $\tau_b = 0.4$.

Under each configuration we simulated 1,000 date sets of sample size $n = 200$. An equally spaced grid on τ with size 0.01 was adopted when estimating $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$. We chose $B = 100$ as the number of bootstrap replicates for the variance estimation.

Table 1 presents the estimation results when the Clayton copula was correctly adopted. We report the biases (Bias), empirical standard deviations (EmpSD), average estimated resampling-based standard deviations (AvgSD) of $\hat{\boldsymbol{\beta}}(\tau)$ and $\hat{\boldsymbol{\alpha}}(\tau)$, and coverage rates of 95% Wald confidence intervals of $\boldsymbol{\beta}_0(\tau)$ and $\boldsymbol{\alpha}_0(\tau)$ with $\tau = 0.1, 0.3, 0.5$ and 0.7. These results show that under these set-ups the biases are small, the bootstrap standard errors agree well with the empirical

Table 1. Simulation results on parameter estimation under the Clayton copula. Bias: biases; AvgSD: average estimated resampling-based standard deviations; EmpSD: empirical standard deviations; Cov95: coverage rates of 95% Wald confidence intervals.

τ		Bias	EmpSD	AvgSD	Cov95		Bias	EmpSD	AvgSD	Cov95
10% indep. censoring, 45% dep. censoring to T, 45% dep. censoring to D, model (2.2) for D										
0.1	$\hat{\beta}^{(0)}$	0.02	0.10	0.11	0.95	$\hat{\alpha}^{(0)}$	0.01	0.10	0.11	0.94
	$\hat{\beta}^{(1)}$	0.02	0.19	0.22	0.95	$\hat{\alpha}^{(1)}$	-0.02	0.17	0.19	0.94
	$\hat{\beta}^{(2)}$	-0.02	0.10	0.11	0.96	$\hat{\alpha}^{(2)}$	0.02	0.10	0.11	0.95
0.3	$\hat{\beta}^{(0)}$	0.01	0.09	0.10	0.95	$\hat{\alpha}^{(0)}$	0.02	0.08	0.10	0.94
	$\hat{\beta}^{(1)}$	0.01	0.16	0.19	0.97	$\hat{\alpha}^{(1)}$	-0.02	0.14	0.16	0.94
	$\hat{\beta}^{(2)}$	-0.01	0.08	0.09	0.96	$\hat{\alpha}^{(2)}$	0.01	0.08	0.10	0.96
0.5	$\hat{\beta}^{(0)}$	0.01	0.08	0.09	0.95	$\hat{\alpha}^{(0)}$	0.02	0.09	0.13	0.97
	$\hat{\beta}^{(1)}$	0.02	0.16	0.21	0.97	$\hat{\alpha}^{(1)}$	-0.02	0.14	0.19	0.97
	$\hat{\beta}^{(2)}$	-0.01	0.08	0.09	0.96	$\hat{\alpha}^{(2)}$	0.01	0.08	0.10	0.97
10% indep. censoring, 30% dep. censoring to T, 60% dep. censoring to D, AFT model for D										
0.1	$\hat{\beta}^{(0)}$	0.01	0.09	0.10	0.94	$\hat{\alpha}^{(0)}$	0.02	0.08	0.10	0.96
	$\hat{\beta}^{(1)}$	0.01	0.16	0.18	0.95	$\hat{\alpha}^{(1)}$	-0.03	0.13	0.15	0.95
	$\hat{\beta}^{(2)}$	-0.01	0.09	0.10	0.96	$\hat{\alpha}^{(2)}$	0.03	0.09	0.10	0.96
0.3	$\hat{\beta}^{(0)}$	0.01	0.07	0.08	0.95	$\hat{\alpha}^{(0)}$	0.02	0.08	0.10	0.96
	$\hat{\beta}^{(1)}$	0.00	0.13	0.15	0.96	$\hat{\alpha}^{(1)}$	-0.03	0.13	0.15	0.95
	$\hat{\beta}^{(2)}$	-0.01	0.07	0.08	0.96	$\hat{\alpha}^{(2)}$	0.03	0.09	0.10	0.96
0.5	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.96	$\hat{\alpha}^{(0)}$	0.02	0.08	0.10	0.97
	$\hat{\beta}^{(1)}$	0.00	0.13	0.15	0.97	$\hat{\alpha}^{(1)}$	-0.03	0.13	0.15	0.95
	$\hat{\beta}^{(2)}$	-0.00	0.07	0.08	0.95	$\hat{\alpha}^{(2)}$	0.03	0.09	0.10	0.96
0.7	$\hat{\beta}^{(0)}$	0.00	0.07	0.08	0.96	$\hat{\alpha}^{(0)}$	0.03	0.10	0.11	0.96
	$\hat{\beta}^{(1)}$	0.01	0.14	0.18	0.97	$\hat{\alpha}^{(1)}$	-0.03	0.13	0.15	0.95
	$\hat{\beta}^{(2)}$	-0.00	0.07	0.08	0.96	$\hat{\alpha}^{(2)}$	0.03	0.09	0.10	0.96

ones, and the coverage rates are in general close to the nominal level. For case (I), the convergence rate was 94.5% and on average 3.7 iterations were required to achieve convergence. For case (II) the convergence rate was 99.2%, achieved by an average of 4.4 iterations. In a similar fashion, Table E1 in Appendix E presents the estimation results when data were generated based on the Frank copula model. These results are also satisfactory. For case (I), the convergence rate was 92.7% with 5.5 iterations on average. For case (II) the convergence rate was 99.7%, achieved by an average of 5.3 iterations. With a dataset, the algorithm failing to converge might indicate a lack of sufficient information to estimate regression quantiles within the specified τ-range. In that case, one can consider inference on a narrower τ-range or a more restrictive model for T or D which requires less information to achieve a reasonable fit of the data.

We also compared our approach with a naive application of Peng and Huang (2008) by treating D as independent censoring. Figure 1 displays the mean

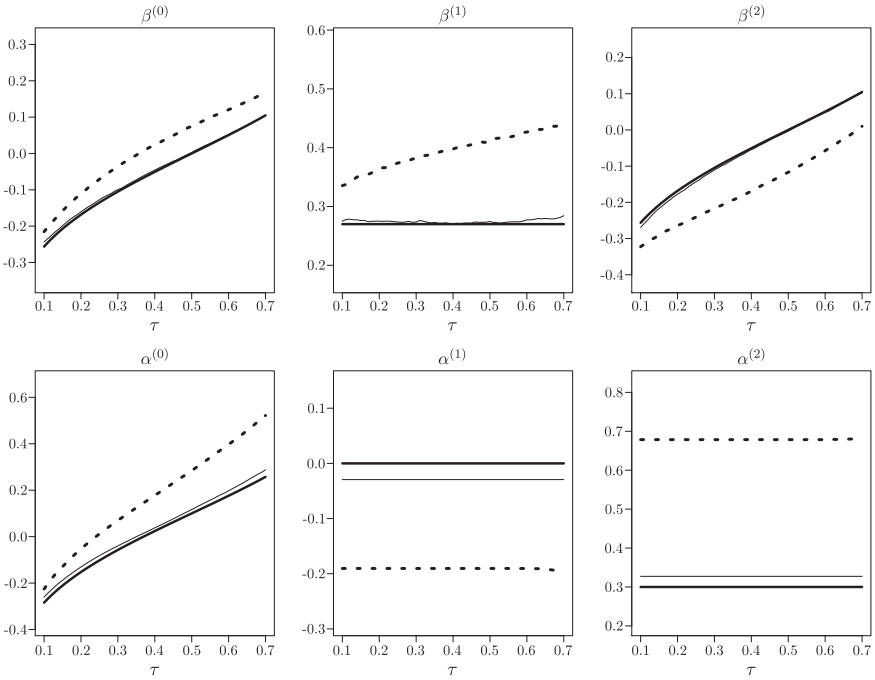

Figure 1. Upper panel: Comparison among true coefficients $\boldsymbol{\beta}_0(\tau)$ (bold solid lines), mean estimates for $\boldsymbol{\beta}_0(\tau)$ from the proposed method (solid lines) under a correctly specified Clayton copula, and mean estimates for $\boldsymbol{\beta}_0(\tau)$ from the naive approach (dotted lines); lower panel: Comparison among true coefficients $\boldsymbol{\alpha}_0(\tau)$ (bold solid lines), mean estimates for $\boldsymbol{\alpha}_0(\tau)$ from the proposed method (solid lines) under a correctly specified Clayton copula, and mean estimates for $\boldsymbol{\alpha}_0(\tau)$ from the naive approach (dotted lines).

estimated coefficients from the proposed approach and those from the naive approach, along with the true coefficients under a correctly specified Clayton copula, assuming an AFT model for D in both approaches. We can see that the proposed estimator $\hat{\boldsymbol{\beta}}(\tau)$ is virtually unbiased and $\hat{\boldsymbol{\alpha}}(\tau)$ has only small bias, while the naive approach can produce substantial bias, suggesting the importance of properly accounting for dependent censoring.

To assess the robustness of our methods, we also carried out estimation procedures with mis-specified copulas and compared the results to those under the correct copulas. Specifically, we focused on configuration (II), the case with 30% dependent censoring. With the true Kendall's tau set to be 0.576, we first generated T and D under the Clayton copula, and then estimated the regression coefficients assuming two types of dependence structure. One was the Frank copula with Kendall's tau= 0.576, which represents the situation of mis-specified copula function with correct degree of association, and the other was the Clayton

copula with Kendall's tau= $0.79, 0.33$ and 0.16, in which the copula function was true but the association parameters were not. Similarly, we also generated T and D under the Frank copula, and examined the estimation when assuming the Clayton copula with Kendall's tau= 0.576 and the Frank copula with Kendall's tau= $0.26, -0.12$ and -0.33, respectively.

Table E2 (Appendix E) summarizes the results when we mis-specified the copula function but correctly specified the association parameter, with the dependent censoring rate set to be 30%. Here the biases are still small and the coverage rates are again close to the nominal level. This suggests that, even with incorrect copula function, we can still obtain unbiased estimation if correct knowledge of the degree of association is accessible. In contrast, Figures E1 and E2 (Appendix E) depict the estimated coefficients for T under correctly specified copula forms with mis-specified association levels. Now the magnitude of the biases increases with the deviation of the assumed association from the true value. For example, when the underlying copula was Clayton with Kendall's tau= 0.576, the resulting biases may be moderate (as large as 0.05) for $\boldsymbol{\beta}_0(\tau)$ by assuming Kendall's tau= 0.79 or 0.33, and more pronounced (as large as 0.09) by assuming Kendall's tau= 0.16.

4. The WASID Example

We applied the proposed method to the WASID study, a double-blind and multicenter clinical trial that compared warfarin and aspirin in treating symptomatic intracranial arterial stenosis, an important cause of stroke. In this trial, 569 patients who had stroke or transient ischemic attack resulting from stenosis of a major intracranial artery were randomized to receive either warfarin or aspirin. In our analysis, T was defined as time from randomization to ischemic stroke, brain hemorrhage, or death, whichever happened first. We refer to this event as "study endpoint". During an average of 1.8-year follow-up, T was observed for 57 patients treated by warfarin and 60 patients treated by aspirin. For various reasons, the study medications were terminated early for 125 patients, among whom 81 were on the warfarin arm and 44 were on the aspirin arm. The follow-up of these patients continued while they received appropriate disease management determined by their primary physicians. The primary analysis reported in Chimowitz et al. (2005) followed an intent-to-treat (ITT) strategy: for patients whose assigned treatments were terminated early, no distinction was made between the follow-up information before and after the withdrawal.

To conduct a secondary on-treatment analysis that confers the treatment effect pertaining to the situation where the originally assigned treatment was not terminated early, our strategy was to censor the time to study endpoint at the time of early termination of study medication. We also looked into the possibility

of adopting a framework recently proposed for handling premature termination of treatment (Zhang et al. (2011)) that requires distinguishing mandatory discontinuation of study medication, and attaches the effect of interest to the dynamic treatment regimen that accounts for mandatory treatment discontinuation. However, per WASID protocol, there is no clear and definite rule to categorize mandatory termination and optional termination of study medication. Decisions for stopping study medication and post-withdrawal treatment were largely based on the discretion of physicians. We did not pursue the analysis in this direction.

For our analysis, one complication was that the withdrawals might be correlated with the underlying disease progression and thus pose dependent censoring to T. We let D denote dependent censoring time, the time from randomization to study withdrawal. Administrative censoring occurred for 146 patients in the warfarin group and 172 patients in the aspirin group. Time to such independent censoring was denoted by C. We considered three covariates: Treatment, 1 for warfarin and 0 for aspirin; Diabetes, the indicator of having diabetes; Stenosis Percentage, the percentage of stenosis by central reader.

We analyzed the WASID data based on some classical approaches, naively treating early drug termination as independent censoring. No treatment effect was detected by the log rank test. Adjusting for Diabetes and Stenosis Percentage, Cox regression also suggested that there was no significant treatment effect. The hazard ratio of warfarin versus aspirin was 0.91 with p-value=0.63. Stenosis Percentage was not found to be significant in predicting time to the study endpoint. Having diabetes was found to have a significant negative effect on the progression to the study endpoint. The corresponding hazard ratio and p value were 2.15 and < 0.001, respectively.

We applied the proposed regression approach, adjusting for the same set of covariates considered in the naive analysis. We specified different r values such that the corresponding Kendall's tau were 0.2, 0.4, 0.6, and 0.8, representing the cases where the positive associations between T and D were weak, moderate, and strong. The link function was chosen to be $\log(\cdot)$. Due to heavy censoring to D by T or C with the censoring rate around 80%, we adopted an AFT model for D to increase numerical stability. For inference, we performed 300 bootstrap resampling for each scenario. We considered both the Clayton and Frank copulas. We only present the results based on the Clayton copula, since the results under the Frank copula were similar.

Figure 2 depicts the estimates for $\boldsymbol{\beta}_0(\tau)$ under the Clayton copula, together with the results from a naive application of Peng and Huang (2008) in which D was treated as independent censoring. In Figure 2, the naive estimate and the proposed estimates for the treatment effect appear to be similar for $\tau < 0.18$ and demonstrate a larger yet moderate divergence for later τ's. In all cases,

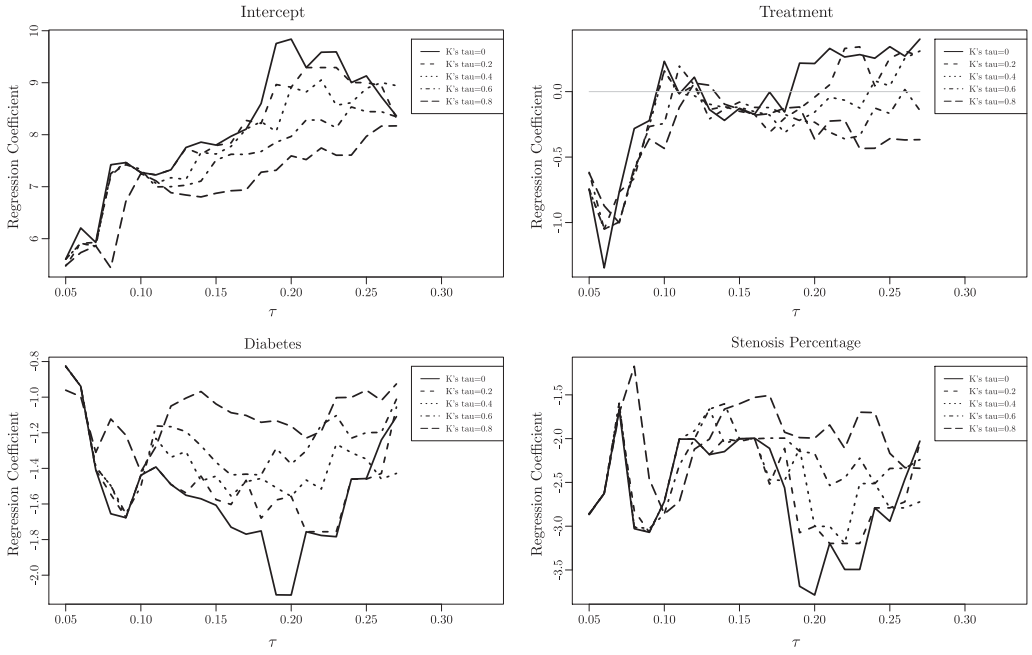

Figure 2. The WASID example: Point estimates of regression coefficients for time to the primary endpoint (ischemic stroke, brain hemorrhage, or death) under the Clayton copula with Kendall's tau=0, 0.2, 0.4, 0.6, and 0.8.

the estimated treatment effects over τ demonstrate a common pattern: negative at lower quantiles, then decreasing in the magnitude and becoming stabilized around 0. For Diabetes and Stenosis Percentage, the departure of the estimates that assume dependent censoring from the naive estimate are more noticeable.

In Table 2, we summarize the standard errors of the naive estimates and the proposed estimates under different specifications of $r > 0$. It can be seen that the proposed estimates have comparable efficiency to the naive estimate obtained by Peng and Huang (2008)'s method. We also performed the second-stage inference procedure on the WASID data. Formal tests on the significance of covariate effects were performed based on the average effects on quantiles of T with τ ranging from 0.05 to 0.25. Results show that the treatment effect was not significant for any choice of r we considered. This is consistent with Chimowitz et al. (2005), who found no benefit of warfarin over aspirin in the WASID trial. However, we found that Diabetes has significant effects under all choices of r (all p-values < 0.001), with average effects -1.58, -1.49, -1.38, -1.28, and -1.11, corresponding to the cases where Kendall's tau =0, 0.2, 0.4, 0.6, and 0.8. This suggests that the diabetic patients may progress significantly faster to the study endpoint compared to nondiabetic patients. This finding is consistent with the

Table 2. The WASID example: standard errors under the Clayton copula with Kendall's tau=0, 0.2, 0.4, 0.6, and 0.8. $\hat{\beta}^{(1)}$, $\hat{\beta}^{(2)}$ and $\hat{\beta}^{(3)}$ are estimated coefficients of Treatment, Diabetes, and Stenosis Percentage on T, respectively.

τ		K's tau=0	K's tau=0.2	K's tau=0.4	K's tau=0.6	K's tau=0.8
0.05	$\hat{\beta}^{(1)}$	0.79	0.77	0.78	0.76	0.72
	$\hat{\beta}^{(2)}$	0.77	0.75	0.73	0.73	0.69
	$\hat{\beta}^{(3)}$	2.10	2.04	2.04	2.04	1.95
0.10	$\hat{\beta}^{(1)}$	0.69	0.69	0.69	0.71	0.74
	$\hat{\beta}^{(2)}$	0.65	0.65	0.64	0.66	0.69
	$\hat{\beta}^{(3)}$	1.83	1.67	1.79	1.76	1.73
0.15	$\hat{\beta}^{(1)}$	0.43	0.41	0.37	0.36	0.35
	$\hat{\beta}^{(2)}$	0.41	0.38	0.37	0.36	0.39
	$\hat{\beta}^{(3)}$	1.45	1.38	1.33	1.26	1.17
0.20	$\hat{\beta}^{(1)}$	0.51	0.46	0.40	0.39	0.33
	$\hat{\beta}^{(2)}$	0.60	0.48	0.44	0.34	0.32
	$\hat{\beta}^{(3)}$	1.76	1.48	1.42	1.24	1.05
0.25	$\hat{\beta}^{(1)}$	0.53	0.50	0.46	0.41	0.35
	$\hat{\beta}^{(2)}$	0.79	0.66	0.54	0.45	0.33
	$\hat{\beta}^{(3)}$	2.24	1.85	1.67	1.46	1.19

naive Cox regression analysis, but is better endorsed by taking into account the potential dependence between T and D.

To illustrate the impact of adjusting for dependent censoring in a more meaningful way, we plot the estimated quantiles of T and D (see Figure 3 and Figure E3 in Appendix E) for each treatment group with and without diabetes, with Stenosis Percentage fixed at its mean value. From Figure E3, it is apparent that the disparity among different estimates is negligible in the diabetes group, but accounting for dependent censoring at different levels can lead to dramatically different estimates for $Q_T(\tau|\boldsymbol{Z})$ in the non-diabetic group. One plausible explanation for this is that non-diabetic patients generally progress to the study endpoint slower than diabetic patients and thus are more prone to the "risk" of early termination of study medication. Consequently, adjusting for dependent censoring for the non-diabetic patients makes a bigger influence on the estimated quantiles of T. From Figure 3, assuming independence between patient withdrawal and the study endpoint tends to give more optimistic estimate for $Q_T(\tau|\boldsymbol{Z})$ compared to cases where T and D were assumed to be positively associated. This is reasonable and intuitive explanation is that an observed D, meaning $T > D$ and T is censored, would be suggestive of a smaller T when T and D are believed to be positively associated than that under independence between T and D. As

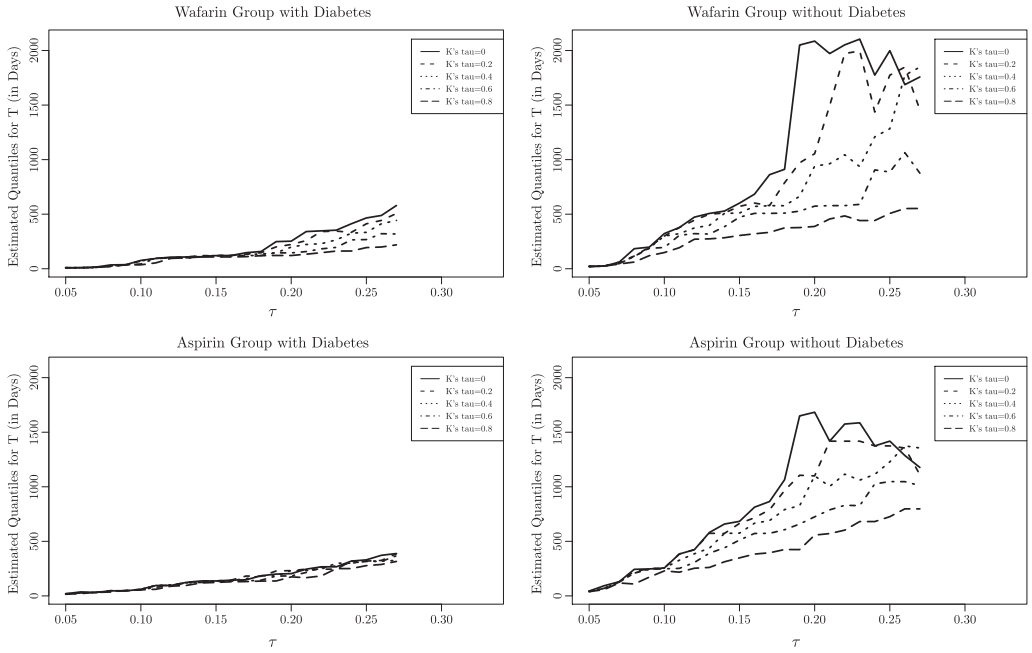

Figure 3. The WASID example: Estimated quantiles of time to the primary endpoint (ischemic stroke, brain hemorrhage, or death) under the Clayton copula with Kendall's tau=0, 0.2, 0.4, 0.6, and 0.8, with the Stenosis Percentage fixed at its mean (63.7%)

a result, the prediction of $Q_T(\tau|\boldsymbol{Z})$ would be more conservative under a positive association assumption. From Figure E3, the warfarin group tends to have smaller D compared to the aspirin group, which means the patients treated by warfarin tend to withdraw earlier than the other group. This is also consistent with Chimowitz et al. (2005), who found a higher rate of adverse events in the warfarin group than in the aspirin group.

In summary, we found no evidence of better clinical efficacy for warfarin compared to aspirin in treating symptomatic intracranial arterial stenosis, which is consistent with previously published results on this trial. In our analysis, we took into account of the dependence between T and D and provided a comprehensive view of the covariate effects under different specifications of the association.

5. Remarks

We propose a quantile regression method for survival data subject to dependent censoring. Under the assumed model for the event time of interest, covariate effects are formulated on the quantiles defined on the marginal survival distribution. We model the dependence structure between the survival time and

the censoring via a copula model. Such an assumption is not verifiable based on observed data, so caution is needed when applying the proposed method. Extensive numerical studies show that the proposed estimation is quite robust to misspecification of the parametric class of the adopted copula, provided the strength of association is reasonably specified.

Some earlier works (Moeschberger (1974); Link (1989); Emoto and Matthews (1990), among others) directly specified the bivariate distribution by parametric or semi-parametric models. Other works (Robins and Rotnitzky (1992); Robins (1993); Robins and Finkelstein (2000); Scharfstein et al. (2001); Scharfstein and Robins (2002)) proposed estimation based on inverse probability of censoring weighting (IPCW) with the flexibility to accommodate time-dependent prognostic factors. Analyses of dependently censored data can also be focused on identifiable *crude quantities* to provide inference in settings that do not exclude censoring events (Gray (1988); Pepe (1991); Lin (1997); Fine and Gray (1999)). Considering *crude* quantities and *net* quantities offers alternative perspectives of survival endpoint of interest, depending on context (Jiang, Chappell, and Fine (2003)). The proposal here concerns *net* conditional quantiles of T, as motivated by the WASID example, and provides a useful complement to existing methods for dependently censored data.

Supplemental Material

Appendices A–E referenced in Sections 2–4 are available online at `http://www3.stat.sinica.edu.tw/statistica`

Acknowledgement

The authors are grateful to the Editor, an associate editor, and two referees for many helpful comments. This work has been supported by National Science Foundation grant DMS-1007660 and by the National Heart, Lung, and Blood Institute of Health under Award number R01HL 113548.

References

Chen, Y.-H. (2010). Semiparametric marginal regression analysis for dependent competing risks under an assumed copula. *J. Roy. Statist. Soc. Ser. B* **72**, 235-251.

Chimowitz, M., Lynn, M., Howlett-Smith, H., Stern, B., Hertzberg, V., Frankel, M., Levine, S., Chaturvedi, S., Kasner, S., benesch, C., Sila, C., Jovin, T. and Romano, J. (2005). Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *The New England J. Medicine* **352**, 1305-1316.

Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141-151.

Dickson, E. R., Grambsch, P. M., Fleming, T. R., Fisher, L. D. and Langworthy, A. (1989). Prognosis in primary biliary cirrhosis: Model for decision making. *Hepatology* **10**, 1-7.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist.* **7**, 1-26.

Emoto, S. E. and Matthews, P. C. (1990). A Weibull model for dependent censoring. *Ann. Statist.* **18**, 1556-1577.

Fine, J. P. and Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *J. Amer. Statist. Assoc.* **94**, 496-509.

Fine, J. P., Yan, J. and Kosorok, M. (2004). Temporal process regression. *Biometrika* **91**, 683-703.

Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika* **74**, 549-555.

Gray, R. J. (1988). A class of k-sample tests for comparing the cumulative incidence of a competing risk. *Ann. Statist.* **16**, 1141-1154.

Honoré, B., Khan, S. and Powell, J. (2002). The identifiability of the competing risks model. *J. Econometrics* **109**, 67-105.

Huang, Y. (2010). Quantile calculus and censored regression. *Ann. Statist.* **38**, 1607-1637.

Huang, X. and Zhang, N. (2008). Regression survival analysis with an assumed copula for dependent censoring: a Sensitivity analysis approach. *Biometrics* **64**, 1090-1099.

Jensen, G., Torp-Pedersen, C., Hildebrandt, P., Kober, L., Nielsen, F., Melchior, T., Joen, T. and Andersen, P. (1997). Does in-hosipital fibrillation affect prognosis after myocardial infarction? *European Heart J.* **18**, 919-924.

Jiang, H., Chappell, R. and Fine, J. P. (2003). Estimating the distribution of nonterminal event time in the presence of mortality or informative dropout. *Controlled Clinical Trials* **24**, 135-146.

Jin, Z., Lin, D. Y., Wei, L. J. and Ying, Z. (2003). Rank-based inference for the accelerated failure time model. *Biometrika* **90**, 341-353.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* 2nd edition. Wiley, New York.

Kaslow, R., Ostrow, D., Detels, R., Phair, J., Polk, B. and Rinaldo, C. (1987). The multicenter aids cohort study: rationale, organization and selected characteristics of the participants. *Amer. J. Epidemiology* **126**, 310-318.

Klein, J. P. and Moeschberger, M. L. (1988). Bounds on net survival probabilities for dependent competing risks. *Biometrics* **44**, 529-538.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

Lin, D. Y. (1997). Non-parametric inference for cumulative incidence functions in competing risks studies. *Statist. Med.* **16**, 901-910.

Link, W. A. (1989). A model for informative censoring. *J. Amer. Statist. Assoc.* **84**, 749-752.

Moeschberger, M. L. (1974). Life tests under dependent competing causes of failure. *Technometrics* **16**, 39-47.

Oakes, D. (1989). Bivariate survival models induced by frailties. *J. Amer. Statist. Assoc.* **84**, 487-493.

Peng, L. and Fine, J. P. (2009). Competing Risks Quantile Regression. *J. Amer. Statist. Assoc.* **104**, 1440-1453.

Peng, L. and Huang, Y. (2008). Survival analysis with quantile regression models. *J. Amer. Statist. Assoc.* **103**, 637-649.

Pepe, M. S. (1991). Inference for events with dependent risks in multiple endpoint studies. *J. Amer. Statist. Assoc.* **86**, 770-778.

Peterson, A. V. (1976). Bounds for a joint distribution function with fixed subdistribution functions: application to competing risks. *Proc. Nat. Acad. Sci. USA* **73**, 11-13.

Portnoy, S. (2003). Censored regression quantiles. *J. Amer. Statist. Assoc.* **98**, 1001-1012.

Powell, J. (1984). Least absolute deviations estimation for the censored regression model. *J. Econometrics* **25**, 303-325.

Powell, J. (1986). Censored regression quantiles. *J. Econometrics* **32**, 143-155.

Robins, J. M. (1993). Information recovery and bias adjust-ment in proportional hazards regression analysis of randomized trials using surrogate markers. In *ASA Proceedings of the Biopharmaceutical Section*, 24-33. American Statistical Association, Alexandria, Virginia.

Robins, J. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology: Methodological Issues* (Edited by N. Jewell and K. Dietz), 297-331. Birkhäiuse, Boston.

Robins, J. M. and Finkelstein, D. M. (2000). Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* **56**, 779-778.

Scharfstein, D. O. and Robins, J. M. (2002). Estimation of the failure time distribution in the presence of informative censoring. *Biometrika* **89**, 617-634.

Scharfstein, D. O., Robins, J. M., Eddings, W. and Rotnitzky, A. (2001). Inference in randomized studies with informative censoring and discrete time-toevent endpoints. *Biometrics* **57**, 404-413.

Slud, E. V. and Rubinstein, L. V. (1983). Dependent competing risks and summary survival curves. *Biometrika* **70**, 643-649.

Thorogood, J., Persijn, G., Schreuder, G., D'amaro, J., Zantvoort, F., Van Houwelingen, J. and Van Rood, J. (1990). The effect of hla matching on kidney graft survival in separate posttransplantation intervals. *Transplantation* **50**, 146-150.

Tsiatis, A. A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Nat. Acad. Sci. USA* **72**, 20-22.

Verweij, P. and Van Houwelingen, H. (1995). Time-dependent effects of fixed covariates in cox regression. *Biometrics* **51**, 1550-1556.

Wang, H. J. and Wang, L. (2009). Locally weighted censored quantile regression. *J. Amer. Statist. Assoc.* **104**, 1117-1128.

Yan, J. (2007). Enjoy the joy of copulas: with a package copula. *J. Statist. Software* **21**, 1-21.

Yang, S. (1999). Censored median regression using weighted empirical survival and hazard functions. *J. Amer. Statist. Assoc.* **94**, 137-145.

Ying, Z., Jung, S. H. and Wei, L. J. (1995). Survival analysis with median regression models. *J. Amer. Statist. Assoc.* **90**, 178-184.

Zhang, M., Tsiatis, A. A., Davidian, M., Pieper, K. S. and Mahaffey, K. W. (2011). Inference on treatment effects from a randomized clinical trial in the presence of premature treatment discontinuation: the SYNERGY trial. *Biostatistics* **12**, 258-269.

Zheng, M. and Klein, J. P. (1995). Estimates of marginal survival for dependent competing risks based on an assumed copula. *Biometrika* **82**, 127-138.

Division of Biostatistics, University of Texas Health Science Center at Houston, Houston, TX 77030, USA.

E-mail: Shuang.Ji@uth.tmc.edu

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA.

E-mail: lpeng@emory.edu

Department of Biostatistics, University of Pittsburgh, Pittsburgh, PA 15260, USA.

E-mail: rul12@pitt.edu

Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, USA.

E-mail: mlynn@sph.emory.edu

(Received October 2012; accepted September 2013)