



NIH PUBLIC ACCESS

Author Manuscript

Lifetime Data Anal. Author manuscript; available in PMC 2010 November 18.

Published in final edited form as:

Lifetime Data Anal. 2010 October ; 16(4): 491–508. doi:10.1007/s10985-009-9140-6.

Joint covariate-adjusted score test statistics for recurrent events and a terminal event

Rui Song andDepartment of Statistics, Colorado State University, Fort Collins, CO 80526, USA
song@stat.colostate.edu**Jianwen Cai**

Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7420, USA cai@bios.unc.edu

Abstract

Recurrent events data are frequently encountered and could be stopped by a terminal event in clinical trials. It is of interest to assess the treatment efficacy simultaneously with respect to both the recurrent events and the terminal event in many applications. In this paper we propose joint covariate-adjusted score test statistics based on joint models of recurrent events and a terminal event. No assumptions on the functional form of the covariates are needed. Simulation results show that the proposed tests can improve the efficiency over tests based on covariate unadjusted model. The proposed tests are applied to the SOLVD data for illustration.

Keywords

Frailty; Proportional hazards; Proportional rates; Recurrent events data; Semiparametric model

1 Introduction

In many clinical trials, events may occur repeatedly for individual subjects and the occurrence of such recurrent events could be stopped by a terminal event. For instance, recurrences of hospitalizations could be terminated by the death of a patient. It is often of interest to assess the treatment efficacy simultaneously with respect to both the recurrent events and the terminal event in many applications.

As an example, let us consider the Studies of Left Ventricular Dysfunction (SOLVD), conducted between June 1986 and March 1989, to determine the effectiveness of Enalapril in reducing mortality and hospitalizations for heart failure in patients with chronic congestive heart failure and low ejection fractions (The SOLVD Investigators 1991). A total of 2569 patients were randomly assigned to placebo or Enalapril. The patients experienced between zero and twenty-seven hospitalizations during the course of the study. The death rate was about 39.7% for the placebo group and 35.2% for the SOLVD group. Since death could be related with the hospitalization process, these high death rates are not negligible.

Several methods have been proposed for the analysis of recurrent events data, for example, the work of Andersen and Gill (1982); Wei et al. (1989); Pepe and Cai (1993); Lawless and Nadeau (1995) and Lin et al. (2000). All these methods assume that the censoring mechanism is independent of the recurrent event processes. Some efforts have been made in recent years on modeling the recurrent events and the terminal event, such as Huang and Wang (2004) (HW hereinafter), Liu et al. (2004); Ye et al. (2007); Huang and Liu (2007); Rondeau (2007) and Liu and Huang (2008). These works account for the dependence between the recurrent events

and the terminal event by allowing a common frailty variable to have a multiplicative effect on the recurrent event intensity, rates/means function and on the hazard function respectively. To assess the treatment efficacy, simultaneous test statistics can be built on the score statistics of the joint models for both the recurrent events and the terminal event.

Jointly assessing the treatment efficacy on the recurrent events and death has been treated previously using log-rank type of statistics, see Cook and Lawless (1997) and Ghosh and Lin (2000). There are several differences between these two methods and the proposed method. Firstly, both methods of Cook and Lawless (1997) and Ghosh and Lin (2000) build on the assumption that the recurrent event processes are terminated by death in the sense that they are not defined after death. In this paper, we view the recurrent event process as latent, unobservable after death; that is, death is viewed as dependent censoring. This was also the approach taken by Ghosh and Lin (2003) and Huang and Wang (2004). Secondly, Ghosh and Lin (2000) proposed a one degree freedom test by taking a linear combination of the marginal log-rank statistics for death and for recurrent events, to perform simultaneous inference on both endpoints. When there is a treatment effect in one endpoint but not in the other, or the treatment effects for both endpoints are in opposite directions, such a combined statistic does not have satisfactory power. In addition, the clinical interpretation in these circumstances is not very clear and could be misleading sometimes. Moreover, both methods of Cook and Lawless (1997) and Ghosh and Lin (2000) build on the marginal treatment effects for the endpoints and does not consider the correlation between the recurrent event process and the terminal event, which however often exists in practice. With the intention to test the treatment effects on both endpoints while dealing with the dependence between the terminal event and the recurrent events and incorporating auxiliary covariates information, we consider joint modeling based covariates-adjusted score statistics in this paper.

It is well known that incorporating covariates in log-rank type of test can improve efficiency (Tsiatis et al. 1985; Slud 1991; Kong and Slud 1997; Li 2001). The rationale of this approach builds on the behavior of such tests under model misspecification (Struthers and Kalbfleisch 1986). That is, the unadjusted test is actually based on a different model other than the data generating process, if the process indeed depends on some covariates. The efficiency of such covariate-adjusted tests highly depends on the prespecified working model: these tests may not be as efficient as the covariate-unadjusted tests if the prespecified working model is wrong. The interpretation of the unadjusted test and that of the covariate-adjusted test is thus different: the covariate-adjusted test is trying to test a conditional treatment effect; while the unadjusted test is trying to test a marginal treatment effect.

Recent development of semiparametric theory (Bickel et al. 1998; Tsiatis 2006) provides a new framework of covariate adjustment in treatment comparison. In this framework, the model space can be enlarged by incorporating more covariates which are independent of the treatment assignment, but correlated with the outcome. Different from the covariate-adjusted approach mentioned above, the approach does not involve study of model misspecification. Since the auxiliary covariates can provide more information of the outcome, this method can provide more efficient inference compared with the covariate-unadjusted modeling while possessing the same model interpretation.

In this paper, we propose covariate-adjusted score statistics based on HW's model using this approach. The inference is focused on the comparison of the treatment effects on the intensity of the recurrent events process and on the cumulative hazards of the terminal event. We propose efficient influence functions in a certain class of functions and develop the inference procedure. Simulation results demonstrate that the proposed tests can significantly improve efficiency compared with the original HW's model. Compared with the score tests based on HW's model with the treatment indicator as the only regressor, our test can incorporate auxiliary covariates

information, hence are more efficient while keeping the same interpretation and similar computational advantage as HW's method. Moreover, compared with the traditional covariate adjustment methods, the proposed method provides a better solution when the functional form of the auxiliary covariate is unknown. We also give rigorous justification and asymptotics of the inference procedure, by applying the semiparametric theory to the multivariate failure time data. We demonstrate both empirically and theoretically the efficiency gains of the proposed tests over tests based on HW's model with the treatment indicator as the only regressor. We also empirically demonstrate the efficiency gains of the proposed method over HW's model with the adjustment covariates as the regressors.

The remainder of the paper is organized as follows. In Sect. 2, the data set-up and HW's joint model are introduced. The proposed efficient influence functions, the corresponding estimating equations and the inference procedure are presented in Sect. 3. Section 4 gives some simulation evidence and the proposed method is applied to the data from the SOLVD study in Sect. 5. The article concludes with some discussion in Sect. 6.

2 The data and model assumptions

Assume that there are n independent subjects randomly assigned to two treatments with known probabilities ρ and $1 - \rho$, and accordingly define $X_i = 0$ or 1 , for subject $i = 1, \dots, n$. Let $T_i \equiv (T_{i1}, T_{i2}, \dots, T_{iM_i})$, where $T_{i1} < T_{i2} < \dots < T_{iM_i}$ are the ordered event times of interest which constitute the recurrent event process, and M_i denotes the number of recurrent events for subject i . Let D_i be the terminal event time and C_i be the censoring time. The research interest is focused on T_i and D_i in time interval $[0, \tau]$, where τ is the time for the end of the study and both event processes potentially could be observed beyond τ .

Define $Y_i \equiv \min(C_i, D_i, \tau)$, the time when the observation of the recurrent event process is terminated, and $\Delta_i \equiv I(D_i \leq Y_i)$, the failure indicator function for the terminal event. The observed data for subject i , $O_i \equiv (T_i, Y_i, \Delta_i, X_i)$ are independent replicates of $O \equiv (T, Y, \Delta, X)$. We also utilize when convenient the following counting process notation:

$$N_{ik}^{R0}(t) \equiv I\{T_{ik} \leq t\}, N_i^{R0}(t) \equiv \sum_{k=1}^{M_i} N_{ik}^{R0}(t), N_i^R(t) \equiv N_i^{R0}(t \wedge Y_i), N_i^{D0}(t) \equiv I\{D_i \leq t\} \text{ and } N_i^D(t) \equiv N_i^{D0}(t \wedge Y_i).$$

To account for correlations among T_i , D_i and C_i , a nonnegative latent variable γ is used without specifying its distribution. This latent variable γ satisfies $E(\gamma) = E(\gamma|X) = 1$ to assure the model identifiability. Conditioning on the covariate X_i and the latent variable γ_i , which is an i.i.d. realization of γ , $\{T_i, D_i, C_i\}$ are assumed to be mutually independent.

Let \mathcal{F}_t^R be the σ -field generated by the recurrent event process $\{N^{R0}(s) : 0 \leq s \leq t\}$, and $\lambda(t)$ be the intensity function of $N^{R0}(t)$ associated with \mathcal{F}_t^R , that is, $E\{dN^{R0}(t) | \mathcal{F}_{t-}^R\} = \lambda(t) dt$, where $dN^{R0}(t)$ is the increment $N^{R0}\{(t + dt)-\} - N^{R0}(t-)$ of N^{R0} over the small interval $[t, t + dt)$. It is assumed that $\lambda(t) = \gamma \lambda_0(t) \exp(X\alpha_0)$, $t \in [0, \tau]$, where the baseline intensity function $\lambda_0(t)$ is a continuous function, and the log intensity ratio α_0 denotes the treatment effect for the recurrent events. Similarly, let \mathcal{F}_t^D be the σ -field generated by the death process $\{N^{D0}(s) : 0 \leq s \leq t\}$, and $h(t)$ be the hazard function of $N^{D0}(t)$ associated with \mathcal{F}_t^D , that is, $E\{dN^{D0}(t) | \mathcal{F}_{t-}^D\} = h(t) dt$, where $dN^{D0}(t)$ is the increment $N^{D0}\{(t + dt)-\} - N^{D0}(t-)$ of N^{D0} over the small interval $[t, t + dt)$. It is assumed that $h(t) = \gamma h_0(t) \exp(X\beta_0)$, $t \in [0, \tau]$, where $h_0(t)$ is a continuous baseline hazard function and the log hazard ratio β_0 denotes the treatment effect for the terminal event.

Let $\Lambda_0(\cdot)$ and $H_0(\cdot)$ be the cumulative baseline intensity function and cumulative baseline hazard function respectively, the likelihood function based on $\{\gamma_i\}$ and the observed data $\{O_i\}$ takes the form $L_n(\alpha, \beta, \lambda_0(\cdot), h_0(\cdot)) \equiv$

$$\prod_{i=1}^n \left[\prod_{k=1}^{M_i} \left\{ \gamma_i e^{X_i \alpha} \lambda_0(T_{ik}) \right\} e^{-\gamma_i \Lambda_0(Y_i) e^{X_i \alpha}} \left\{ \gamma_i e^{X_i \beta} h_0(Y_i) \right\}^{\Delta_i} e^{-\gamma_i H_0(Y_i) e^{X_i \beta}} \right]. \tag{1}$$

Since $E(N^R(Y_i)|X_i, Y_i, \gamma_i) = E(M_i|X_i, Y_i, \gamma_i) = \gamma_i \exp(X_i \alpha_0) \Lambda_0(Y_i)$, it follows that

$E(M_i \Lambda_0^{-1}(Y_i) | X_i, Y_i) = \exp(X_i \alpha_0)$, which suggests the following estimating equation for α_0 :

$$n^{-1} \sum_{i=1}^n X_i \left\{ M_i \Lambda_0(Y_i)^{-1} - \exp(X_i \alpha) \right\} = 0. \tag{2}$$

An estimating equation for β_0 can be constructed from the usual partial score function of β for the terminal event:

$$U(\beta) \equiv \sum_{i=1}^n \Delta_i \left\{ X_i - \frac{\sum_{j=1}^n X_j \gamma_j \exp(X_j \beta) I(Y_j \geq Y_i)}{\sum_{j=1}^n \gamma_j \exp(X_j \beta) I(Y_j \geq Y_i)} \right\} = 0. \tag{3}$$

We note that, however, Eqs. 2 and 3 are not directly applicable, since $\{\gamma_j\}$ are not observed and $\Lambda_0(\cdot)$ is unknown. Huang and Wang (2004) proposed plug-in estimator $\widehat{\Lambda}_0(\cdot)$ and $\widehat{\gamma}_j$ which are easy to compute as follows.

To find an unbiased estimator of $\Lambda(\cdot)$, we note that the likelihood in Eq. 1 can be decomposed as $L_n =$

$$\left\{ \prod_{i=1}^n \prod_{k=1}^{M_i} \frac{\lambda_0(T_{ik})}{\Lambda_0(Y_i)} \right\} \times \prod_{i=1}^n \left[\prod_{k=1}^{M_i} \left\{ \gamma_i e^{X_i \alpha} \Lambda_0(Y_i) \right\} e^{-\gamma_i \Lambda_0(Y_i) e^{X_i \alpha}} \left\{ \gamma_i e^{X_i \beta} h_0(Y_i) \right\}^{\Delta_i} e^{-\gamma_i H_0(Y_i) e^{X_i \beta}} \right] \equiv L_{1n}(\lambda_0(\cdot)) \times L_{2n}(\alpha, \beta, \lambda_0(\cdot), h_0(\cdot)).$$

The first part L_{1n} arises from the conditional distribution of the event times, given the observed number of events M_i , and it does not require information on X_i and the unobserved γ_i . Define a probability density function $f(t) = \lambda_0(t)/\Lambda_0(\tau)$ on $t \in (0, \tau]$ and the corresponding cumulative probability function $F(t)$. Conditional on $(\gamma_i, Y_i, X_i, M_i)$, the observed recurrent event times are the order statistics of a set of independent and identically distributed random variables with likelihood L_{1n} . $F(t)$ can be estimated by a nonparametric maximum likelihood estimator of L_{1n} , with:

$$\widehat{F}(t) = \prod_{s_{(l)} > t} \left(1 - \frac{d_{(l)}}{N_{(l)}} \right), \tag{4}$$

where $\{s_{(l)}\}$ are the ordered and distinct values of the event times $\{T_{ik}\}$, $d_{(l)}$ is the number of events occurring at $s_{(l)}$, and $N_{(l)}$ is the total number of events with event time and censoring time satisfying $T_{ik} \leq s_{(l)} \leq Y_i$. The class of estimating function for α can be upgraded as

$$n^{-1} \sum_{i=1}^n \bar{X}_i \left\{ M_i \widehat{F}(Y_i)^{-1} - \exp \left(\bar{X}_i \bar{\alpha} \right) \right\} = 0,$$

where $\bar{X}_i = (1, X_i)$ and $\bar{\alpha} = (\log(\Lambda_0(\tau)), \alpha)$. The estimator is denoted as $\widehat{\alpha}_{hw}$.

To resolve the problem of unknown γ in the estimation of β , again note that $E(M_i | X_i, Y_i, \gamma_i) = \gamma_i \exp(X_i \alpha_0) \Lambda_0(Y_i)$. The subject specific frailty γ_i can thus be estimated as:

$$\widehat{\gamma}_i = \frac{M_i}{\widehat{F}(Y_i) \exp \left\{ \bar{X}_i \widehat{\alpha}_{hw} \right\}}$$

Plugging $\widehat{\gamma}_i$ into the score function (3), another estimating equation for β becomes:

$$\sum_{i=1}^n \Delta_i \left\{ X_i - \frac{\sum_j X_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)}{\sum_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)} \right\} = 0,$$

with the estimator denoted as $\widehat{\beta}_{hw}$. Define functions $G(t) \equiv E(\gamma I(Y_1 \geq t))$, $R(t) \equiv G(t) \Lambda_0(t)$, $Q(t) \equiv \int_0^t G(u) d\Lambda_0(u)$, and

$$b_i(t) \equiv \sum_{k=1}^{M_i} \left\{ \int_t^{\tau} \frac{I(T_{ik} \leq u \leq Y_i) dQ(u)}{R^2(u)} - \frac{I(t < T_{ik} \leq \tau)}{R(T_{ik})} \right\}.$$

It was shown in HW that under certain regulatory conditions,

$\sqrt{n}(\widehat{\alpha}_{hw} - \alpha_0) = n^{-1/2} \sum_{i=1}^n f_i(\alpha_0) + o_p(1)$, with the influence function $f_i(\alpha)$ being the second

entry of the vector function $E \left[-\partial e_1 / \partial \bar{\alpha} \right]^{-1} e_i$, where

$e_i = -E \left\{ F(Y)^{-1} \bar{X}^T M b_i(Y) \right\} + \bar{X}_i \left\{ M_i F^{-1}(Y_i) - \exp \left(\bar{X}_i \bar{\alpha}_0 \right) \right\}$. The asymptotic variance of $\widehat{\alpha}_{hw}$ is thus $E \left[f_1^2 \right]$.

To study the asymptotic properties of $\widehat{\beta}_{hw}$, further define

$$\begin{aligned} \psi_i(\beta) \equiv & \int_0^\tau \left[X_i - \frac{E \{ X \gamma e^{X\beta} I(Y \geq t) \}}{E \{ \gamma e^{X\beta} I(Y \geq t) \}} \right] d(\Delta_i I(Y_i \leq t) - E(\Delta I(Y \leq t))) \\ & + \int_0^\tau \left[\frac{\psi_{3i}(t; \beta) E \{ X \gamma e^{X\beta} I(Y \geq t) \}}{E \{ \gamma e^{X\beta} I(Y \geq t) \}^2} - \frac{\psi_{4i}(t; \beta)}{E \{ \gamma e^{X\beta} I(Y \geq t) \}} \right] d(E(\Delta I(Y \leq t))), \end{aligned}$$

where

$$\psi_{3i}(t; \beta) \equiv E \left\{ M \Lambda_0^{-1}(y) e^{X(b-\alpha)} I\{Y \geq t\} (X f_i(\alpha) + b_i(Y)) \right\} + M_i \Lambda_0^{-1}(Y_i) e^{X_i(\beta-\alpha)} I(Y_i \geq t) - E \left\{ \gamma e^{X\beta} I(Y \geq t) \right\}$$

, and

$$\psi_{4i}(t; \beta) \equiv E \left\{ MX \Lambda_0^{-1}(Y) e^{X(\beta-\alpha)} I(Y \geq t) (X f_i(\alpha) + b_i(Y)) \right\} + M_i X_i \Lambda_0^{-1}(Y_i) e^{X_i(\beta-\alpha)} I(Y_i \geq t) - E \left\{ X \gamma e^{X\beta} I(Y \geq t) \right\}$$

. Through some linearization arguments, it can be shown that under certain regularity

$$\text{conditions, } \sqrt{n}(\widehat{\beta}_{hw} - \beta_0) = n^{-1/2} \sum_{i=1}^n g_i(\beta_0) + o_p(1), \text{ where the influence function } g_i(\beta) = E$$

$[\partial U(\beta)/\partial \beta]^{-1} \psi_i(\beta)$. The asymptotic variance of $\widehat{\beta}_{hw}$ is thus $E[g_1^2]$. Since it is not easy to directly implement the plug-in estimator of f_i and g_i , bootstrap variance estimators of α and β is used in Huang and Wang (2004).

Compared with other joint models of the recurrent and the terminal event processes (Liu et al. 2004; Ye et al. 2007), HW's model relaxes the condition of independent non-death censoring and is computationally simpler at the expense of efficiency. To be specific, both Liu et al. (2004) and Ye et al. (2007) involve procedures solving for the high dimensional parameters: $\Lambda_0(\cdot)$ and $H_0(\cdot)$, while HW method uses a plug-in estimator of the shape function of $\Lambda_0(\cdot)$ and only needs to solve two estimating equations for α and β . As a consequence, Liu et al. (2004) and Ye et al. (2007) maximize the full likelihood, while HW method maximizes a conditional likelihood, hence may not be as efficient as the former two methods, which was empirically demonstrated by Ye et al. (2007). On the other hand, the frailty distribution in HW method is left unspecified, whereas Liu et al. (2004) and Ye et al. (2007) both have the gamma distribution assumption. As HW method allows greater generality by leaving the frailty distribution function unspecified, however, the plug-in estimation of γ for each individual is rather loose and leads to larger variability for the estimation of β . In the next section, with intention to improve efficiency while keeping the generality and computational advantage of HW method, we propose score test statistics through incorporating the information of auxiliary covariates.

3 A general class of influence functions

In this section, we provide a class of influence functions based on HW's model, and an auxiliary p -dimensional covariate Z . By "auxiliary covariate", we refer to variables other than the treatment indicator, that are correlated with the outcome processes. For ease of illustration, we set $p = 1$. We assume that Z is independent of X , and is correlated with both recurrent events and the terminal event processes. Conditional on (X, Z, γ) , $(N^R(\cdot), D, C)$ are mutually independent. As a referee point out, if censoring C affects D or $N^R(\cdot)$ in ways other than γ , then the proposed method can not relax the independent non-death censoring assumption. Other model assumptions in Sect. 2 also apply here.

We first introduce some relevant results in semiparametric efficiency framework for ease of readers. Assume the observed data $\{X_i\}$, $i = 1, \dots, n$ are i.i.d. random variables and the distribution function of X has the form $p_X(x; \theta)$ with respect to some dominating measure, indexed by an unknown parameter $\theta \equiv (\beta, \Lambda)$. The parameter of interest is β and Λ is a nuisance parameter. If an estimator $\widehat{\beta}_n$ asymptotically normally converges to the true value θ_0 with root-n rate, that is, $\sqrt{n}(\widehat{\beta}_n - \beta_0) = n^{-1/2} \sum_{i=1}^n \phi(X_i; \theta_0) + o_p(1)$, the mean zero function $\phi(X_i; \theta)$ is called the influence function and can be used for the inference of $\widehat{\beta}_n$. The influence functions associated with parameter β belong to $L_2^0(P)$, a Hilbert space \mathcal{S} of all functions with zero mean and finite variance. The efficient influence function $\phi^*(X; \theta)$ is defined as the influence function with the smallest variance, i.e., $\text{var}(\phi^*(X; \theta)) = \min_{\phi \in \mathcal{S}} \text{var}(\phi(X; \theta))$. The efficient influence function is associated with the efficient estimator which is the estimator with the smallest asymptotic variance achieving the information bound. The efficient influence function is orthogonal to the nuisance tangent space associated with Λ .

To find more efficient estimators of the parameter of interest, α and β , we utilize the augmented model space with auxiliary covariates Z and take two steps. In step 1, we provide a class of influence functions for α and β in this augmented model space by first characterizing the tangent space of the whole parameter, \mathcal{H} , and its orthogonal complement \mathcal{H}^\perp . Let $\mathcal{T}_1 \equiv \{l_1(Z) : E\{l_1(Z)\} = 0\}$, $\mathcal{T}_2 \equiv \{l_2(Z, X) : E\{l_2(Z, X) | Z\} = 0\}$ and $\mathcal{T}_3 \equiv \{l_3(Z, X, O) : E\{l_3(Z, X, O) | Z, X\} = 0\}$. Since X is independent of Z , the conditional distribution of X given Z is identical to the marginal distribution of X , which is binary with given mean ρ . Therefore \mathcal{T}_2 can be completely specified, for it does not involve any unknown parameters. As a consequence, a straightforward application of Theorem 4.5 of Tsiatis (2006), which is re-stated in the appendix, yields that \mathcal{H} is a direct sum \mathcal{T}_1 and \mathcal{T}_3 . Its orthocomplement, \mathcal{H}^\perp , is \mathcal{T}_2 .

Since the linear space \mathcal{T}_2 can be equivalently expressed as the space $\{h(Z, X) - E\{h(Z, X) | Z\} : h(Z, X) - E\{h(Z, X) | Z\}\}$, and X is a binary indicator function, the double index function $h(Z, X)$ can be simplified into sum of two functions indexed only by Z : $h(Z; X = 1)1(X = 1) + h(Z; X = 0)1(X = 0)$. Let $h_1(Z) \equiv h(Z; X = 1)$, and $h_2(Z) \equiv h(Z; X = 0)$, we can have the expression that $h(Z; X) = Xh_1(Z) + (1-X)h_2(Z)$. It follows that any function $l_2(Z, X)$ in \mathcal{T}_2 can be expressed as $h(Z, X) - E\{h(Z, X) | Z\} = Xh_1(Z) + (1-X)h_2(Z) - \{E(X|Z)h_1(Z) + E((1-X)|Z)h_2(Z)\} = (X - \rho)h_1(Z) - (X - \rho)h_2(Z) = (X - \rho)h_3(Z)$. We thus show that any function in \mathcal{H}^\perp takes the form $(X - \rho)h_1(Z)$, where $h_1(Z)$ is an arbitrary function of Z .

As stated in Sect. 2, f_1 and g_1 are influence functions for α and β . By Theorem 4.3 of Tsiatis (2006) in the appendix, conditional on γ , the space of all influence functions of α and β take the form

$$\{f_1 + (X - \rho)h_1(Z)\} \quad \text{and} \quad \{g_1 + (X - \rho)h_2(Z)\}, \tag{5}$$

where $h_1(\cdot)$ and $h_2(\cdot)$ are arbitrary functions of Z .

In step 2, we further pick out the efficient influence functions among the class of functions (5). Following the semiparametric efficiency theory (Bickel et al. 1998), the efficient influence function is the projection of an arbitrary influence function on the tangent space of \mathcal{H} . We first compute $\text{proj}\{f_1 | \mathcal{H}^\perp\}$ and $\text{proj}\{g_1 | \mathcal{H}^\perp\}$, the projection of f_1 and g_1 onto the space \mathcal{H}^\perp . The efficient influence function for α and β are $f_1 - \text{proj}\{f_1 | \mathcal{H}^\perp\}$ and $g_1 - \text{proj}\{g_1 | \mathcal{H}^\perp\}$, respectively.

To find $\text{proj}\{f_1 | \mathcal{H}^\perp\}$, we need to find $h_1^*(Z)$ such that

$$E\left[\{f_1 - (X - \rho)h_1^*(Z)\}(X - \rho)h(Z) | \mathcal{H}^\perp\right] = 0, \text{ for all } h(Z), \text{ i. e., we require that}$$

$$E\left[\{f_1 - (X - \rho)h_1^*(Z)\}(X - \rho) | Z\right] = 0. \text{ Since } Z \text{ and } X \text{ are independent, simple algebra yields that}$$

$$h_1^*(Z) = \frac{1}{\rho(1 - \rho)} E\{(X - \rho)f_1(O; \alpha) | Z\}.$$

The efficient influence functions of α in the class of (5) is $f^* \equiv f_1 - (X - \rho)h_1^*(Z)$. Similarly we can show that the efficient influence functions of β in the class of (5) is

$$g^* \equiv g_1 - (X - \rho)h_2^*(Z), \text{ where}$$

$$h_2^*(Z) = \frac{1}{\rho(1-\rho)} E \{ (X - \rho) g_1(O;\beta) | Z \}.$$

When there are no further assumptions about the conditional distribution of D and $N(\cdot)$ given Z , it is generally difficult to compute f^* and g^* directly. Here we suggest to use a ‘‘sieved’’ approach, which restricts the search for the estimates of a and β to those with influence functions of the form $\{f_1 + (X - \rho)q^T(Z)a\}$, and $\{g_1 + (X - \rho)q^T(Z)b\}$, $a, b \in \mathbb{R}^k$, where k -dimensional vector $q(Z)$ consists of basis functions. For example, if we use the $(k - 1)$ -order polynomial basis, $q(Z) = \{1, Z, Z^2, \dots, Z^{k-1}\}^T$; one may also choose a spline basis of discretization basis with $q(Z) = \{I(Z \leq t_1), I(t_1 \leq Z \leq t_2), \dots, I(t_{k-2} \leq Z < t_{k-1}), I(Z \geq t_{k-1})\}$ for $t_1 < t_2 < \dots < t_{k-1}$. If the basis is sufficiently rich so that the linear space spanned by $q(Z)$ approximates the space of all possible h well, the resulting estimators should be close to ‘‘optimal’’ if they are ‘‘optimal’’ within the restricted class. A similar approach is also used by Leon et al. (2003).

Since the ‘‘meat’’ part of the variance of $\widehat{\alpha}_{hw}$ and $\widehat{\beta}_{hw}$ consists pieces that are not easy to implement, we restrict our search to minimize the variance components that are easy to compute. Let μ_n denote the sample mean of X_i , $i = 1, \dots, n$, our estimation procedure can be summarized as follows:

1. solve the least square problem

$$\min_a \frac{1}{n} \sum_i \left[X_i \left\{ M_i - \widehat{F}(Y_i) e^{X_i \bar{\alpha}} \right\} - (X_i - \mu_n) q^T(Z_i) a \right]^2, \tag{6}$$

for fixed $\bar{\alpha}$, where $\widehat{F}(Y_i)$ can be obtained as in (4). The minimizer is denoted as \hat{a} .

2. solve $\bar{\alpha}$ in the estimating equation

$$\sum_{i=1}^n \left[X_i \left\{ M_i \widehat{F}(Y_i)^{-1} - \exp \left(X_i \bar{\alpha} \right) \right\} - (X_i - \mu_n) q^T(Z_i) \bar{a} \right] = 0.$$

The estimator is denoted as $\widehat{\alpha}_p$.

3. solve the least square problem

$$\min_b \frac{1}{n} \sum_{i=1}^n \left[\Delta_i \left\{ X_i - \frac{\sum_j X_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)}{\sum_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)} \right\} - (X_i - \mu_n) q^T(Z_i) b \right]^2,$$

where $\widehat{\gamma}_i = M_i / \left[\widehat{F}(Y_i) \exp \left\{ X_i \widehat{\alpha}_p \right\} \right]$, $i = 1, \dots, n$. The minimizer is denoted as \hat{b} .

4. solve β in the estimating equation

$$\sum_{i=1}^n \left[\Delta_i \left\{ X_i - \frac{\sum_j X_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)}{\sum_j \widehat{\gamma}_j \exp(X_j \beta) I(Y_j \geq Y_i)} \right\} - (X_i - \mu_n) q^T(Z_i) \widehat{b} \right] = 0.$$

The estimator is denoted as $\widehat{\beta}_p$.

Both HW's estimating equations and the proposed estimating equations are in the general influence function classes. Since the general influence function classes are mean zero at the true value, both methods are unbiased.

The following arguments reveal that the proposed estimator is at least as efficient as HW's

estimator. Since $\partial e_1(\alpha) / \partial \alpha = X_1^2 \exp\left(\frac{-T}{X} \bar{\alpha}\right)$, the asymptotic variance of $\widehat{\alpha}_{hw}$ is

$$\text{Var}(\widehat{\alpha}_{hw}) = \left\{ EX^2 \exp\left(\frac{-T}{X} \bar{\alpha}_0\right) \right\}^{-2} E e_1(\alpha_0)^2 = \left\{ EX^2 \exp\left(\frac{-T}{X} \bar{\alpha}_0\right) \right\}^{-2} E(A+B(\alpha_0))^2, \text{ where } A \equiv -E\{X M b_1(Y) F(Y)^{-1}\}$$

accounts for the variance due to the estimation of $F(t)$, and it does not involve

$$a. \text{ The second part } B(\alpha) \equiv X \left[M F^{-1}(Y) - \exp\left(\frac{-T}{X} \bar{\alpha}\right) \right].$$

Let $e_1^* \equiv e_1 + (X_1 - \rho) q^T(Z_1) a^0$, where a^0 is the limite of the minimizer \hat{a} in (6). The asymptotic variance of $\widehat{\alpha}_p$ is

$$\text{Var}(\widehat{\alpha}_p) = \left\{ EX^2 \exp\left(\frac{-T}{X} \bar{\alpha}_0\right) \right\}^{-2} E e_1^{*2} = \left\{ EX^2 \exp\left(\frac{-T}{X} \bar{\alpha}_0\right) \right\}^{-2} E(A+B(\alpha_0) + (X_1 - \rho) q^T(Z_1) a^0)^2,$$

since $\partial e_1^*(\alpha) / \partial \alpha = X_1^2 \exp\left(\frac{-T}{X} \bar{\alpha}\right)$. It follows that $\text{Var}(\widehat{\alpha}_p) \leq \text{Var}(\widehat{\alpha}_{hw})$ since $E(B(\alpha_0) + (X_1 - \rho) q^T(Z_1) a^0)^2 \leq E(B(\alpha_0))^2$ and $E\{(X_1 - \rho) q^T(Z_1)\} = 0$. The proof for $\text{Var}(\widehat{\beta}_p) \leq \text{Var}(\widehat{\beta}_{hw})$ shares a similar argument and we omit the details here. Similar as in Huang and Wang (2004), due to the complexity of evaluating the variance directly, we propose using bootstrap method instead. We tried several bootstrap sizes as 250, 200, 100 and 50. Empirical results are satisfactory with bootstrap size as small as 50. In our experience, the computation time of implementing the proposed method is almost the same as that of HW's method with both covariate adjusted and unadjusted, since our procedure only has two steps more than HW's method: step 1 and step 3. These two steps usually take about 10 iterations to converge in one second in R.

Since the computation expense is rather small compared with HW's method while efficiency improvement could be substantial and the model interpretation is unchanged, we recommend to use the proposed method when there are some auxiliary covariates, which are independent of the treatment indicator but are related to the outcome processes.

4 Simulation studies

In this section we carry out simulations to illustrate the proposed method. For each setting, we generated 1000 datasets. The sample size n was set as 200. The study was restricted within time interval $[0, 10]$. The treatment indicator X was generated taking values 1 or 0 with probability 0.5. Given the treatment X and the frailty γ , a subject's underlying recurrent event process $\{N(t), t \in [0, 10]\}$ is a nonstationary Poisson process with the corresponding intensity function $\gamma \lambda_0(t) \exp(x\alpha)$, where the baseline intensity function is $\lambda_0(t) = 1$. The subject's failure time D has a hazard function $\gamma h_0(t) \exp(x\beta)$, where the baseline hazard function $h_0(t) = t/400$.

Given the frailty γ , the censoring time C was taken to follow an exponential distribution with mean 1 in the treatment group or with mean $3/\gamma^2$ in the control group. This censoring mechanism can be interpreted as follows when considering the frailty as an unobserved health indicator. Sick patients with a high occurrence rate in the control group will drop out early due to large values of frailty. On the other hand, the dropout of the patients in the treatment group is noninformative for both the recurrent event process and failure times, since the treatment has effectively reduced the event occurrence rate. The frailty γ follows a gamma distribution with unit mean and variance 0.5. Note that we do not restrict the cumulative intensity function to be one at the end of the study. It is standardized to be mean 1 instead.

To generate the nonhomogeneous Poisson process, we use the thinning method proposed in Lewis and Shedler (1979) and Kuhl and Bhaïrgond (2000), which can be summarized as follows:

- Step 1. Set $t = t_{i-1}$, the previous event time, where $t_0 = 0$.
- Step 2. Generate random variables U_1 and U_2 uniformly distributed on $[0,1]$ independently.
- Step 3. Replace t by $t - (1/\lambda^*) \log U_1$, where $\lambda^* \equiv \max_t \lambda(t)$.
- Step 4. If $U_2 \leq \lambda(t)/\lambda^*$, set $t_i = t$ and stop; else go back to step 2 and go on.

We consider one dimensional auxiliary covariate Z , which is independent of the treatment assignment X , but correlated with terminal event D and recurrent event process $N(t)$. We generate Z from standard normal distribution. The following three models are considered to demonstrate our proposed approach.

In the first model, the terminal event time was generated as $D = \sqrt{-800 \exp(-X\beta) \log\{\phi(Z)\} / \gamma}$, where $\phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable. In the following we explain why D is indeed generated from a proportional hazards model with hazard function $H(t|X, \gamma) = H_0(t) \exp\{X\beta\}\gamma$ and the baseline hazard $H_0(t) = t^2/800$. Conditional on covariates X and frailty γ , the survival function of a proportional hazards model satisfies $S(t|X, \gamma) = \exp\{-H(t|X, \gamma)\} = \exp\{-H_0(t) \exp\{X\beta\}\gamma\}$. Assume $D_0 \sim S(t|X, \gamma)$, utilizing the fact that $S(D_0|X, \gamma) \sim U[0, 1]$ and $\phi(Z) \sim U[0, 1]$, we generate the death time D_0 through $\phi(Z) = S(D_0|X, \gamma)$, that is, $\phi(Z) = \exp\{-H(t|D_0, \gamma)\} = \exp\{-H_0(D_0) \exp\{X\beta\}\gamma\}$. Since Z is standard normal and $\phi(\cdot)$ is the CDF of standard normal, $\phi(Z)$ follows a uniform $(0,1)$ distribution. Since $S(D_0|X, \gamma)$ also follows a uniform $(0, 1)$ distribution, we set $\phi(Z) = S(D_0|X, \gamma)$. Through this manner, we generate the survival time while incorporating the effect of Z at the same time. Consequently $D_0 = \sqrt{-800 \exp(-X\beta) \log\{\phi(Z)\} / \gamma}$. A sample correlation between the time to survival and Z was approximately -0.3 . The effect of Z on time to the first event (provided it happens) was incorporated in a similar manner. This leads to a sample correlation of approximately -0.5 between the time to the first recurrent event and baseline covariate Z . For this model, we consider three sub scenarios with $(\alpha, \beta) = (0, 0), (-0.4, -0.8)$ and $(-1, -1.5)$ to represent the null hypothesis, local alternative and large alternative respectively.

In the second model, we put Z as a linear regressor in the log intensity ratio and the hazard ratio:

$$\begin{aligned} \Lambda(t|X, Z, \gamma) &= \Lambda_0(t) \exp\{\alpha X + \alpha_1 Z\}, \quad \text{and} \\ H(t|X, Z, \gamma) &= H_0(t) \exp\{\beta X + \beta_1 Z\}. \end{aligned} \tag{7}$$

In the third model, the quadratic form of Z is put as a regressor to represent a U-shaped effect of the adjusted covariate:

$$\begin{aligned}\Lambda(t|X, Z, \gamma) &= \Lambda_0(t) \exp\{\alpha X + \alpha_1 Z^2\}, \quad \text{and} \\ H(t|X, Z, \gamma) &= H_0(t) \exp\{\beta X + \beta_1 Z^2\}.\end{aligned}\tag{8}$$

The values α_1 and β_1 are set to 0.5. In contrast with that of the first model, the covariate Z takes explicit forms in the second and the third model. For these two models, we consider two sub scenarios with $(\alpha, \beta) = (0, 0)$ and $(-0.4, -0.8)$ to represent the null and alternative hypothesis respectively.

In addition to the proposed method, two other methods are performed for all three models to gauge the difficulties of the problem: HW's method without covariate adjustment and the adjusted covariate as a linear regressor in the model. The standard error is obtained from the 50 times nonparametric bootstrap with replacement. The results for the three models are recorded in Tables 1, 2 and 3 respectively. The relative efficiencies for the proposed method over HW method in all three models are presented in Table 4.

For all of the cases considered here, all estimators are unbiased, the means of the standard error estimates agree well with the sample standard errors and the empirical type I errors are close to the nominal level 0.05. In the first model, without specifying the functional form of Z , the proposed estimator is the most efficient among the three methods. Sometimes over 50% efficiency gain is achieved. In the second model, where the adjusted covariate takes a linear form in the model, the proposed method is almost as efficient as the results from fitting the true model in terms of the estimation of α and β . In the third model, where the adjusted covariate takes a U-shaped form in the model, the proposed method is again the most efficient one in terms of the estimation of α and β and achieving highest power among the three methods. The efficiency gains of the proposed method seem to increase with the degree of mis-specification of the adjustment covariate effect, as the efficiency gains of the third model is more significant than that of the second model. In general, the proposed method for covariate adjustment seems to be more robust and efficient in our study.

5 A data example: SOLVD study

We illustrate the proposed method to the data from the SOLVD clinical trial, a randomized double-blind trial conducted between June 1986 and March 1989, to compare enalapril to placebo (The SOLVD Investigators 1991). Enalapril was an angiotensin-converting-enzyme inhibitor. The investigators are interested in the effect of enalapril in reducing the mortality and hospitalization in patients with chronic heart failure and ejection fractions ≤ 0.35 .

The study enrolled 2569 patients, 1284 patients were randomly assigned to receive placebo and 1285 were randomized to receive enalapril doses of 2.5–20 mg per day. The follow-up time of patients ranged from 22 to 55 months, with the average 41.4 months. During the time patients were monitored, data on all hospitalizations, survival time and the baseline level of ejection fraction (%) ($EF \times \%$) were recorded. There were 510 deaths in the placebo group with death rate 39.7%, as compared with 452 in the enalapril group, with death rate 35.2%. Table 5 summarizes numbers of hospital admissions and deaths for treatment and control groups.

We are interested in testing the treatment effect in terms of reducing the number of hospitalizations and the mortality. The results of applying joint model of Huang and Wang (2004) with the treatment indicator as the only covariate is provided in Table 6. We also considered the baseline auxiliary covariate, level of ejection fraction ($EF \times \%$) with the histogram in Fig. 1. The ejection fraction can be considered to be independent of the treatment

indicator due to the small sample correlation coefficient -0.007 . This covariate may be auxiliary to the outcome as reflected from the significance of HW's model where the ejection fraction is considered as a regressor. The results for the proposed method and covariate adjustment as the regressor are also shown in Table 6.

Nonparametric bootstrap with bootstrap size 50 were used to estimate the standard error of $\widehat{\alpha}$ and $\widehat{\beta}$. Both estimated covariate effects are statistically significant with p values <0.001 for all three methods. The analysis indicates that enalapril is effective in terms of reducing the mortality and hospitalizations for heart failure in patients with chronic congestive heart failure and low ejection fractions.

The estimation results for three methods are fairly close for both hospitalization and death. The estimation error of the proposed method is smaller than HW's method without adjusting covariates, as expected, although the efficiency improvement is not significantly large. We conjecture that this may be because the prediction effect of the ejection fraction is relatively small, due to the rather small magnitude of the estimation of the log intensity ratio and the log hazard ratio (-0.024 and 0.040) together with the range of the ejection fraction from 5 to 35.

6 Discussion

In this paper, we propose a joint covariate-adjusted score test statistic based on a joint model of correlated recurrent events and a terminal event. The test of interest is focused on the comparison of the treatment effect on the intensity of recurrent events processes and the cumulative hazard of the terminal event. We propose efficient influence functions for the treatment indicators and develop an effective inference procedure. Compared with traditional covariate-adjusted methods, this method does not assume the functional form of the adjusted covariates and is more robust. In practice, except the case where sufficient evidence supports that the adjusted covariate appear in the true model as a linear regressor, we would recommend using the proposed method, as it is generally more robust and efficient compared with the traditional covariate adjustment approach.

In the numerical studies, we consider one dimensional auxiliary covariates Z mainly due to computation concerns, since the polynomial basis or spline basis are only vectors of $k - 1$ elements in the sieved approach and can be easily solved by a least square estimation. The computation could be more complicated when Z is composed of several covariates. How to choose an appropriate expansion (approximation) for adjusting multivariate auxiliary covariates is a practical and challenging problem. With the current methodology that only one dimensional covariate can be adjusted, the following two step procedure could be a possible remedy: in the first step, we can reduce the dimension of the covariates with certain dimension reduction method, such as choosing the first principle component; in the second step, the proposed covariate adjustment methodology can be performed with the resulted one dimensional component from the first step. This two-step procedure could utilize the information of the multi-dimensional covariates and deserves future research as well.

Acknowledgments

This research is partly supported by National Institutes of Health NHLBI Grant R01 HL-57444.

Appendix

We give Theorems 4.3 and 4.5 of Tsiatis (2006) for ease of readers.

Theorem 4.3 If a semiparametric RAL estimator for β exists, the efficient influence function of this estimator must belong to the space of influence functions, the linear variety $\{\phi(Z) + \mathcal{T}^\perp\}$, where $\phi(Z)$ is the influence function of any semiparametric RAL estimator for β and \mathcal{T} is the semiparametric tangent space, any if an RAL estimator for β exists that achieves the semiparametric efficiency bound (i.e., a semi-parametric efficient estimator), then the influence function of this estimator must be the unique and well-defined element

$$\phi_{\text{eff}}(Z) = \phi(Z) - \text{proj}\{\phi(Z) | \mathcal{T}^\perp\} = \text{proj}\{\phi(Z) | \mathcal{T}\}.$$

Theorem 4.5 The tangent space \mathcal{F} for the nonparametric model, i.e., the entire Hilbert space \mathcal{H} , is equal to

$$\mathcal{F} = \mathcal{H} = \mathcal{F}_1 \oplus \dots \oplus \mathcal{F}_m,$$

where

$$\mathcal{T}_1 = \{\alpha_1^{q \times 1}(Z^{(1)}) \in \mathcal{H} : E\{\alpha_1^{q \times 1}(Z^{(1)})\} = 0^{q \times 1}\}$$

and

$$\mathcal{T}_j = \{\alpha_j^{q \times 1}(Z^{(1)}, \dots, Z^{(j)}) \in \mathcal{H} : E\{\alpha_j^{q \times 1}(Z^{(1)}, \dots, Z^{(j)}) | Z^{(1)}, \dots, Z^{(j-1)}\} = 0^{q \times 1}\},$$

$j = 2, \dots, m$, and $\mathcal{T}_j, j = 1, \dots, m$ are mutually orthogonal spaces. Equivalently, the linear space \mathcal{T}_j can be defined as the space

$$\{h_{*j}^{q \times 1}(Z^{(1)}, \dots, Z^{(j)}) \in \mathcal{H} : E\{h_{*j}^{q \times 1}(Z^{(1)}, \dots, Z^{(j)}) | Z^{(1)}, \dots, Z^{(j-1)}\} = 0^{q \times 1}\},$$

for all square-integrable functions $h_{*j}^{q \times 1}(\cdot)$ of $Z^{(1)}, \dots, Z^{(j)}$.

In addition, any element $h(Z^{(1)}, \dots, Z^{(m)}) \in \mathcal{H}$ can be decomposed into orthogonal elements

$$h = h_1 + \dots + h_m,$$

where

$$\begin{aligned} h_1(Z^{(1)}) &= E\{h(\cdot) | Z^{(1)}\}, \\ h_j(Z^{(1)}, \dots, Z^{(j)}) &= E\{h(\cdot) | Z^{(1)}, \dots, Z^{(j)}\} - E\{h(\cdot) | Z^{(1)}, \dots, Z^{(j-1)}\}, \end{aligned}$$

for $j = 2, \dots, m$, and $h_j(\cdot)$ is the projection of h onto \mathcal{T}_j , i.e., $h_j(\cdot) = \text{proj}\{h(\cdot) | \mathcal{T}_j\}$.

References

- Andersen PK, Gill RD. Cox's regression model for counting processes: a large sample study (Com: P1121–1124). *Ann Stat* 1982;10:1100–1120.
- Bickel, PJ.; Klaassen, CAJ.; Ritov, Y., et al. *Efficient and adaptive estimation for semiparametric models*. Springer-Verlag; New York: 1998.
- Cook RJ, Lawless JF. Comment on “an overview of statistical methods for multiple failure time data in clinical trials”. *Stat Med* 1997;16:841–843.
- Ghosh D, Lin DY. Nonparametric analysis of recurrent events and death. *Biometrics* 2000;56:554–562. [PubMed: 10877316]
- Ghosh D, Lin DY. Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* 2003;59:877–885. [PubMed: 14969466]
- Huang CY, Wang MC. Joint modeling and estimation for recurrent event processes and failure time data. *J Am Stat Assoc* 2004;99:1153–1165.
- Huang X, Liu L. A joint frailty model for survival and gap times between recurrent events. *Biometrics* 2007;63:389–397. [PubMed: 17688491]
- Kong FH, Slud E. Robust covariate-adjusted Logrank tests. *Biometrika* 1997;84:847–862.
- Kuhl, M.; Bhairgond, P. Nonparametric estimation of nonhomogeneous poisson processes using wavelets. In: Joines, JA.; Barton, RR.; Kang, K.; Fishwick, PA., editors. *Proceedings of the 2000 winter simulation conference*; Orlando. 2000.
- Lawless JF, Nadeau JC. Nonparametric estimation of cumulative mean functions for recurrent events. *Technometrics* 1995;37:158–168.
- Leon S, Tsiatis AA, Davidian M. Semiparametric estimation of treatment effect in a Pretest-posttest study. *Biometrics* 2003;59:1046–1055. [PubMed: 14969484]
- Lewis P, Shedler G. Simulation of nonhomogeneous poisson processes by thinning. *IBM J Res Dev* 1979;32:403–413.
- Li Z. Covariate adjustment for non-parametric tests for censored survival data. *Stat Med* 2001;20:1843–1853. [PubMed: 11406845]
- Lin DY, Wei LJ, Yang I, et al. Semiparametric regression for the mean and rate functions of recurrent events. *J R Stat Soc B* 2000;62:711–730.
- Liu L, Huang X. The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Stat Med* 2008;27:2665–2683. [PubMed: 17910008]
- Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. *Biometrics* 2004;60:747–756. [PubMed: 15339298]
- Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *J Am Stat Assoc* 1993;88:811–820.
- Rondeau V. Joint frailty models for recurring events and death using maximum penalized likelihood estimation: application on cancer events. *Biostatistics* 2007;8(4):708–721. [PubMed: 17267392]
- Slud E. Relative efficiency of the log rank test within a multiplicative intensity model. *Biometrika* 1991;78:621–630.
- Struthers CA, Kalbfleisch JD. Misspecified proportional hazard models. *Biometrika* 1986;73:363–369.
- The SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;325:293–302. [PubMed: 2057034]
- Tsiatis, AA. *Semiparametric theory and missing data*. Springer; New York: 2006.
- Tsiatis AA, Rosner GL, Trichler DL. Group sequential tests with censored survival data adjusting for covariates. *Biometrika* 1985;72:365–373.
- Wei LJ, Lin DY, Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J Am Stat Assoc* 1989;84:1065–1073.
- Ye Y, Kalbfleisch JD, Schaubel DE. Semiparametric analysis of correlated recurrent and terminal events. *Biometrics* 2007;63:78–87. [PubMed: 17447932]

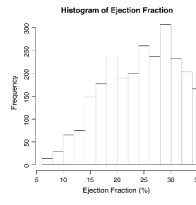


Fig. 1.
The histogram of ejection fraction in SOVLD data

Table 1

Simulation results for the first model based on 1000 simulation samples with Monte Carlo error 0.013

| Method | True value | Mean | Empirical standard deviation | SE | Type I error/Power | |
|-----------------------------------|------------|------|------------------------------|-------|--------------------|-------|
| HW's method | α | 0 | 0.009 | 0.247 | 0.222 | 0.063 |
| | β | 0 | -0.012 | 0.436 | 0.419 | 0.054 |
| | α | -0.4 | -0.402 | 0.247 | 0.226 | 0.262 |
| | β | -0.8 | -0.802 | 0.419 | 0.428 | 0.457 |
| | α | -1 | -1.037 | 0.292 | 0.273 | 0.979 |
| | β | -1.5 | -1.542 | 0.499 | 0.501 | 0.889 |
| Proposed method | α | 0 | 0.007 | 0.237 | 0.218 | 0.064 |
| | β | 0 | -0.009 | 0.363 | 0.379 | 0.040 |
| | α | -0.4 | -0.400 | 0.233 | 0.207 | 0.297 |
| | β | -0.8 | -0.804 | 0.316 | 0.327 | 0.731 |
| | α | -1 | -1.036 | 0.277 | 0.255 | 0.992 |
| | β | -1.5 | -1.548 | 0.393 | 0.395 | 0.983 |
| Adjusted covariate as a regressor | α | 0 | 0.007 | 0.238 | 0.219 | 0.061 |
| | β | 0 | 0.079 | 0.390 | 0.401 | 0.037 |
| | α | -0.4 | -0.399 | 0.233 | 0.208 | 0.284 |
| | β | -0.8 | -0.813 | 0.473 | 0.465 | 0.422 |

Table 2

Simulation results for the second model

| Method | True value | Mean | Empirical standard deviation | SE | Type I error/Power | |
|-----------------------------------|------------|------|------------------------------|-------|--------------------|-------|
| HW's method | α | 0 | -0.008 | 0.267 | 0.234 | 0.065 |
| | β | 0 | 0.009 | 0.377 | 0.398 | 0.043 |
| Proposed method | α | -0.4 | -0.382 | 0.250 | 0.244 | 0.361 |
| | β | -0.8 | -0.789 | 0.446 | 0.425 | 0.471 |
| Proposed method | α | 0 | -0.007 | 0.246 | 0.222 | 0.062 |
| | β | 0 | 0.009 | 0.377 | 0.398 | 0.058 |
| Adjusted covariate as a regressor | α | -0.4 | -0.380 | 0.235 | 0.234 | 0.371 |
| | β | -0.8 | -0.788 | 0.446 | 0.424 | 0.478 |
| Adjusted covariate as a regressor | α | 0 | -0.008 | 0.245 | 0.221 | 0.060 |
| | β | 0 | 0.012 | 0.376 | 0.397 | 0.050 |
| Adjusted covariate as a regressor | α | -0.4 | -0.382 | 0.235 | 0.233 | 0.378 |
| | β | -0.8 | -0.793 | 0.445 | 0.423 | 0.491 |

Table 3

Simulation results for the third model

| Method | True value | Mean | Empirical standard deviation | SE | Type I error/Power | |
|-----------------------------------|------------|------|------------------------------|-------|--------------------|-------|
| HW's method | α | 0 | -0.036 | 0.443 | 0.427 | 0.061 |
| | β | 0 | 0.053 | 0.521 | 0.491 | 0.063 |
| Proposed method | α | -0.4 | -0.431 | 0.477 | 0.457 | 0.317 |
| | β | -0.8 | -0.827 | 0.528 | 0.533 | 0.418 |
| Proposed method | α | 0 | -0.011 | 0.342 | 0.368 | 0.055 |
| | β | 0 | 0.033 | 0.440 | 0.408 | 0.066 |
| Adjusted covariate as a regressor | α | -0.4 | -0.401 | 0.335 | 0.301 | 0.365 |
| | β | -0.8 | -0.801 | 0.410 | 0.435 | 0.461 |
| Adjusted covariate as a regressor | α | 0 | -0.041 | 0.391 | 0.366 | 0.064 |
| | β | 0 | 0.062 | 0.477 | 0.495 | 0.057 |
| Adjusted covariate as a regressor | α | -0.4 | -0.412 | 0.428 | 0.439 | 0.350 |
| | β | -0.8 | -0.807 | 0.499 | 0.462 | 0.427 |

Table 4

Relative efficiencies with respect to HW's model with only treatment indicator as the regressor for the settings recorded in Table 1

| Model | Null | | Alternative | |
|------------------------------------|----------------------------|---------------------------|----------------------------|---------------------------|
| | <i>α</i> | <i>β</i> | <i>α</i> | <i>β</i> |
| <i>Model 1</i> | | | | |
| Auxiliary covariate as a regressor | 1.077 | 1.249 | 1.124 | 0.788 |
| Proposed model (Local Alt) | 1.086 | 1.442 | 1.124 | 1.758 |
| Proposed model (Larger Alt) | | | 1.111 | 1.612 |
| <i>Model 2</i> | | | | |
| Auxiliary covariate as a regressor | 1.188 | 1.005 | 1.132 | 1.004 |
| Proposed model | 1.178 | 1.000 | 1.132 | 1.000 |
| <i>Model 3</i> | | | | |
| Auxiliary covariate as a regressor | 1.283 | 1.092 | 1.241 | 1.119 |
| Proposed model | 1.677 | 1.402 | 2.028 | 1.659 |

Table 5

Frequency of hospitalization and death for congestive heart failure

| Treatment group | Number of subjects | Number of hospitalizations | | | | Average number of hospitalizations | |
|-------------------|--------------------|----------------------------|------|------|-------|------------------------------------|-------|
| | | 0 | 1 | 2 | 3 ≥ 4 | | |
| Placebo | 1284 | 814 | 236 | 118 | 59 | 57 | 0.731 |
| (%) | 100 | 63.4 | 18.4 | 9.2 | 4.6 | 4.4 | |
| Placebo (death) | 510 | 266 | 113 | 70 | 32 | 29 | 0.974 |
| (%) | 100 | 52.2 | 22.2 | 13.7 | 6.3 | 5.7 | |
| Enalapril | 1285 | 953 | 175 | 69 | 40 | 48 | 0.508 |
| (%) | 100 | 74.2 | 13.6 | 5.4 | 3.1 | 3.7 | |
| Enalapril (death) | 452 | 281 | 80 | 38 | 25 | 28 | 0.790 |
| (%) | 100 | 62.2 | 17.7 | 8.4 | 5.5 | 6.2 | |

Table 6

Estimation results for SOLVD data

| Covariate adjustment? | Parameter | Estimates | SE | Covariate estimates | Covariate standard error |
|-----------------------|---------------------|-----------|--------|---------------------|--------------------------|
| No | $\hat{\alpha}_{hw}$ | -0.428 | 0.0938 | | |
| As a regressor | $\hat{\alpha}_c$ | -0.430 | 0.0925 | -0.024 | 0.0064 |
| Proposed method | $\hat{\alpha}_p$ | -0.433 | 0.0928 | | |
| No | $\hat{\beta}_{hw}$ | -0.207 | 0.0857 | | |
| As a regressor | $\hat{\beta}_c$ | -0.213 | 0.0814 | -0.040 | 0.0058 |
| Proposed method | $\hat{\beta}_p$ | -0.198 | 0.0822 | | |