Published in final edited form as:

Biostat Epidemiol. 2018 ; 2(1): 84–98. doi:10.1080/24709360.2018.1445430.

Semiparametric Additive Model for Estimating Risk Difference in Multicenter Studies

DONGLIN ZENG¹, NOORIE HYUN², and JIANWEN CAI³

¹Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599

²Division of Biostatistics, Institute for Health and Society, Medical College of Wisconsin, Milwaukee, WI, 53226

³Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599

SUMMARY

Many cancer studies are conducted in multiple centers. While they have the advantage of more patients and larger population, center-to-center heterogeneity could be significant such that it cannot be ignored in analysis. In this paper, we propose semiparametric additive risk models with a general link function to estimate risk effects while accounting for center-specific baseline function. We propose an estimating equation for inference and show that the derived estimators are consistent and asymptotically normal. Simulation studies demonstrate good small-sample performance of the proposed method. We apply the method to analyze data from the Study of Left Ventricular Dysfunction (SOLVD) in 1990 and discuss application to one-to-one matched design.

Keywords

Additive risk models; Estimating equation; multi-center study; one-to-one matched design; proportional hazards model; recurrent event

1 Introduction

Many cancer studies are conducted in multiple centers, with advantages to not only enroll sufficient number of patients but also enable investigators to assess risk effects in larger populations. However, since significant heterogeneity could exist among patients from different centers due to center-specific characteristics, simply treating combined data as from a homogeneous population can lead to severely biased risk effect estimates and reduced power for evaluating their significance. Therefore, a valid inference requires a flexible risk model and estimation procedure to account for the heterogeneity from multicenter studies.

Commonly used method to allow center-specific baselines is stratified proportional hazards regression model which assumes that the hazard ratios between covariates are constants but baseline functions differ among centers (Holt and Prentice, 1974; Huster, Brookmeyer and

Self, 1989; Lee, Wei and Amato, 1992; Cai, Zhou and Davis, 1997; Glidden and Vittinghoff, 2004). Estimation of the covariate effects is based on maximizing the summation of log-partial likelihood functions using data from each center. Alternatively, Lee, Wei and Ying (1993) considered a stratified accelerated failure time model, where residue terms in the model may have different distributions for each center but within the same center, they are assumed to be exchangeable and possess the same marginal distribution. Based on the latter property, estimating equations using some linear rank statistics were first constructed for each center and then combined for parameter estimation.

In practice, when evaluating the effect of some risk factors, risk difference is often preferred to hazards ratios due to its direct interpretation in terms of the number of events changed with the change of risk factor levels. This is particularly meaningful when events are rare. To estimate risk difference, many literature have discussed fitting additive risk models for either modelling survival events or recurrent events (e.g., Aalen, 1980; Breslow and Day, 1980; Pocock, Gore and Kerr, 1982; Cox and Oakes, 1984; Buckley, 1984; Breslow and Day, 1987; Aalen, 1989; Huffner and McKeague, 1991; Andersen et al, 1993; Lin and Ying, 1994; Lin, Oakes and Ying, 1998; Schaubel, Zeng and Cai, 2006; Martinussen and Scheike, 2002; Pan, He and Song, 2015; Ding and Sun, 2015). Particularly, the method in Lin and Ying (1994) considered an additive risk model for a marginal intensity model with independent observations; Schaubel et al (2006) applied such a model to model the rate function of recurrent events. None of them has studied the additive risk models using data from multicenter studies.

In this paper, we propose semiparametric additive risk models to study the covariate effects from multicenter studies while accounting for center heterogeneity via center-specific baseline functions. We assume constant risk differences but allow different baseline functions across centers. Moreover, our models are imposed upon a general counting process so they are applicable to both survival endpoints and recurrent events. In Section 2, we introduce a semiparametric additive risk model by assuming covariate effects to be constant across centers and describe the corresponding inference procedure. We also give the asymptotic properties of the proposed estimators in Section 2. In Section 3, simulation studies are conducted and the proposed method is applied to analyze a real data example. In Section 4, as an extreme example, we examine power analysis in the one-to-one matched design. Some concluding remarks and further generalization are given in Section 5.

2 Semiparametric Additive Risk Model and Inference with Fixed Effects

We consider n independent centers, each with n_i subjects, $i = 1, \dots, n$. For the j th subject in the i th center, let N_{ij}^* denote the counting process of event time and $z_{ij}(t)$ denote its covariates. To capture the difference between centers, we denote C_i as the i th center-specific variables which may not be observable. We further let C_{ij} denote the potential censoring time for subject j in center i and assumed C_{ij} is independent of N_{ij}^* conditional on C_i and z_{ij} .

Our model assumes that the rate function of N_{ij}^* given z_{ij} and C_i satisfies

$$E[dN_{ij}^*(t) | Y_{ij}(t), z_{i\cdot}, C_i] = Y_{ij}(t) \{d\Lambda(t; C_i) + \psi(z_{ij}(t)^T \beta) dt\}, \quad (1)$$

where $Y_{ij}(t)$ is the at-risk process, z_i denotes the covariates of all the subjects in center i , $\Lambda(t; C_i)$ is the center-specific hazard rate function, and ψ is a known link function. Particularly, if $N_{ij}^*(t)$ is the counting process associated with a survival outcome, then $Y_{ij}(t)$ indicates whether the subject is still alive at time t ; if $N_{ij}^*(t)$ is the counting process associated with a recurrent event, then $Y_{ij}(t)$ indicates whether the subject may experience this event at time t (usually, $Y_{ij}(t) = 1$). Note that in model (1), we do not impose any structure on $\Lambda(t; C_i)$ so $\Lambda(t; C_i)$ can be very different from center to center. Parameter β represents additive effect of the covariate factors at the scale defined by the link function ψ . For example, when $\psi(x) = x$, this effect is additive on the event rate; when $\psi(x) = \exp(x)$, the additive effect is then on the log-scale of the event rate. Moreover, model (1) implies that the center-specific event rate from center i is fully captured by the center-specific baseline function.

Let τ be the study duration and C_{ij} be the potential censoring time for subject j in center i . Then $R_{ij}(t) = Y_{ij}(t)I(C_{ij} \wedge \tau - t)$ is the observed at-risk process. The observations from all the subjects are given as $(N_{ij}(t) = N_{ij}^*(t)R_{ij}(t), R_{ij}(t), z_{ij}), i = 1, \dots, n, j = 1, \dots, n_i$. Our goal is to estimate β . To this end, let $\Lambda_0(t; C_i)$ and β_0 denote the true value of $\Lambda(t; C_i)$ and β respectively. It is clear that

$$E[dN_{ij}(t) | R_{ij}(t), z_{i\cdot}, C_i] = R_{ij}(t)d\Lambda_0(t; C_i) + R_{ij}(t)\psi(z_{ij}(t)^T \beta_0)dt. \quad (2)$$

Therefore, one intuitive way is to treat the above expression as a simple regression problem where $d\Lambda_0(t; C_i)$ is the jump size of $\Lambda_0(t; C_i)$ at time t , so for a fixed β , we can estimate $d\Lambda_0(t; C_i)$ only using the data from center i as

$$\frac{\sum_{j=1}^{n_i} R_{ij}(t)(dN_{ij}(t) - \psi(z_{ij}(t)^T \beta)dt)}{\sum_{j=1}^{n_i} R_{ij}(t)}.$$

Equivalently, we estimate $\Lambda_0(t; C_i)$ using

$$\int_0^t \frac{\sum_{j=1}^{n_i} R_{ij}(s)(dN_{ij}(s) - \psi(z_{ij}(s)^T \beta)ds)}{\sum_{j=1}^{n_i} R_{ij}(s)}.$$

After substituting the above expression back into (2), we obtain

$$E\left[dN_{ij}(t) \middle| R_{ij}(t) = 1, z_{i\cdot}, C_i\right] = R_{ij}(t) \frac{\sum_{k=1}^{n_i} R_{ik}(t)(dN_{ik}(t) - \psi(z_{ik}^T \beta)dt)}{\sum_{k=1}^{n_i} R_{ik}(t)} + R_{ik}(t)\psi(z_{ij}(t)^T \beta)dt.$$

Thus, we propose the following class of estimating equations for estimating β :

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \\ \times \left\{ dN_{ij}(t) - \psi(z_{ij}(t)^T \beta)dt - \frac{\sum_{k=1}^{n_i} R_{ik}(t)(dN_{ik}(t) - \psi(z_{ik}(t)^T \beta)dt)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} = 0,$$

where $\omega_i(t)$ is any given center-specific weight function. Since

$$\sum_{j=1}^{n_i} R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} = 0,$$

the estimating equation is equivalent to

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \\ \times \left\{ dN_{ij}(t) - \psi(z_{ij}(t)^T \beta)dt \right\} = 0. \quad (3)$$

Under the special choice of $\psi(x) = x$, equation (3) explicitly gives the estimator

$$\hat{\beta}_n = \left[\sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} z_{ij}(t)^T dt \right]^{-1} \\ \times \left[\sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} dN_{ij}(t) \right].$$

In the appendix, we show that under some regularity conditions, $\sqrt{n}(\hat{\beta}_n - \beta_0)$ converges in distribution to a normal distribution with mean zero. The asymptotic covariance can be consistently estimated by

$$\begin{aligned}
& \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\}^{\otimes 2} dt \right]^{-1} \\
& \times \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \right. \\
& \quad \left. \times (dN_{ij}(t) - \psi(z_{ij}(t)^T \hat{\beta}_n) dt) \right]^{\otimes 2} \\
& \times \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\}^{\otimes 2} dt \right]^{-1}.
\end{aligned}$$

Thus, any inference for $\hat{\beta}$, such as testing $\beta_0 = 0$ and constructing confidence interval for β_0 , can be based on this asymptotic normal distribution.

3 Numerical Study

We have conducted a number of simulation studies with survival outcomes to examine the performance of our proposed estimator under model (1). In each simulation study, we generated n centers with different sizes. Within center i , two fixed covariates were generated for each subject, denoted as (Z_{1ij}, Z_{2ij}) : Z_{1ij} was from a Benoulli variable with probability 0.5; $Z_{2ij} = Z_{1ij}(U_{ij} + i/n)$, where U_{ij} was independently generated from a uniform distribution in $[0, 1]$. The covariate Z_2 means the center variable conditional on $Z_1 = 1$. The true covariate effects were chosen to be 0.5 and 1.0 respectively. For different simulation studies, different baseline hazard rate functions were used for different centers and the link function $\psi(x)$ was equal to x or $\exp(x)$. The right-censoring time was generated from a uniform distribution in $[0, 4]$ and τ was set to 2, yielding a censoring rate varying from 10% to 25%.

Table 1 and Table 2 summarize the results from all the simulation studies with $n = 25, 50, 100, 200$ based on 1000 iterations, each table corresponding to different link function $\psi(x)$. The results show that for $\psi(x) = x$, i.e., the standard additive hazards model, the estimators perform well when the number of centers is at least 50: the bias is small and the inference is correct; when $\psi(x) = \exp(x)$, the estimators still perform well even if the number of the centers is only 25. All the simulations converged for $\psi(x) = x$; however, when the link function $\psi(x) = \exp(x)$, we report that about 12% simulation did not converge with small number of centers $n = 25$ and small cluster size $m = 2$; about 5% did not converge for $n = 50$ and $m = 2$; about 3% did not converge for $n = 25$ and $m = 6$; the non-convergence rate was below 2% for $n = 100$ and $m = 2$; while all the other settings with non-small center sizes had no convergence problem. Thus, we conclude that models with nonlinear link function are appropriate for multicenter-studies when the center sizes are moderate in order to obtain stable results.

To further examine the robustness and efficiency of the proposed method, we conducted an additional simulation study with $\psi(x) = x$ to compare our estimator to the estimator ignoring the heterogeneity across the centers. The latter was obtained using the method in Lin and Ying (1994) and was expressed as

$$\left[\sum_{i=1}^n \sum_{j=1}^{n_i} \int R_{ij}(t) \{Z_{ij}(t) - \bar{Z}(t)\}^{\otimes 2} dt \right]^{-1} \left[\sum_{i=1}^n \sum_{j=1}^{n_i} \{Z_{ij}(t) - \bar{Z}(t)\} dN_{ij}(t) \right],$$

where

$$\bar{Z}(t) = \frac{\sum_{i=1}^n \sum_{j=1}^{n_i} R_{ij}(t) Z_{ij}(t)}{\sum_{i=1}^n \sum_{j=1}^{n_i} R_{ij}(t)}.$$

In the simulation study, we considered $n = 100$ and 200 with different cluster sizes as in Table 1. The results from 1000 repetitions are summarized in Table 3. From the table, we find that when the marginal baseline hazards are the same among the centers, i.e., this refers to the first scenario since $E[dN_{ij}(t)|Z_{ij}] = (t + Z_{ij}^T \beta_0) dt$, our estimator is less efficient than Lin and Ying's but such inefficiency vanishes as the center size increases; when the marginal baseline hazards are different among the centers, Lin and Ying's estimator has very large bias due to violation of the assumption that the baseline hazard functions are identical across centers. The superiority of our estimator to Lin and Ying's estimator becomes more significant as the center size increases.

We further apply our method to analyze data from SOLVD (SOLVD Investigators, 1990) Treatment Trial. This study was a randomized, double-masked, placebo-controlled trial conducted between 1986 and 1991. The participants were of age 21 to 80 years old, inclusive, with overt symptoms of congestive heart failure and left ventricular ejection fraction less than 35%. The latter is a measure of the efficiency of the heart in ejecting blood and is a number between 0 and 100%. The study was done at 23 medical centers in US, Canada and Belgium and the average number of patients per center was slightly over 100. The event of interest was the number of years to the first hospitalization for congestive heart failure or death from all causes, whichever happened first. The goal was to examine the effect of treatment by enalapril versus placebo but the participants' age, gender and ejection fraction could be potential confounders so should also be adjusted for in the analysis. We fit our proposed model to this data and the results are summarized in Table 4: it shows that the estimates and the inference are not sensitive to the choice of the link function. Both results show that the treatment significantly reduced the risk of the first hospitalization due to the heart failure or death; the elder patients tended to have higher risk than the younger ones; the higher ejection fraction was significantly associated with lower risk; however, there was no significant difference between genders. Our findings are similar to Cai et al (1999) but at different scales.

4 Application to One-to-One Matched Design

We apply the proposed method to study power analysis in an extreme situation of multicenter studies based on one-to-one matched design. One-to-one matched design, in which one subject with condition is matched to another subject without condition in order to study the effect of the condition on some outcome, is commonly used in practice to remove potential confounding effects, for example, cross-over design, a special case of such a matched design, is often used in clinical trials to compare pre- and post-treatment outcomes (Jones and Kenward, 2015). By treating the matched group as one center, the data from this design can be viewed as one special case of the multi-center study considered in this paper.

In a one-to-one matched study, a center contains two subjects with one assigned to treatment and the other assigned to control. We code the first subject's covariate as 1 while the other 0. Assume the usual additive risk model with $\psi(x) = x$. We aim to derive the sample size or power formula for comparison between treatment and control under the alternative hypothesis in which $\beta_0 = \delta/\sqrt{n}$.

We let $\omega_i(t) = 1$ and assume that the censoring time for the matched subjects will be independent and with the same survival function $S_c(t)$. From the expression in the appendix, the asymptotic covariance for $\sqrt{n}\hat{\beta}_n$ can be approximated by

$$\begin{aligned} & E \left[\frac{1}{n} \sum_{i=1}^n \int \left[I(Y_{1i} \geq t) \left(1 - \frac{I(Y_{1i} \geq t)}{I(Y_{1i} \geq t) + I(Y_{2i} \geq t)} \right) \right. \right. \\ & \quad \left. \left. + I(Y_{2i} \geq t) \left(\frac{I(Y_{1i} \geq t)}{I(Y_{1i} \geq t) + I(Y_{2i} \geq t)} \right) \right]^2 dt \right]^{-2} \\ & \times E \left[\frac{1}{n} \sum_{i=1}^n \left[\int \frac{I(Y_{1i} \wedge Y_{2i} \geq t)}{I(Y_{1i} \geq t) + I(Y_{2i} \geq t)} (dN_{1i}(t) - dN_{2i}(t)) \right]^2 \right], \end{aligned}$$

where Y_{1i} and Y_{2i} denote the last observed event times of the first and the second subjects in the i th pair, respectively. Further simplification gives that the above expression is equal to

$$\frac{E \left[n^{-1} \sum_{i=1}^n \{N_{1i}(Y_{1i} \wedge Y_{2i}) - N_{2i}(Y_{1i} \wedge Y_{2i})\}^2 \right]}{E \left[n^{-1} \sum_{i=1}^n Y_{1i} \wedge Y_{2i} \right]^2}.$$

When the time-to-event is survival and no pairs have the events at the same time, $[N_{1i}(Y_{1i} \wedge Y_{2i}) - N_{2i}(Y_{1i} \wedge Y_{2i})]^2 = I_{1i}(Y_{1i} < Y_{2i}) + I_{2i}(Y_{1i} > Y_{2i})$, where I_{ji} is the censoring indicator for subject j in center i . Thus, we conclude that under the alternative hypothesis,

$$\hat{\beta} \sim \text{Normal} \left(\frac{\delta}{\sqrt{n}}, \frac{\bar{\Delta}_n}{n\bar{Y}_n^2} \right),$$

where

$$\bar{\Delta}_n = \frac{1}{n} \sum_{i=1}^n [\Delta_{1i} I(Y_{1i} < Y_{2i}) + \Delta_{2i} I(Y_{1i} < Y_{2i})],$$

i.e., the average number of first failures in each center, and

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_{1i} \wedge Y_{2i},$$

i.e., the average length of $Y_{1i} \wedge Y_{2i}$ across all the centers. As the result, the power under the alternative hypothesis is given by

$$1 - \Phi \left(\Phi_{1-\alpha}^{-1} - \delta \bar{Y}_n / \sqrt{\bar{\Delta}_n} \right)$$

where Φ is the cumulative distribution function of the standard normal distribution and α is the type I error. Particularly, if under the null hypothesis, Y_{1i} and Y_{2i} are exchangeable, then $\bar{\Delta}_n$ can be approximated by

$$\frac{1}{2n} \sum_{i=1}^n (\Delta_{1i} + \Delta_{2i})$$

and $\bar{Y}_n$ can be approximated by

$$\frac{2}{n} \sum_{i=1}^n Y_{1i} I(Y_{1i} < Y_{2i}),$$

i.e., the average of the events across centers.

As one example, we plot in Figure 1 the power curves associated with different numbers of clusters in the third setting of Table 1. From the plot, we conclude that to achieve the power of 80%, the minimal effect sizes to be detected are in turn about 0.65, 0.45, 0.3 and 0.2 for $n = 25, 50, 100$ and 200 .

5 Discussion

We have proposed an additive risk model for analyzing event data from multiple centers where the center variation can be dramatic. We also provide a simple power formula for one-to-one matched design where matching variables can be unknown. For example, the matching variables can be many and are often not immediately available to analysts.

Our model allows nonlinear link functions when covariate effects are not additive. Our estimators for the linear link performed well when either the number of clusters or the average cluster size was not small. When the link function was nonlinear and the center size

was small, the estimates were relatively unstable due to low event rates. A larger center size may be necessary in this situation. Furthermore, in our approach, we assume the cluster size is bounded but the number of clusters goes to infinity. However, it is also seen in many studies that the cluster size can be very large. When this happens, we can still show that the proposed estimators are consistent since the proposed estimation equation remains unbiased at the true value of the parameters; however, the asymptotic covariance can be complicated and depend on the within-cluster correlations. Different consideration should be taken for the case when the number of the clusters is bounded but the cluster size goes to infinity or the case when both the number of the clusters and the cluster size go to infinity at some rates.

Model (1) assumes that after adjusting for baseline information for each center, the covariate effects are constant across all the centers. Since the constant effect across center may be questionable in practice, we generalize model (1) to the following model:

$$E[dN_{ij}^*(t) | Y_{ij}(t), z_{ij}(t), C_i] = Y_{ij}(t)(d\Lambda(t; C_i) + \psi(z_{ij}(t))^T(\beta + \xi_i))dt, \quad (4)$$

where ξ_i is center-specific random effect and is assumed to follow a parametric distribution $\phi(\cdot; \gamma)$ with parameter γ . For identifiability, we assume that ξ_i has mean zero. Therefore, in the more general model, we allow the covariate effect within each center to have a random variation; equivalently, a random interaction between covariates and center is incorporated in model (4). To estimate β and γ , as before, we can first estimate $d\Lambda(t; C_i)$ using

$$\frac{\sum_{j=1}^{n_i} R_{ij}(t)(dN_{ij}(t) - \psi(z_{ij}(t))^T(\beta + \xi_i))dt}{\sum_{j=1}^{n_i} R_{ij}(t)}.$$

After substituting the above expression into a similar expression to (2), we obtain

$$\begin{aligned} E[dN_{ij}(t) | R_{ij}(t), z_{ij}(t), C_i] &= R_{ij}(t) \frac{\sum_{k=1}^{n_i} R_{ik}(t) dN_{ik}(t)}{\sum_{k=1}^{n_i} R_{ik}(t)} \\ &\quad + R_{ij}(t) \left[\psi(z_{ij}(t))^T(\beta + \xi_i) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \psi(z_{ik}(t))^T(\beta + \xi_i)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right] dt. \end{aligned}$$

To remove the random variable ξ_i , we integrate both sides over ξ with respect to its density and obtain

$$E\left[dN_{ij}(t)\middle|R_{ij}, z_i, C_i\right] = R_{ij}(t) \frac{\sum_{k=1}^{n_i} R_{ik}(t) dN_{ik}(t)}{\sum_{k=1}^{n_i} R_{ik}(t)} + R_{ij}(t) \left[\tilde{\psi}(z_{ij}(t); \beta, \gamma) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \tilde{\psi}(z_{ik}(t); \beta, \gamma)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right] dt,$$

Where

$$\tilde{\psi}(z_{ij}(t); \beta, \gamma) = E_{\xi_i} \left[\psi(z_{ij}(t)^T (\beta + \xi_i)) \right].$$

We thus obtain the same expression as given in Section 2, except that $\psi(z_{ij}(t)^T \beta)$ is now replaced by $\tilde{\psi}(z_{ij}(t); \beta, \gamma)$. Hence, similar to equation (3), we can propose the following estimating equation for estimating β and γ :

$$\sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left[\nabla \tilde{\psi}(z_{ij}(t); \beta, \gamma) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \nabla \tilde{\psi}(z_{ik}(t); \beta, \gamma)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right] dt \quad (5) \\ \times \left\{ dN_{ij}(t) - R_{ij}(t) \tilde{\psi}(z_{ij}(t); \beta, \gamma) \right\} = 0,$$

where $\nabla \tilde{\psi}(z_{ij}(t); \beta, \gamma)$ denotes the gradient of $\tilde{\psi}$ with respect to β and γ . In a special case when ξ_i follows a multivariate-normal distribution with mean zero and covariance Σ and $\psi(x) = \exp(x)$, we can easily calculate

$$\tilde{\psi}(z_{ij}(t); \beta, \Sigma) = \exp \left\{ z_{ij}(t)^T \beta + z_{ij}(t)^T \Sigma z_{ij}(t) / 2 \right\}.$$

We notice that when $\psi(x) = x$, equation (5) reduces to (3). Thus, β is still estimable but not γ . The asymptotic properties of $\hat{\beta}$ which solves (5) can be derived using the same procedure as in the appendix.

APPENDIX

Asymptotic properties

We derive the asymptotic distribution for assumptions: $\hat{\beta}$. We need the following assumptions:

(A.1) There exists two positive-definite matrices Σ_1 and Σ_2 such that

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) E \left[\left[R_{ij}(t) \{ z_{ij}(t) \psi'(z_{ij}(t))^T \beta_0 \} - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t))^T \beta_0}{\sum_{k=1}^{n_i} R_{ik}(t)} \right]^{\otimes 2} \right] dt \rightarrow \sum_1, \\ & \frac{1}{n} \sum_{i=1}^n E \left[\left\{ \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) (z_{ij}(t) \psi'(z_{ij}(t))^T \beta_0) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t))^T \beta_0}{\sum_{k=1}^{n_i} R_{ik}(t)} \right. \right. \\ & \quad \left. \left. \times (dN_{ij}(t) - R_{ij}(t) \psi(z_{ij}(t))^T \beta_0) dt \right\}^{\otimes 2} \right] \rightarrow \sum_2. \end{aligned}$$

(A.2) There exists a $\alpha_1 > 0$ such that with probability one, $P(C_i \geq \tau | z_i, C_i) > \alpha_1$.

(A.3) The link function ψ is twice-continuously differentiable.

(A.4) There exists a constant M such that $\max_{i=1, \dots, n, j=1, \dots, n_i} \sup_{t \in [0, \tau]} |z_{ij}(t)| \leq M$ with probability one.

In the special case of $\psi(x) = x$, condition (A.1) can be reduced to the linear independence of $[1, z_{ij}(t)]$ for some i and j and $t \in [0, \tau]$. We denote the left-hand side in (3) as $l_n(\beta)$. Then it holds that

$$\begin{aligned} & \frac{1}{n} E[l_n(\beta_0)] \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) E \left[\left[R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t))^T \beta_0 \right\} - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t))^T \beta_0}{\sum_{k=1}^{n_i} R_{ik}(t)} \right] \right. \\ & \quad \left. \times E[dN_{ij}(t) - R_{ij}(t) \psi(z_{ij}(t))^T \beta_0] dt \middle| R_{ij}(t), z_{ij}(t), C_i \right] \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) E \left[\left[R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t))^T \beta_0 \right\} - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t))^T \beta_0}{\sum_{k=1}^{n_i} R_{ik}(t)} \right] d\Lambda_0(t; C_i) \right] \\ &= 0. \end{aligned}$$

The score function for $l_n(\beta)$ in (3), $\dot{l}_n(\beta)$ is

$$\begin{aligned}
& \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ \psi''(z_{ij}(t)^T \beta) z_{ij}(t) z_{ij}(t)^T - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \psi''(z_{ik}(t)^T \beta) z_{ik}(t) z_{ik}(t)^T}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} dN_{ij}(t) \\
& - \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left\{ [\psi''(z_{ij}(t)^T \beta) \psi(z_{ij}(t)^T \beta) + \{\psi'(z_{ij}(t)^T \beta)\}^2] z_{ij}(t) z_{ij}(t)^T dt \right. \\
& - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \psi''(z_{ik}(t)^T \beta) z_{ik}(t) z_{ik}(t)^T}{\sum_{k=1}^{n_i} R_{ik}(t)} \psi(z_{ij}(t)^T \beta) dt \\
& \left. - \frac{\sum_{k=1}^{n_i} R_{ik}(t) \psi'(z_{ik}(t)^T \beta) z_{ik}(t)}{\sum_{k=1}^{n_i} R_{ik}(t)} \psi'(z_{ij}(t)^T \beta) z_{ij}(t)^T dt \right\}.
\end{aligned}$$

Additionally, when n is large enough, for any γ ,

$$\begin{aligned}
& \frac{1}{n} E[\gamma^T l_n(\beta_0) \gamma] \\
& \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) E \left\| R_{ij}(t) \gamma^T \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta_0) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta_0)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \right\|^2 dt \\
& \rightarrow \gamma^T \sum_1 \gamma \geq \alpha_0 \|\gamma\|^2,
\end{aligned}$$

for some $\alpha_0 > 0$. Therefore, for any $\epsilon > 0$, as a map from $\{\beta: \|\beta - \beta_0\| < \epsilon\}$ to real space,

$E[l_n(\beta)]/n$ has a range which contains $\{x: \|x\| < \delta(\epsilon)\}$ for some positive constant $\delta(\epsilon)$ and for all large n .

By the Hoeffding's inequality and condition (A.4), we have that for any $\epsilon > 0$,

$$P\left(\left|\frac{1}{n} l_n(\beta) - \frac{1}{n} E[l_n(\beta)]\right| > \epsilon\right) \leq \exp\{-\alpha_1 n \epsilon^2\},$$

where α_1 is a constant depending on β . Therefore, the first Borel-Cantelli lemma gives

$$\frac{1}{n} l_n(\beta) - \frac{1}{n} E[l_n(\beta)] \rightarrow_{a.s.} 0.$$

Since according to conditions (A.3) and (A.4), $\dot{l}_n(\beta)/n$ is uniformly bounded when β belongs to any compact set $\mathcal{B}$, we conclude that

$$\sup_{\beta \in \mathcal{B}} \left| \frac{1}{n} l_n(\beta) - \frac{1}{n} E[l_n(\beta)] \right| \rightarrow_{a.s.} 0.$$

Hence, for any $s > 0$, as a map from $\{\beta: \|\beta - \beta_0\| < \epsilon\}$, the range of $I_n(\beta)/n$ should also contain the ball $\{x: \|x\| < \delta(\epsilon)/2\}$ when n is large. Thus, there exists an $\hat{\beta}_n$ such that $I_n(\hat{\beta}_n) = 0$ and $\|\hat{\beta}_n - \beta_0\| < \epsilon$. That is, $I_n(\beta) = 0$ has one and unique solution $\hat{\beta}_n$. The arbitrary choice of ϵ gives that $\hat{\beta}_n$ converges to β_0 almost surely. We have proved that there exists a sequence of estimators $\{\hat{\beta}_n\}$ such that $I_n(\hat{\beta}_n) = 0$ and $\hat{\beta}_n \rightarrow_{a.s.} \beta_0$. Particularly, we note that when $\psi(x) = x$, $\hat{\beta}_n$ is also the unique solution to $I_n(\beta) = 0$.

To derive the asymptotic distribution of $\hat{\beta}_n$, we perform the Taylor expansion of the left-hand side in (3) and obtain

$$\begin{aligned} & \sqrt{n} \left\{ \frac{1}{n} i_n(\beta_0) + o_p(1) \right\} (\hat{\beta}_n - \beta_0) \\ &= \frac{1}{\sqrt{n}} i_n(\beta_0) = \frac{1}{\sqrt{n}} l_n(\beta_0) - \frac{1}{n} E[l_n(\beta_0)] \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) [R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta_0) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta_0)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \\ & \quad \times \{ dN_{ij}(t) - R_{ij}(t) \psi(z_{ij}(t)^T \beta_0) dt \} \\ & \quad - E \left[R_{ij}(t) \left\{ z_{ij}(t) \psi'(z_{ij}(t)^T \beta_0) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \beta_0)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right\} \right. \\ & \quad \left. \times \{ dN_{ij}(t) - R_{ij}(t) \psi(z_{ij}(t)^T \beta_0) dt \} \right] \Bigg]. \end{aligned}$$

According to condition (A.1) and the fact that $i_n(\beta_0)/n - E[i_n(\beta_0)/n] \rightarrow_{a.s.} 0$, we have $i_n(\beta_0)/n \rightarrow \Sigma_1$. The Liapounov central limit theorem and condition (A.1) gives that the right-hand side converges in distribution to a normal distribution with mean zero and covariance Σ_2 . Therefore, we obtain

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \rightarrow_d N(0, \Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1}).$$

Clearly, the asymptotic covariance $\Sigma_1^{-1} \Sigma_2 \Sigma_1^{-1}$ can be consistently estimated using the following sandwiched estimator:

$$\begin{aligned}
& \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left[z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right]^{\otimes 2} dt \right]^{-1} \\
& \times \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) (z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)}) \right. \\
& \quad \left. \times (dN_{ij}(t) - R_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) dt) \right]^{\otimes 2} \\
& \times \left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^{n_i} \int \omega_i(t) R_{ij}(t) \left[z_{ij}(t) \psi'(z_{ij}(t)^T \hat{\beta}_n) - \frac{\sum_{k=1}^{n_i} R_{ik}(t) z_{ik}(t) \psi'(z_{ik}(t)^T \hat{\beta}_n)}{\sum_{k=1}^{n_i} R_{ik}(t)} \right]^{\otimes 2} dt \right]^{-1}.
\end{aligned}$$

REFERENCE

- Aalen O (1980). A model for nonparametric regression analysis of counting processes. In Lecture Notes in Statistics, 2, 1–25. New York: Springer.
- Aalen OO (1989). A linear regression model for the analysis of life times. *Statistics in Medicine*, 8, 907–925. [PubMed: 2678347]
- Andersen PK, Borgan Ø, Gill RD, and Keiding N (1993). *Statistical Models Based on Counting Processes* New York: Springer-Verlag.
- Breslow NE and Day NE (1980). *Statistical Models in Cancer Research, 1, The Design and Analysis of Case-Control Studies* Lyon: IARC.
- Breslow NE and Day NE (1987). *Statistical Models in Cancer Research, 2, The Design and Analysis of Cohort Studies* Lyon: IARC.
- Buckley JD (1984). Additive and multiplicative models for relative survival rates. *Biometrics*, 40, 51–62. [PubMed: 6733234]
- Cai J and Zeng D (2011). Additive mixed effect model for clustered failure time data. *Biometrics*, 67, 1340–1351. [PubMed: 21418052]
- Cai J, Zhou H, and Davis CE (1997). Estimating the mean hazard ratio parameters for clustered survival data with random clusters. *Statistics in Medicine*, 16, 2009–2020. [PubMed: 9304770]
- Cox DR and Oakes D (1984). *Analysis of Survival Data*, 21 CRC press.
- Ding J and Sun L (2015). Additive mixed effect model for recurrent gap time data. *Lifetime Data Analysis*, online, 1–31. [PubMed: 25097158]
- Glidden DV and Vittinghoff E (2004). Modelling clustered survival data from multicentre clinical trials. *Statistics in Medicine*, 23, 369–388. [PubMed: 14748034]
- Holt JD and Prentice RL (1974). Survival analysis in twin studies and matched pari experiments. *Biometrika*, 61, 17–30.
- Huffer FW and McKeague IW (1991). Weighted least squares estimation for Aalen’s additive risk model. *Journal of the American Statistical Association*, 86, 114–129.
- Huster WJ, Brookmeyer R, and Self SG (1989). Modeling paried survival data with covariates. *Biometrics*, 45, 145–156. [PubMed: 2655727]
- Jones B and Kenward MG (2015). *Design and analysis of cross-over trials*, 3rd ed. New York: Chapman & Hall/CRC.
- Lee EW, Wei LJ, Amato DA, and Leurgans S (1992). *Survival Analysis: State of the Art*, chapter: Cox-type regression analysis for large numbers of small groups of correlated failure time observations. *Nato Science*, 211, 237–247.

- Lee EW, Wei LJ, and Ying Z (1993). Linear regression analysis for highly stratified failure time data. *Journal of the American Statistical Association*, 88, 557–565.
- Lin DY, Oakes D, and Ying Z (1998). Additive hazards regression with current status data. *Biometrika*, 85, 289–298.
- Lin DY and Ying Z (1994). Semiparametric analysis of the additive risk model. *Biometrika*, 81, 61–71.
- Martinussen T and Scheike TH (2002). Efficient estimation in additive hazards regression with current status data. *Biometrika*, 89, 649–658.
- Pan D, He H, Song X, and Sun L (2015). Regression analysis of additive hazards model with latent variables. *Journal of the American Statistical Association*, 110, 1148–1159.
- Pocock SJ, Gore SM and Kerr GR (1982). Long term survival analysis: the curability of breast cancer. *Statistics in Medicine*, 1, 93–104. [PubMed: 7187092]
- Schaubel DE, Zeng D and Cai J (2006). A semiparametric additive rates model for recurrent event data. *Lifetime Data Analysis*, 12, 389–406. [PubMed: 17031499]
- SOLVD Investigators. (1990). Studies of left ventricular dysfunction (SOLVD)—Rationale, design and methods: Two trials that evaluate the effect of enalapril in patients with reduced ejection fraction. *The American Journal of Cardiology*, 66, 315–322. [PubMed: 2195865]

Figure 1:
Power curves when $\lambda(t; C_i) = \xi_i I(i < n/2) + \xi_i t I(i \geq n/2)$, $\xi_i \sim \Gamma(1, 1)$.

Table 1:

Simulation results with $\psi(x) = x$

		center size=2				center size=6				center size=20			
<i>n</i>	par.	Bias	SE	ESE	CP	Bias	SE	ESE	CP	Bias	SE	ESE	CP
$\lambda(t; C_i) = t\xi_i, \quad \xi_i \sim \Gamma(1, 1)$													
25	β_1	-.006	1.105	.957	.91	.021	.758	.737	.93	.001	.334	.310	.92
	β_2	.112	1.167	1.024	.91	.030	.740	.724	.93	-.001	.342	.314	.92
50	β_1	-.055	.955	.870	.92	-.009	.372	.367	.94	.005	.216	.210	.95
	β_2	.103	.936	.855	.93	.022	.390	.390	.94	.004	.221	.220	.95
100	β_1	.014	.596	.568	.93	.001	.309	.293	.94	-.014	.152	.148	.93
	β_2	.031	.639	.606	.94	.002	.319	.304	.93	.008	.161	.156	.94
200	β_1	-.003	.388	.386	.95	-.001	.195	.191	.95	-.006	.111	.107	.94
	β_2	.010	.399	.392	.95	.009	.203	.203	.95	.009	.117	.113	.94
$\lambda(t; C_i) = \xi_i I(i < n/2) + \xi_i t I(i \geq n/2), \quad \xi_i \sim \Gamma(1, 1)$													
25	β_1	-.020	1.559	1.378	.93	.022	.901	.872	.94	-.002	.402	.371	.92
	β_2	.139	1.471	1.324	.93	.035	.836	.814	.94	.002	.389	.353	.92
50	β_1	-.010	1.249	1.154	.94	-.007	.469	.462	.95	.005	.259	.248	.94
	β_2	.152	1.131	1.043	.94	.022	.456	.452	.95	.004	.250	.245	.94
100	β_1	.015	.729	.719	.95	-.004	.376	.356	.92	-.017	.178	.176	.94
	β_2	.032	.735	.709	.94	.006	.363	.346	.94	.011	.178	.175	.95
200	β_1	.006	.499	.507	.96	.001	.239	.233	.94	-.005	.129	.127	.95
	β_2	.005	.477	.473	.96	.007	.232	.230	.95	.010	.130	.126	.94
$\lambda(t; C_i) = \sqrt{i} \bar{Z}_i^T \beta_0 t/2, \quad \bar{Z}_i = n_i^{-1} \sum_{j=1}^{n_i} Z_{ij}$													
25	β_1	.094	1.046	.905	.92	.029	.803	.777	.94	.005	.332	.310	.93
	β_2	.003	1.214	1.072	.92	.020	.814	.788	.93	-.003	.365	.329	.92
50	β_1	.026	.982	.901	.94	.005	.373	.380	.95	.009	.235	.227	.94
	β_2	.022	1.030	.960	.94	.006	.431	.435	.96	-.000	.256	.249	.94
100	β_1	.018	.680	.670	.95	.007	.342	.334	.94	-.013	.168	.170	.96
	β_2	.026	.798	.774	.94	-.007	.378	.367	.94	.003	.188	.189	.96
200	β_1	.016	.471	.467	.94	.002	.244	.241	.95	-.001	.138	.135	.95
	β_2	-.002	.542	.527	.94	.004	.270	.272	.95	.006	.154	.149	.94

Table 2:Simulation results with $\psi(x) = \exp(x)$

<div><div>center size=2</div><div>center size=6</div><div>center size=20</div></div>													
<i>n</i>	par.	Bias	SE	ESE	CP	Bias	SE	ESE	CP	Bias	SE	ESE	CP
$\lambda(t; C_i) = t\xi_i, \quad \xi_i \sim \Gamma(1, 1)$													
25	β_1	.036	1.098	.907	.93	.020	.735	.610	.93	−.024	.283	.254	.89
	β_2	−.004	1.306	1.056	.94	.003	.583	.481	.91	.021	.242	.211	.90
50	β_1	−.034	.836	.805	.91	−.036	.406	.369	.95	−.008	.174	.164	.95
	β_2	.022	.673	.631	.89	.022	.332	.304	.95	.002	.149	.140	.94
100	β_1	−.018	.530	.502	.93	−.006	.257	.241	.93	−.014	.121	.121	.95
	β_2	.024	.463	.429	.93	.004	.221	.105	.93	.009	.105	.104	.95
200	β_1	−.010	.380	.352	.95	−.000	.162	.160	.95	−.004	.086	.085	.95
	β_2	.011	.319	.295	.95	.003	.138	.137	.95	.005	.076	.074	.95
$\lambda(t; C_i) = \xi_i I(i < n/2) + \xi_i t I(i \geq n/2), \quad \xi_i \sim \Gamma(1, 1)$													
25	β_1	.006	1.205	1.423	.93	.004	.774	.655	.92	−.019	.314	.281	.90
	β_2	−.004	.977	1.130	.93	.016	.603	.510	.91	.020	.260	.227	.90
50	β_1	−.033	.959	.918	.91	−.032	.469	.440	.95	−.004	.193	.180	.96
	β_2	.033	.743	.695	.89	.026	.367	.346	.95	.002	.160	.149	.94
100	β_1	−.025	.600	.568	.93	−.014	.288	.268	.93	−.016	.132	.133	.95
	β_2	.032	.504	.469	.92	.005	.240	.221	.93	.012	.111	.111	.95
200	β_1	−.005	.438	.406	.95	−.000	.181	.178	.95	−.004	.094	.094	.96
	β_2	.004	.353	.329	.95	.002	.148	.147	.95	.004	.081	.079	.95
$\lambda(t; C_i) = \sqrt{i} \bar{Z}_i^T \beta_0 t / 2, \quad \bar{Z}_i = n_i^{-1} \sum_{j=1}^{n_i} Z_{ij}$													
25	β_1	−.032	1.122	1.328	.91	.006	.695	.619	.92	−.016	.272	.253	.92
	β_2	−.019	.954	1.087	.93	−.002	.555	.493	.91	.014	.237	.212	.91
50	β_1	−.015	.867	.840	.90	−.033	.406	.403	.95	.001	.178	.168	.95
	β_2	.025	.691	.659	.90	.031	.333	.327	.94	.001	.153	.143	.94
100	β_1	−.034	.543	.537	.94	−.004	.269	.252	.94	−.015	.125	.126	.96
	β_2	.046	.473	.459	.94	.001	.230	.215	.94	.010	.109	.109	.96
200	β_1	−.005	.417	.381	.94	−.002	.172	.171	.95	−.002	.092	.091	.95
	β_2	.001	.349	.320	.95	.003	.145	.146	.96	.006	.080	.079	.95

Table 3:

Results of comparing our method vs Lin and Ying's method

		center size=2			center size=6			center size=20		
<i>n</i>	par.	Bias	SE	<i>R</i> _{MSE}	Bias	SE	<i>R</i> _{MSE}	Bias	SE	<i>R</i> _{MSE}
$\lambda(t; C_i) = t\xi_i, \quad \xi_i \sim \Gamma(1, 1)$										
100	β_1	.032	.513	1.347	.003	.284	1.185	−.013	.151	1.010
	β_2	.015	.652	1.340	.002	.297	1.152	.008	.159	1.022
200	β_1	.016	.329	1.387	.001	.182	1.153	−.005	.107	1.079
	β_2	−.002	.345	1.339	.006	.189	1.145	.010	.112	1.078
$\lambda(t; C_i) = \xi_i I(i < n/2) + \xi_i t I(i \geq n/2), \quad \xi_i \sim \Gamma(1, 1)$										
100	β_1	.319	.604	1.139	.329	.345	.622	.291	.193	.262
	β_2	−.286	.615	1.176	−.332	.337	.588	−.308	.189	.243
200	β_1	.361	.408	.839	.306	.221	.401	.301	.134	.155
	β_2	−.348	.393	.794	−.315	.213	.374	−.308	.132	.152
$\lambda(t; C_i) = \sqrt{i} \bar{Z}_i^T \beta_0 t / 2, \quad \bar{Z}_i = n_i^{-1} \sum_{j=1}^{n_i} Z_{ij}$										
100	β_1	−.861	.558	.439	−.777	.325	.165	−.732	−.163	.050
	β_2	.955	.616	.493	.774	.341	.200	.717	.171	.065
200	β_1	−1.178	.398	.143	−.916	.217	.067	−.894	.128	.023
	β_2	1.199	.411	.183	.945	.223	.077	.908	.135	.028

Note: *R*_{MSE} is the ratio of the mean square errors of our estimator vs Lin and Ying's estimator.

Table 4:

Results of analyzing SOLVD data

Covariates	$\psi(x) = x$			$\psi(x) = e^x$		
	Estimate	SE	P-value	Estimate	SE	P-value
Treatment vs placebo	-.0637	.0136	<0.001	-.0713	.0159	<0.001
Age	.0102	.0027	<0.001	.0116	.0028	<0.001
Male vs female	-.0182	.0153	0.234	-.0191	.0171	0.264
Ejection fraction	-.0041	.0007	<0.001	-.0046	.0009	<0.001