Published in final edited form as:

Biometrics. 2017 December ; 73(4): 1150–1160. doi:10.1111/biom.12693.

Joint Modeling of Longitudinal, Recurrent Events and Failure Time Data for Survivor's Population

Qing Cai^{1,*}, Mei-Cheng Wang^{1,**}, and Kwun Chuen Gary Chan^{2,***}

¹Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, U.S.A

²Department of Biostatistics, University of Washington, Seattle, Washington 98195, U.S.A

Summary

Recurrent events together with longitudinal measurements are commonly observed in follow-up studies where the observation is terminated by censoring or a primary failure event. In this article, we developed a joint model where the dependence of longitudinal measurements, recurrent event process and time to failure event is modeled through rescaling the time index. The general idea is that the trajectories of all biology processes of subjects in the survivors' population are elongated or shortened by the rate identified from a model for the failure event. To avoid making disputing assumptions on recurrent events or biomarkers after the failure event (such as death), the model is constructed on the basis of survivors' population. The model also possesses a specific feature that, by aligning failure events as time origins, the backward-in-time model of recurrent events and longitudinal measurements shares the same parameter values with the forward time model. The statistical properties, simulation studies and real data examples are conducted. The proposed method can be generalized to analyze left-truncated data.

Keywords

Backward process model; Counting process; Informative censoring; Informative sampling; Left truncation; Time-adjusted model

1. Introduction

Recurrent events together with longitudinal measurements are frequently encountered in follow-up studies. In biomedical applications, two types of longitudinal measurements are commonly observed: (i) repeated measurements collected at sampling times, and (ii) marker measurements observed when recurrent events occur. In case (i), longitudinal measurements are assumed to be a stochastic process which exists continuously over time, such as CD4 cell counts in HIV studies or other disease-related biomarkers, where recurrent events are

* qcai3@jhu.edu. ** mcwang@jhu.edu. *** kcgchan@u.washington.edu.

Supplementary Materials

Web Appendices A (Proof of Equivalence of Model (2) and (3)), B (Proof of Equation (7)), C (Proof of Theorem 3.1), D (Simulations on Sensitivity Analysis) and E (Simulation Studies of Baseline Processes Estimation) referenced in Sections 2, 3, 5, and 7 are available with the online version of this article at the *Biometrics* website on Wiley Online Library. The instruction for BIOCARD data application (Web Appendix F) and a R sample code for simulation studies are also included.

sampling times. In case (ii), typical examples are studies with repeated marker measurements observed upon the occurrence of recurrent events, where both recurrent events and marker measurements are of scientific interests; an example is the medical charge upon the occurrence of hospitalization. For either case, the observation of longitudinal measurements and recurrent events could be terminated by censoring or a primary failure event such as death. Despite the difference in the data generating mechanisms, these two kinds of data share the same notations. This article presents a semiparametric joint model framework for both types of longitudinal measurements, (i) and (ii), with covariates information.

In the absence of informative sampling times, many authors (Wulfsohn and Tsiatis, 1997; Henderson et al., 2000; Xu and Zeger, 2001; Song, Davidian, and Tsiatis, 2002; Vonesh, Greene, and Schluchter, 2006; Song and Wang, 2008) used the shared frailty model to analyze longitudinal measurements and time-to-event data, where multiple layers of models with shared underlining random variables are created for different data components. The general idea of shared frailty models is that different data components are assumed to be independent of each other conditioning on the shared random variables, and hence the correlation within the data structure is explained by the shared random variables. However, since these models assume that the sampling times are independent, none of them can be applied to longitudinal measurements of type (ii). Abundant attempts have been adopted to jointly model longitudinal measurements, recurrent events, and time to failure events data. The shared frailty models (Sun, Sun, and Liu, 2007; Liu, Huang, and O'Quigley, 2008; Liu and Huang, 2009; Kim et al., 2012; Sun et al., 2012) form a popular approach as well. However, some of these models may not be applicable to randomly sampled measurements, that is, type (i) data, due to specific model restrictions for recurrent event processes. In addition, various approaches have been proposed for jointly modeling longitudinal measurements and recurrent events with uncorrelated failure event (Lin and Ying, 2001; Lin, Scharfstein, and Rosenheck, 2004; Sun et al., 2005; Liang, Lu, and Ying, 2009), and modeling recurrent events together with informative failure event jointly (Wang, Qin, and Chiang, 2001; Huang and Wang, 2004; Liu, Wolfe, and Huang, 2004; Ye, Kalbfleisch, and Schaubel, 2007; Zeng and Cai, 2010). Most of these approaches consider the data correlation by using the shared frailty method, except the models in Lin and Ying (2001), Lin et al. (2004) and Sun et al. (2005).

In this article, we consider a joint model of longitudinal measurements, recurrent events and time to failure events, which is a useful alternative to the shared frailty method and is applicable to both two types of longitudinal measurements (i) and (ii). The time rescaling technique was previously adopted by Ghosh and Lin (2003) and Huang and Wang (2003) for modeling recurrent events together with informative failure events. Luo, Wang, and Huang (2008) also provides a detailed comparison of different models. Our model has the features that the trajectories of all biology processes are targeted on subjects in the survivors' population, where a subject's failure time is elongated or shortened by the rate identified from a failure time model for the failure event. Hence, our method can be developed without disputing assumption on the recurrent events or biomarkers after death. In contrast with the shared frailty method, our model does not involve latent variable and the proposed methodological procedure is computationally simpler.

Besides the forward time model, by aligning failure events as time origins, we will also study statistical model and inference of recurrent events and longitudinal measurements backward in time. For many diseases, including Alzheimer's Disease, the changes in biomarker performance before diagnosis of disease is widely recognized as an important index for predicting the disease (Hall et al., 2000; Wilson et al., 2007). The conventional stochastic process models are always forward in time and not designed to study terminal behavior of processes. However, backward process models are more relevant for studying these terminal behavior (Chan and Wang, 2010; Chan and Wang, In Press). In this article, we will show that our model, though developed in forward time scale, has the consistent interpretation in the backward models under proper assumptions.

In Section 4, the method is generalized to the case where left truncation in failure time data is present. It handles the sampling designs when the study recruits only subjects who have experienced the initiating event but have not experienced the failure event. Under left truncation, it is well known that the time from initiating event to failure event tends to be longer than the failure time from the target population. We will use a data example from the Alzheimer's Disease to illustrate the proposed method.

The article is organized as follows. Section 2.1 introduces the joint model and discusses the interpretation of the joint model. An alternative model is presented in Section 2.2, and we show that the model has consistent interpretation in backward time perspective. Section 3 describes the estimation procedures and develops statistical properties of the proposed semiparametric estimators. The generalization to left truncation is discussed in Section 4. In Section 5, we evaluate the performance of the estimating methods by simulation studies. Real data from an Alzheimer's Disease study and an AIDS study are used as illustrating examples for data analysis in Section 6. Additional discussions are included in Section 7 to conclude the article.

2. Survivors Model

2.1. A Time-Adjusted Forward Model

Let T be the time from an initial event to the failure event which is in continuous time scale, $\mathbf{Z}$ a $p \times 1$ vector of covariates, and $R^*(t)$ the counting process of the number of sampling times at or before time t . The longitudinal process $Y(t)$ is measured repeatedly at time t where $dR^*(t) = 1$. Suppose that the study is conducted in the time interval $[0, \tau]$, and potentially, the data information of $\{T, R^*(t), Y(t)\}$ is observed for $0 \leq t \leq \tau$. In reality, due to limitation in experimental design or other reasons, the data are subject to right censoring and only partially observed. Let C_R be the censoring time for observing the sampling times and the longitudinal measurements $\{R^*(\cdot), Y(\cdot)\}$. Let C_T be the censoring time for the failure event. When framing the proposed model, we shall consider only the recurrent events and the longitudinal measurements occurring before the failure event, that is, $t \leq T$.

Consider the following assumption:

$$(A1) \quad \{T, R^*(\cdot), Y(\cdot)\} \text{ is independent of } (C_T, C_R) \text{ conditioning on } \mathbf{Z}.$$

Under Assumptions (A1), we consider a time-adjusted forward model to characterize the joint relationship of $\{T, R^*(\cdot), Y(\cdot)\}$: At any $t \geq 0$, assume

$$\begin{cases} \ln T = \beta'_T \mathbf{Z} + \ln T_0, & \text{where } \mathbf{Z} \text{ is independent of } T_0 \\ E \left\{ R^* \left(e^{\beta'_R \mathbf{Z}} dt \right) \mid T \geq t e^{\beta'_R \mathbf{Z}}, \mathbf{Z} \right\} = e^{\beta'_R \mathbf{Z}} E \{ R_0^*(dt) \mid T_0 \geq t \}, \\ E \left\{ Y \left(t e^{\beta'_Y \mathbf{Z}} \right) \mid T \geq t e^{\beta'_R \mathbf{Z}}, R^* \left(e^{\beta'_R \mathbf{Z}} dt \right) = 1, \mathbf{Z} \right\} = E \{ Y_0(t) \mid T \geq t, R_0^*(dt) = 1 \} + \beta'_Y \mathbf{Z}, \end{cases} \quad (1)$$

where β_T , β_R , and β_Y is a $p \times 1$ vector of parameters, and T_0 , $R_0^*(\cdot)$ and $Y_0(\cdot)$ are respectively the baseline failure time, the baseline recurrent event process and the baseline longitudinal process. The three components $\{T_0, R_0^*(\cdot), Y_0(\cdot)\}$ are distribution-free and possibly correlated. Of note, the notations $\{T_0, R_0^*(\cdot), Y_0(\cdot)\}$ represent individual level variables or processes, and $\{T_{0i}, R_{0i}^*(\cdot), Y_{0i}(\cdot)\}$ will be used in later sections as variables or

processes for subject i . In Model (1), the notation $R^* \left(e^{\beta'_R \mathbf{Z}} dt \right)$ is used to indicate the jumps of the recurrent event process on the adjusted time scale, that is,

$$R^* \left(e^{\beta'_R \mathbf{Z}} dt \right) = R^* \left(e^{\beta'_R \mathbf{Z}} t \right) - R^* \left(e^{\beta'_R \mathbf{Z}} t^- \right).$$

Model (1) has an important feature as a *survivors' model*, in which only those who have not experienced the failure event are included when modeling the recurrent events and longitudinal measurements. The model assumes that the trajectories of a subject's recurrent event and longitudinal measurement processes in the survivors' population are elongated or shortened by the regression function identified from the accelerated failure time model (Kalbfleisch and Prentice, 1980). More precisely, the time scale in $\{R^*(\cdot), Y(\cdot)\}$ is transformed from $[0, T_0]$ to $[0, T]$ using $e^{\beta'_R \mathbf{Z}}$ as the multiplicative adjustment. Besides the covariate effect on time-rescaling, covariates are also allowed to influence the rate function of $R^*(\cdot)$ by a multiplicative term $\exp(\beta'_R \mathbf{Z})$ and influence the trajectory function $Y(\cdot)$ by an additive term $\beta'_Y \mathbf{Z}$, respectively.

Note that Assumption (A1) specifies the independent censoring assumption. The condition $(T > t)$ in Model (1) has the practical interpretation that $\{R^*(\cdot), Y(\cdot)\}$ is required to exist only for the time before the failure event, and therefore allows Model (1) to capture the covariate effects on the survivors' population and avoids the arguable issue of longitudinal variables or recurrent events after the failure time. Additionally, the survivors' population changes when t varies, and this special characteristic is indicated by the condition $(T > t)$. The conditioning event $R^* \left(\exp(\beta'_R \mathbf{Z}) dt \right) = 1$ in the third layer of Model (1) is a natural condition for type (ii) data, since the marker measurement $Y \left(t \exp(\beta'_R \mathbf{Z}) \right)$ exists and is observed given the occurrence of a recurrent event. For type (i) longitudinal measurement, Model (1) models

the mean trajectory and covariate effect only for the observed $Y(\cdot)$, but the result generalizes to the underlying $Y(\cdot)$ if $Y(\cdot)$ and $R(\cdot)$ are independent given Z . Similar conditions are employed for type (i) data in Lin and Ying (2001).

As mentioned in Section 1, Model (1) is applicable to the two types of longitudinal measurements (i) and (ii), where the rate of sampling process in the rescaled time index is changed by a multiplier $\exp(\beta'_T Z)$ compared to baseline. For the special case when the rate of sampling points is pre-planned and constant overtime, the rate of $R^*(t)$ is not influenced by covariates and we have $\beta_R = \beta_T$. In this case, the non-zero β_R is the rescaling effect rather than the covariate effect. Note that the case $\beta_R = \mathbf{0}$ implies that the recurrent event process is affected by the covariates only through rescaling the time to failure event, where the recurrent event process is stretched or compressed. Thus, a positive or negative β_R reflects the additional inflation or deflation of the frequency of $R^*(\cdot)$ explained by the covariates.

2.2. An Alternative Model and Its Backward Property

By rescaling time index, we construct an alternative model, termed as Model (2). This model is slightly stronger than Model (1), and possesses an attractive feature of having consistent interpretation in both the forward and backward models as will be explained later. Under the independent censoring assumption (A1), Model (2) assumes

$$\left\{ \begin{array}{l} \ln T = \beta'_T Z + \ln T_0, \quad \text{where } Z \text{ is independent of } T_0 \\ E \left\{ R^* \left(t e^{\beta'_T Z} \right) | T, Z \right\} = e^{\beta'_R Z} E \{ R_0^*(dt) | T_0 \}, \quad t \in [0, T_0] \\ E \left\{ Y \left(t e^{\beta'_T Z} \right) | T, R^* \left(e^{\beta'_T Z} dt \right) = 1, Z \right\} = E \{ Y_0(t) | T_0, R_0^*(dt) = 1 \} + \beta'_Y Z, \quad t \in [0, T_0] \end{array} \right.$$

(2)

Model (2) extends the joint model of failure time and recurrence events studied by Huang and Wang (2003) to a joint model with the additional longitudinal measurement $Y(\cdot)$. Note that by conditioning on T , the baseline failure time T_0 plays a role similar to a subjective-specific random effect or frailty, and the model identifies the regression effect on Z . Huang and Wang (2003) indicated that their recurrent event model extends to

$$E \left\{ R^* \left(e^{\beta'_T Z} dt \right) | T \geq t e^{\beta'_T Z}, Z \right\} = e^{\beta'_R Z} E \{ R_0^*(dt) | T_0 \geq t \}.$$

In our case, under Model (2), the longitudinal measurement satisfies

$$\begin{aligned}
& E \left\{ Y \left(t e^{\beta'_{\tau} \mathbf{Z}} \right) | T \geq t e^{\beta'_{\tau} \mathbf{Z}}, R^* \left(e^{\beta'_{\tau} \mathbf{Z}} dt \right) = 1, \mathbf{Z} \right\} \\
&= E \left[I_{\left[T \geq t e^{\beta'_{\tau} \mathbf{Z}} \right]} E \left\{ Y \left(t e^{\beta'_{\tau} \mathbf{Z}} \right) | T, R^* \left(e^{\beta'_{\tau} \mathbf{Z}} dt \right) = 1, \mathbf{Z} \right\} | R^* \left(e^{\beta'_{\tau} \mathbf{Z}} dt \right) = 1, \mathbf{Z} \right] \times \left\{ P \left(T \geq t e^{\beta'_{\tau} \mathbf{Z}} \right) | R^* \left(e^{\beta'_{\tau} \mathbf{Z}} dt \right) = 1, \mathbf{Z} \right\}^{-1} \\
&= E \left[I_{\left[T_0 \geq t \right]} \left\{ E \left(Y_0(t) | T_0, R_0^*(dt) = 1 \right) + \beta'_{\gamma} \mathbf{Z} \right\} | R_0^*(dt) = 1, \mathbf{Z} \right] \left\{ P \left(T_0 \geq t | R_0^*(dt) = 1, \mathbf{Z} \right) \right\}^{-1} \\
&= E \left\{ \left(Y_0(t) | T_0 \geq t, R_0^*(dt) = 1 \right) + \beta'_{\gamma} \mathbf{Z} \right\}.
\end{aligned}$$

Thus, the Tregression model of $Y(\cdot)$ in Model (1) would be validated by Model (2), and a similar argument extends to the model for $R^*(\cdot)$ to conclude that the validity of Model (2) implies Model (1). Essentially, Model (1) can be thought of as a marginal model of Model (2), since T is used as a part of conditional statistics in the latter but not in the former.

Interestingly, as a special property, Model (2) implies the validity of a backward time model with the same parameter value. The backward time model can be used to study the terminal behaviors of the biological processes where the failure event is aligned as the time origin and time is counted backward. Backward time models offer a natural and direct way to study terminal behavior of recurrent markers. Using the human immunodeficiency virus (HIV) infection as an illustrating example, it is scientifically and clinically interesting to understand the pattern of the frequency and severity of opportunistic infections before death. Evidence suggests that HIV-infected patients experienced higher frequency of AIDS-defining events before death, where the frequencies could vary with gender, risk behaviors or geographic location (Chan et al., 1995). Similar terminal behaviors are studied for Alzheimer's Disease (Wilson et al., 2007), renal disease (Usvyat et al., 2013) and functional decline in the general population (Lunney et al., 2003). To set notation, let the backward time index be denoted by t^B which stands for the time counted retrospectively from the failure event. Define the backward process of the longitudinal measurement as $Y^B(t^B; T) = Y(T - t^B)$ and the backward recurrent event process as $R^B(t^B; T) = R^*(T) - R^*((T - t^B)^-)$, where $0 \leq t^B \leq T$.

Under Assumptions (A1), by aligning failure events as time origins and counting time backward, it can be proved that Model (2) holds if and only if the following backward model holds: For any $t \in [0, T_0]$,

$$\begin{cases} \ln T = \beta'_{\tau} \mathbf{Z} + \ln T_0, & \text{where } \mathbf{Z} \text{ is independent of } T_0 \\ E \left\{ R^B \left(e^{\beta'_{\tau} \mathbf{Z}} dt^B; T \right) | T, \mathbf{Z} \right\} = e^{\beta'_{\tau} \mathbf{Z}} E \left\{ R_0^B \left(dt^B; T_0 \right) | T_0 \right\}, \\ E \left\{ Y^B \left(t^B e^{\beta'_{\tau} \mathbf{Z}}; T \right) | T, R^B \left(e^{\beta'_{\tau} \mathbf{Z}} dt^B; T \right) = 1, \mathbf{Z} \right\} = E \left\{ Y_0^B \left(t^B; T_0 \right) | T_0, R_0^B \left(dt^B; T_0 \right) = 1 \right\} + \beta'_{\gamma} \mathbf{Z}. \end{cases}$$

(3)

Model (3) assumes that the trajectories of a subject's backward recurrent event and longitudinal measurement processes are elongated or shortened by the scale identified from

the accelerated failure time model, where covariates influence the rate of $R^B(\cdot)$ by a multiplicative term $\exp(\beta'_R \mathbf{Z})$ and influence the measurement $Y^B(\cdot)$ by an additive term $(\beta'_Y \mathbf{Z})$, respectively. Model (3) is consistent with Model (2) in the sense that they are equivalent to each other and share the same parameter values. This property implies that estimation of Model (3) can be derived through estimation of Model (2), as, what we will establish in the next section. The proof for equivalence between Models (2) and (3) is provided in the Web Appendix A.

3. Estimation

In this section, we introduce and discuss the estimation procedures of Models (1) and (2). Define $X_T = T \wedge C_T$, $X_R = T \wedge C_R$, and $R(t) = R^*(t \wedge X_R)$, where $a \wedge b = \min(a, b)$. Let $\Delta = I_{[T \leq C_T]}$ be the censoring indicator of the failure event. Assume the observations $\{X_{Ti}, X_{Ri}, R_i(\cdot), Y_i(\cdot), Z_i\}$, $i = 1, \dots, n$, are independent and identically distributed (i.i.d.). For most applications, the censoring time for recurrent events and longitudinal measurements is either the same as or less than the censoring time for the failure event. We therefore assume $P(C_R \leq C_T) = 1$.

We construct three-layer estimation functions for β_T , β_R , β_Y , separately but in a certain ordering, for Model (1). As the first step, the weighted log-rank estimating equations (Tsiatis, 1990) is popular among other approaches (Buckley and James, 1979; Ritov, 1990) for estimating β_T in the AFT model. We define the adjusted risk-set indicator of failure event

by $\xi_{Ti}(t; \beta_T) = I_{[X_{Ti} \geq t \exp(\beta'_T \mathbf{Z}_i)]}$ and the adjusted counting process of failure event by $N_{Ti}(t; \beta_T) = I_{[X_{Ti} \geq t \exp(\beta'_T \mathbf{Z}_i)]} \Delta_i$. The estimating function is

$$U_T(\beta_T) = \sum_{i=1}^n \int_0^\infty W_T(t; \beta_T) \left\{ Z_i - \frac{\sum_{j=1}^n \xi_{Tj}(t; \beta_T) Z_j}{\sum_{j=1}^n \xi_{Tj}(t; \beta_T)} \right\} dN_{Ti}(t; \beta_T), \quad (4)$$

where $W_T(t; \beta_T)$ is a non-negative weight function. A zerocrossing of $U_T(\beta_T) = \mathbf{0}$ exists, termed as, $\hat{\beta}_T$ as the estimator of β_T which is strongly consistent and asymptotically normal under regular conditions (Ying, 1993).

In the next step, we use the approach of Huang and Wang (2003) to construct an estimator for β_R . For given β_T , define the adjusted risk-set indicator of recurrent event process by

$\xi_{Ri}(t; \beta_T) = I_{[X_{Ri} \geq t \exp(\beta'_T \mathbf{Z}_i)]}$ and the adjusted counting process of recurrent event process as $N_{Ri}(t; \beta_T) = R_i(t \exp(\beta'_T \mathbf{Z}_i)) = R_i^*(t \exp(\beta'_T \mathbf{Z}_i) \wedge X_{Ri})$. The estimating function for β_R is

$$U_R(\beta_R; \beta_T) = \sum_{i=1}^n \int_0^\infty W_R(t; \beta_T, \beta_R) \{Z_i - \tilde{Z}(t; \beta_T, \beta_R)\} dN_{Ri}(t; \beta_T), \quad (5)$$

where $\tilde{Z}(t; \beta_T, \beta_R) = \frac{\sum_{j=1}^n Z_j \xi_{Rj}(t; \beta_T) e^{\beta'_R Z_j}}{\sum_{j=1}^n \xi_{Rj}(t; \beta_T) e^{\beta'_R Z_j}}$ and $W_R(t; \beta_T, \beta_R)$ is a non-negative weight function. If $W_R(t; \beta_T, \beta_R)$ does not depend on β_R , $U_R(\beta_T, \beta_R)$ is a monotone function of β_R and its zero-crossing is a consistent estimator of β_R (Huang and Wang, 2003).

Replacing β_R by $\hat{\beta}_T$, we solve the equation $U_R(\beta_R; \hat{\beta}_T) = 0$ instead and denote the zero-crossing solution as $\hat{\beta}_R$.

In the third step of estimation, we estimate the parameters β_Y via an estimation equation of $(\beta_Y, \beta_T, \beta_R)$. The rate function of recurrent events process in survivors' population at baseline is $E\{dR_0^*(t) | T_0 \geq t\} = P(T_0 \geq t)^{-1} dE\{R_0^*(t)\}$, and the baseline cumulative recurrence rate function of survivors' population is $\Lambda_{R_0}(t) = \int_0^t P(T_0 \geq u)^{-1} dE\{R_0^*(u)\}$. To simplify notation, we define $\alpha_0(t) = E\{Y_0(t) | T_0 \geq t, R_0^*(dt) = 1\}$ which is an unspecified function of t , and define

$$\mathcal{A}_0(t) = \int_0^t E\{Y_0(t) R_0^*(du) | T_0 \geq u\} = \int_0^t \alpha_0(u) d\Lambda_{R_0}(u) = \int_0^t \frac{\alpha_0(u)}{P\{T_0 \geq u\}} dE\{R_0^*(u)\}. \quad (6)$$

All the baseline information is structurally combined in $\mathcal{A}_0(t)$, because one can prove that

$$E\left[\left\{Y_i\left(te^{\beta'_T Z_i} dt\right) - \beta'_Y Z_i\right\} dN_{Ri}(t; \beta_T) | \xi_{Ri}(t; \beta_T)\right] = \xi_{Ri}(t; \beta_T) e^{\beta'_R Z_i} d\mathcal{A}_0(t). \quad (7)$$

Therefore, a mean-zero stochastic process for the i th subject can be expressed as

$$M_{Yi}(t; \beta_T, \beta_R, \beta_Y, \mathcal{A}_0) = \int_0^t \left\{Y_i\left(ue^{\beta'_T Z_i}\right) - \beta'_Y Z_i\right\} dN_{Ri}(u; \beta_T) - \int_0^t \xi_{Ri}(u; \beta_T) e^{\beta'_R Z_i} d\mathcal{A}_0(u),$$

which leads to two estimation equations:

$$\sum_{i=0}^n M_{Yi}(t; \beta_T, \beta_R, \beta_Y, \mathcal{A}_0) = 0, \quad (8)$$

$$\sum_0^t \int_0^t W_Y \left(t; \beta_T, \beta_R, \beta_Y \right) \mathbf{Z}_i dM_{Yi} \left(t; \beta_T, \beta_R, \beta_Y, \mathcal{A}_0 \right) = 0 \quad (9)$$

Based on (8), we develop an estimator for $d\mathcal{A}_0(t)$ and $\mathcal{A}_0(t)$, respectively, as

$$d\hat{\mathcal{A}}_0(t) = \frac{\sum_{i=1}^n \left\{ Y_i \left(t e^{\beta'_Y \mathbf{Z}_i} dt \right) - \beta'_Y \mathbf{Z}_i \right\} dN_{Ri} \left(t; \beta_T \right)}{\sum_{j=1}^n \xi_{Rj} \left(t; \beta_T \right) e^{\beta'_R \mathbf{Z}_j}},$$

$$\hat{\mathcal{A}}_0(t) = \sum_{i=1}^n \int_0^t \frac{Y_i \left(u e^{\beta'_Y \mathbf{Z}_i} du \right) - \beta'_Y \mathbf{Z}_i}{\sum_{j=1}^n \xi_{Rj} \left(t; \beta_T \right) e^{\beta'_R \mathbf{Z}_j}} dN_{Ri} \left(u; \beta_T \right). \quad (10)$$

Replacing $\mathcal{A}_0(t)$ by $\hat{\mathcal{A}}_0(t)$ in (9), we create an estimating function of β_Y with given $\{\beta_T, \beta_R\}$:

$$U_Y \left(\beta_Y; \beta_T, \beta_R \right) = \sum_{i=1}^n \int_0^\infty W_Y \left(t; \beta_T, \beta_R, \beta_Y \right) \left\{ \mathbf{Z}_i - \tilde{\mathbf{Z}} \left(t; \beta_T, \beta_R \right) \right\} \left\{ Y_i \left(t e^{\beta'_Y \mathbf{Z}_i} \right) - \beta'_Y \mathbf{Z}_i \right\} dN_{Ri} \left(t; \beta_T \right). \quad (11)$$

Simple algebraic calculation yields

$$U_Y \left(\beta_Y; \beta_T, \beta_R \right) = \sum_{i=1}^n \int_0^\infty W_Y \left(t; \beta_T, \beta_R, \beta_Y \right) \times \left(\mathbf{Z}_i - \tilde{\mathbf{Z}} \left(t; \beta_T, \beta_R \right) \right) dM_{Yi} \left(t; \beta_T, \beta_R, \beta_Y, \mathcal{A}_0 \right).$$

So, by similar arguments of Lin et al. (2000), $n^{-1/2} U_Y(\beta_Y; \beta_T, \beta_R)$ will converge weakly to a zero mean Gaussian process with continuous sample paths. A zero-root of $U_Y(\beta_Y; \beta_T, \beta_R) =$

$\mathbf{0}$ will be a consistent estimator of β_Y . Replacing β_T and β_R by $\hat{\beta}_T$ and $\hat{\beta}_R$ respectively, we

estimate β_Y by the zero-root solution of the equation $U_Y \left(\hat{\beta}_Y; \hat{\beta}_T, \hat{\beta}_R \right) = \mathbf{0}$. If the weight function $W_Y(t; \beta_T, \beta_R, \beta_Y)$ does not depend on β_Y and

$\left\{ \sum_{i=1}^n \int_0^\infty W_Y \left(t; \hat{\beta}_T, \hat{\beta}_R \right) \left(\mathbf{Z}_i - \tilde{\mathbf{Z}} \left(t; \hat{\beta}_T, \hat{\beta}_R \right) \right) \mathbf{Z}'_i dN_{Ri} \left(t; \hat{\beta}_T \right) \right\}$ is invertible, there is a closed form of the solution as

$$\hat{\beta}_Y = \left\{ \sum_{i=1}^n \int_0^\infty W_Y(t; \hat{\beta}_T, \hat{\beta}_R) \left(Z_i - \tilde{Z}(t; \hat{\beta}_T, \hat{\beta}_R) \right) Z_i' dN_{Ri}(t; \hat{\beta}_T) \right\}^{-1} \left\{ \sum_{i=1}^n \int_0^\infty W_Y(t; \hat{\beta}_T, \hat{\beta}_R) \left(Z_i - \tilde{Z}(t; \hat{\beta}_T, \hat{\beta}_R) \right) Y_i \left(t e^{\beta_T' Z_i} \right) dN_{Ri}(t; \hat{\beta}_T) \right\},$$

which substantially relaxes the burden of computation.

Under Model (2), the estimation procedures are essentially the same as those under Model (1) because Model (1) is induced from Model (2): By re-defining

$$d\mathcal{A}_0(t) = E[Y_0(t) R_0^*(dt) | T_0 \geq t, R_0^*(dt) = 1]$$

the regression parameters can be estimated by solving estimating equations (4), (5), and (11) (see the Web Appendix B). The estimation of the backward model, (3), can also be achieved via Model (1).

We develop asymptotic properties of the regression estimators with constant weight functions, which are shared by Models (1) and (2). Define

$\beta = (\beta_T', \beta_R', \beta_Y')'$, $\hat{\beta} = (\hat{\beta}_T', \hat{\beta}_R', \hat{\beta}_Y')'$, and the trivariate estimating function $U(\beta) = (U_T(\beta_T)', U_R(\beta_R', \beta_T)', U_Y(\beta_Y', \beta_T', \beta_R'))'$. Similar to Ying (1993) and Lin et al. (2000), we introduce the following conditions:

- (C1) β is restricted in a compact set.
- (C2) $\{Z_i, T_i, R_i^*(\cdot), Y_i(\cdot)\}$, $i = 1, \dots, n$ are i.i.d. with uniform bound.
- (C3) The densities of T and $dR^*(t)$ and their first order differential functions are bounded.
- (C4) The censoring time C_T and C_R have uniformly bounded densities, termed as g_T and g_R , respectively, example, there is B_c such that $|g_T(t)| < B_c$ and $|g_R(t)| < B_c$ for all t .
- (C5) $E|\min\{\ln T - \beta_T' Z, C_T, C_R\}|^\phi < \infty$ for some $\phi > 0$.

Theorem 1

Under conditions (C1-5), $n^{-1/2}U(\hat{\beta})$ converges weakly to a multivariate normal random variable with mean zero and variance denoted as Σ , and $U(\beta)$ is asymptotically linear in the sense that there exists a matrix A_n such that for every sequence $d_n > 0$ with $d_n \rightarrow 0$ in probability, we have

$$\sup_{\|b - \beta\| \leq d_n} \|U(b) - U(\beta) - A_n n(b - \beta)\| / (\sqrt{n} + n\|b - \beta\|) = o_p(1).$$

If the eigenvalues of A_n are all bounded away from zero for all large enough n and $A_n \rightarrow A$ where A is nonsingular, there exists a closed neighborhood $\mathcal{N}$ containing β as its interior point such that $\hat{\beta}$ is strongly consistent and $n^{1/2}(\hat{\beta} - \beta)$ converges to $N(0, A^{-1} \Sigma (A^{-1})')$ weakly in $\mathcal{N}$.

The proof of Theorem 1 is presented in the Web Appendix C. Of note, the matrix A involves the true distributions of recurrent events, longitudinal measurements and censoring with covariates, and is hard to estimate in practice. For data applications, we will use the Bootstrap approaches to estimate the confidence intervals.

4. Extension to Left Truncation

In biomedical studies, left-truncated sampling is commonly adopted, where only those subjects who have experienced the initiating event but have not experienced the failure event are recruited and followed until the occurrence of failure event or censoring; see Wang, Brookmeyer, and Jewell (1993), among others. The failure time of interest is still the time from initiating event to failure event in the target population, and this type of sampling results in left-truncated and right-censored failure time data along with accompanied recurrent events and longitudinal measurements. For example, in Alzheimer's Disease studies, the failure time T is age at onset of symptoms of Mild Cognitive Impairment (MCI), and the study recruited and followed only those subjects who were disease-free at baseline. The observed data sample then forms an example of left-truncated and right-censored data.

Our Model (1) and the subsequently developed inferential approach can be generalized to left-truncated and right-censored data under the conditional independent left truncation and right-censoring assumption, that is, $\{T, R^*(\cdot), Y(\cdot)\}$ is independent of (C_T, C_R, L) where L is the left truncation time, given covariates $\mathbf{Z}$. Suppose the observations $\{L_i, X_{Ti}, X_{Ri}, R_i(\cdot), Y_i(\cdot), \mathbf{Z}_i\}$, $i = 1, \dots, n$, are i.i.d. subject to $L_i < X_{Ti} \wedge X_{Ri}$. To estimate Model (1) under the left-truncated and right-censored sampling, we reconstruct the risk set indicators and the counting processes (Andersen et al., 2012). In particular We define

$\xi_{Ti}^L(t; \beta_T) = I_{[X_{Ti} \geq \exp(\beta'_T \mathbf{Z}_i) \geq L_i]}$, $N_{Ti}^L(t; \beta_T) = I_{[L_i \leq X_{Ti} \leq \exp(\beta'_T \mathbf{Z}_i)]} \Delta_i$,
 $N_{Ri}^L(t; \beta_T) = R_i^* \left(\exp(\beta'_T \mathbf{Z}_i) \wedge X_{Ri} \right) I_{[L_i \leq \exp(\beta'_T \mathbf{Z}_i)]}$. Estimation procedures similar to those in Section 3 can be achieved using the following three-layer estimation functions:

$$U_T^L(\beta_T) = \sum_{i=1}^n \int_0^\infty W_T(t; \beta_T) \left\{ \mathbf{Z}_i \frac{\sum_{j=1}^n \xi_{Tj}^L(t; \beta_T) \mathbf{Z}_j}{\sum_{j=1}^n \xi_{Tj}^L(t; \beta_T)} \right\} dN_{Tj}^L(t; \beta_T),$$

$$U_T^L(\beta_T; \beta_R) = \sum_{i=1}^n \int_0^\infty W_R(t; \beta_T; \beta_R) \left\{ \mathbf{Z}_i - \tilde{\mathbf{Z}}^L(t; \beta_T; \beta_R) \right\} dN_{Ri}^L(t; \beta_T),$$

$$U_T^L(\beta_Y; \beta_T; \beta_R) = \sum_{i=1}^n \int_0^\infty W_Y(t; \beta_T; \beta_R; \beta_Y) \{Z_i - \tilde{Z}^L(t; \beta_T; \beta_R)\} \left\{ Y_i \left(t e^{\beta'_T Z_i} \right) - \beta'_Y Z_i \right\} dN_{Ri}^L(t; \beta_T)$$

$$\tilde{Z}^L(t; \beta_T; \beta_R) = \frac{\sum_{j=1}^n Z_j \xi_{Rj}^L(t; \beta_T) e^{\beta'_R Z_j}}{\sum_{j=1}^n \xi_{Rj}^L(t; \beta_T) e^{\beta'_R Z_j}}.$$

where

5. Simulation Studies

We conducted simulation studies to assess the finite sample performance of the proposed methods. For each simulation study, we repeatedly generated 1000 simulated data sets. In each data set, the i th individual's sample is generated by the following procedure:

- Generate the covariate vector Z_i .
- Generate the potential baseline failure time T_{0i} independently with Z_i . Since $T_{0i} > 0$, we generate a random variable V_i independent with Z_i and define $T_{0i} = \exp(V_i)$. Thus, conditioning on T_{0i} is equivalent to conditioning on V_i .
- Define the failure time as $T_i = \exp(\beta'_T Z_i) T_{0i}$.
- Conditioning on $\{V_i, Z_i\}$, generate the sampling time process $R_{0i}^*(\cdot)$ on $[0, \tau]$ as a nonhomogeneous Poisson process with intensity

$$\lambda_i(t; Z_i, V_i) = \exp(-\beta'_R Z_i) \exp(-\beta'_R Z_i) h(t \exp(-\beta'_R Z_i), V_i),$$
 where $h(\cdot, \cdot)$ is a prespecified positive function. Of note, this intensity function satisfies the requirement of Model (2) because

$$E\{R_{0i}^*(t \exp(\beta'_T Z_i)) | T_i, Z_i\} = \exp(\beta'_T Z_i) \int_0^t h(u, V_i) du$$
 and

$$E\{R_{0i}^*(t) | T_{0i}, Z_i=0\} = \int_0^t h(u, V_i) du.$$
- Given V_i and Z_i at each sampling time t , generate the longitudinal variable by

$$Y_i(t) = \beta'_Y Z_i + a(t \exp(-\beta'_T Z_i)) + \varepsilon_i(t),$$
 where $\varepsilon_i(t)$ has normal distribution $N(\mu(V_i), \sigma^2(V_i))$. Here $a(\cdot)$, $\mu(\cdot)$, and $\sigma^2(\cdot)$ are prespecified functions.
- Generate censoring times $\{C_{Ti}, C_{Ri}\}$ which depend on the covariate Z_i but do not depend on V_i . Here, we set $C_{Ti} = C_{Ri} = C_i$.

In Table 1, all the specific distributions used to generate the simulated data are listed. The simulation procedure follows Model (2) exactly and is determined by the prespecified components: τ , distributions of $\{Z_i, T_{0i}, C_i\}$, functions $\{h(\cdot), \mu(\cdot), \sigma^2(\cdot), a(\cdot)\}$, and parameters $\{\beta_T, \beta_R, \beta_Y\}$. The simulation procedure is also valid for Model (1). We examined several settings, with continuous and discrete covariates, different covariate dimensions, various associations within $\{T, R^*(\cdot), Y(\cdot)\}$, and different sample sizes. For the estimating procedure, we set weighting functions as $W_T(t; \beta_T) = n^{-1} \sum_{i=1}^n \xi_{Ti}(t; \beta_T)$ which leads to the Gehan estimating function, and $W_R(\cdot) = W_Y(\cdot) = 1$.

The simulation results are summarized in Table 1. Scenario 0 simulated the special case of type (i) longitudinal measurements with the constant sampling rate overtime. Note that $R^*(\cdot)$ was not influenced by $\mathbf{Z}$ in this case, and we let β_R equal β_T for the reason mentioned in the Section 2.1. The sampling process considered in Scenario 1 followed a stationary Poisson process and was independent after conditioning on $\mathbf{Z}$. In Scenarios 2-3, T were correlated with both $R^*(\cdot)$ and $Y(\cdot)$. In Scenarios 1 and 2 involved a continuous and a discrete one-dimensional covariate separately. Scenario 3 mimicked a two-arm clinical trial with one covariate from Bernoulli distribution and another covariate from uniform distribution. We used the empirical bias, standard errors and 95% bootstrap confidence intervals for $(\beta_T, \beta_R, \beta_Y)$ to evaluate the performance of the estimation method. The bootstrap procedure was based on 1000 replications, where subjects were sampled with replacement and parameters $\{\beta_T, \beta_R, \beta_Y\}$ were estimated for each replication. As shown in Table 1, the estimating method performed well in all the situations considered here. In particular, the 95% bootstrap confidence interval of β_Y has empirical coverage probabilities ranging from 0.925 to 0.954 which accords with the asymptotic normality property.

Simulations studies on sensitivity analysis under the Scenario 2 illustrated that ignoring dependence of data with failure times could lead significantly biased results if the dependence existed (see the Web Appendix D).

6. Real Data Examples

6.1. BIOCARD Data Analysis

As an example of type (i) longitudinal measurements, we consider the application of our models on the data from the Biomarkers of Cognitive Decline Among Normal Individuals (BIOCARD) cohort study which aims at identifying biomarkers associated with the development of Alzheimer's Disease (AD) procession. The study was administrated by NIH from 1995 to 2005, and was re-established by a research group at Johns Hopkins School of Medicine after being stopped for four years. Subjects enrolled in the study were cognitively normal at baseline and data including cognitive performance testing scores were collected annually or per half year during the study.

The $\epsilon 4$ allele of the apolipoprotein E (ApoE) gene is the main genetic risk factor associated with AD dementia (Farrer et al., 1997). Our main object here is to estimate the effect of the ApoE4 gene on the time onset of clinical symptoms and the cognitive performance testing score *Logical Memory IIA - delayed*. Here, we only considered the data collected at the even follow-up years in 1995–2005. Totally, we had 236 subjects, consisting of 71 ApoE4 carriers and 165 ApoE4 non-carriers, and the overall censoring rate is 73.7%. Figure 1 shows the BIOCARD data of uncensored subjects in forward and backward time scales. We first analyzed the data as the right-censored case where the individual enrollment time was the time origin. Let T_i be the time from entry to the onset of symptoms for MCI, $R_i(\cdot)$ be the sampling process since entry and $Y_i(\cdot)$ be the cognitive score, $i = 1, \dots, 236$. This longitudinal measurement is selected because it is found to be a highly predictive marker for onset of symptoms. Since subjects were enrolled at different baseline ages, the centered baseline age of each individual was considered as covariate. By defining the failure time T_i as the age at onset of symptoms, the second set of data analysis is conducted by treating the

observed data as being left-truncated and right-censored, where the truncation time L_j is an individual's age at the time when she or he entered the study. In this analysis ApoE4 status was the only covariate.

The analysis results are reported in Table 2. For the right-censored data setting, the ApoE4 gene type has a time rescaling effect identified by the AFT model. The positive ApoE4 gene type accelerates the progress of disease and, as a result, the time to onset of symptoms is shortened by the rate $\exp(-0.209)$ for subjects with the same baseline age. Furthermore, the score in the ApoE4 carriers population at time $t \exp(-0.209)$ averagely equals to that in the ApoE4 noncarriers population at time t with an additional increase by a coefficient with value 0.560. Of note, β_Y represents the addition shift effect after being adjusted by the time-rescaling effect, and therefore these two effects should be considered together to understand the covariate effect. Neither the time rescaling effect nor the additional shift effect of ApoE4 gene type showed significant effects according to the 95% bootstrap confidence intervals with 1000 bootstrap repetitions. In contrast, the increase of baseline age significantly accelerates the disease process and reduces the score as shown in Table 2.

For the left-truncated and right-censored data setting, using a subject's age at onset of symptoms as the failure time, the positive ApoE4 gene type accelerates the progress to disease and shortens the age of onset of symptoms by rate $\exp(-0.044)$. Furthermore, the ApoE4 gene type influenced the rescaled recurrent event process with an additional multiplier term $\exp(-0.035)$. As to the effect of ApoE4 on the score, it provides an additional shift effect by 0.831, but nonsignificantly.

6.2. CPCRA Data Analysis

To illustrate the application of the proposed method on type (ii) longitudinal measurement, we analyzed data from a clinical trial conducted by Terry Bein Community Programs for Clinical Research on AIDS. The study compared two different treatments, didanosine (ddI) and zalcitabine (ddC), for HIV-infected patients who had previously failed treatment with or were intolerant to zidovudine (Abrams et al., 1994). In the trial, 230 patients were randomly selected to receive ddI treatment and 237 to receive ddC, the event of death is the primary endpoint, and patients were followed until death or censoring. During the course of the trial 363 opportunistic infections were observed, of which 172 were in the ddI group and 191 in the ddC group. For each opportunistic infection, a severity score was provided by physicians as an indicator for the disease progression (Neaton et al., 1994). The analysis in Abrams et al. (1994) suggested that ddC may have provided a survival advantage over ddI.

We investigated the effect of treatment where the indicator, Z , was the only covariate in the model (coded as 0 for ddI and 1 for ddC). Let T , $R^*(\cdot)$, and $Y(\cdot)$, respectively, be the time to death, the opportunistic infection process and the severity score.

The analysis results are summarized in Table 3, which shows that the treatment ddC had a significant time rescaling effect and elongated the survival time by the rate $\exp(0.179)$. The infection rate of the ddC treatment group at time $t \exp(0.179)$ is equal to that of the ddI treatment group at time t with an additional multiplier $\exp(0.156)$, which is not significant. Moreover, the mean severity score of the ddC treatment group at time $t \exp(0.179)$ is equal

to that of the ddI treatment group at time t with an additional increase 0.268 which is also not significant.

7. Discussion

In this article, we developed a joint model for longitudinal measurements, recurrent events, and failure events data where all of the three components of data are treated as outcomes. Without requiring restrict assumptions on recurrent event processes, the proposed model is applicable to both types of longitudinal measurements, (i) and (ii). On the basis of the survivors' population, the model avoids the disputing assumption on the existence of recurrent events or longitudinal measurements after the failure event. As the model does not involve latent variables, computationally it is simpler and easier to adopt when comparing to the shared frailty model. Moreover, the proposed model and estimation inference can be generalized to analyze left-truncated and right-censored data.

The proposed model involves semiparametric structure in each of the three sub-models, for failure time, recurrent events and longitudinal measurements, and the baseline functions in the model are unspecified. Since the main interest of this article is to model and estimate covariates effects, we will only briefly describe the approaches for estimating the baseline functions in the Appendix and show the simulation studies in the Web Appendix E.

Our model possesses a specific feature that the forward time model is equivalent to the backward-in-time model for recurrent events and longitudinal measurements, where the two models share the same regression parameter values. Therefore, our model can be used to study the terminal behavior of biological processes, such as the performance of a biomarker measurement before the diagnosis of disease or the medicine cost distributions before death. Of note, studying the terminal behavior of a longitudinal measurement process is challenging as the censoring cannot be handled using the standard approach (Chan and Wang, In Press). The proposed model possesses equivalence between forward and backward time scales, which is an attractive feature when modeling stochastic processes (such as recurrent events, longitudinal or functional measurements) in the presence of a terminal event. It has the obvious advantage that one can build a forward time model which is also valid in backward time scale.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Appendix: The Inference of Baseline Functions

For Models (1) and (2), with estimated regression parameters, we can estimate $\mathcal{A}_0(t)$ by (10), and estimate $\Lambda_{R_0}(t)$ by a Breslow-type estimator (Breslow, 1972)

$$\hat{\Lambda}_{R_0}(t; \hat{\beta}_T) = \sum_{i=1}^n \int_0^t \frac{dN_{Ri}(u; \hat{\beta}_T)}{\sum_{j=1}^n \xi_{Rj}(u; \hat{\beta}_T) e^{\hat{\beta}'_R \mathbf{Z}_j}}.$$

For Model (1), $\alpha_0(t) = E\{Y_0(t) | T_0 \geq t, R_0^*(dt)=1\}$ can be estimated by Kernel Smoothing method (Ramlau-Hansen, 1983)

$$\hat{\alpha}_0(t; \hat{\beta}_T) = \frac{\int_0^{+\infty} K_h(t-s) d\hat{\mathcal{A}}_0(s; \hat{\beta}_T)}{\int_0^{+\infty} K_h(t-s) d\hat{\Lambda}_{R_0}(s; \hat{\beta}_T)},$$

where $K_h(x) = h^{-1}K(x/h)$ is a kernel function with bandwidth h , $\int_{-1}^1 K(x) dx=1$, and $\int_{-1}^1 xK(x) dx=0$. Using the empirical process theory (Van Der Vaart and Wellner, 1996; Kosorok, 2008) and Kernel Smoothing techniques (Ramlau-Hansen, 1983) one can show that each of the processes $n^{1/2}\{\hat{\mathcal{A}}_0(t; \hat{\beta}_T) - \mathcal{A}_0(t)\}$, $n^{1/2}\{\hat{\Lambda}_{R_0}(t; \hat{\beta}_T) - \Lambda_{R_0}(t)\}$ and $(nh)^{1/2}\{\hat{\alpha}_0(t; \hat{\beta}_T) - \alpha_0(t)\}$ converges weakly to mean-zero Gaussian processes by weak convergence theory for functional parameters. Since these baseline functions are not the focus of our work, we skip details of the asymptotic proofs. The simulation studies are shown in the Web Appendix E.

References

- Abrams DI, Goldman AI, Launer C, Korvick JA, Neaton JD, Crane LR, et al. A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine*. 1994; 330:657–662. [PubMed: 7906384]
- Andersen, PK., Borgan, O., Gill, RD., Keiding, N. Statistical models based on counting processes. New York: Springer; 2012.
- Breslow NE. Contribution to the discussion of the paper by D.R. Cox. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*. 1972; 34:216–217.
- Buckley J, James I. Linear regression with censored data. *Biometrika*. 1979; 66:429–436.
- Chan KCG, Wang MC. Backward estimation of stochastic processes with failure events as time origins. *The Annals of Applied Statistics*. 2010; 4:1602–1620. [PubMed: 21359167]
- Chan IS, Neaton JD, Saravolatz LD, Crane LR, Osterberger J. Frequencies of opportunistic diseases prior to death among HIV-infected persons. *AIDS*. 1995; 9:1145–1152. [PubMed: 8519450]
- Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease: A meta-analysis. *JAMA*. 1997; 278:1349–1356. [PubMed: 9343467]
- Ghosh D, Lin DY. Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics*. 2003; 59:877–885. [PubMed: 14969466]
- Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Statistics in Medicine*. 2000; 19:1555–1566. [PubMed: 10844718]
- Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics*. 2000; 1:465–480. [PubMed: 12933568]
- Huang Y, Wang MC. Frequency of recurrent events at failure time: Modeling and inference. *Journal of the American Statistical Association*. 2003; 98:663–670.
- Huang CY, Wang MC. Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association*. 2004; 99:1153–1165. [PubMed: 24068850]
- Kalbfleisch, JD., Prentice, RL. The statistical analysis of failure time data. Hoboken: John Wiley and Sons; 2011.
- Kim S, Zeng D, Chambless L, Li Y. Joint models of longitudinal data and recurrent events with informative terminal event. *Statistics in Biosciences*. 2012; 4:262–281. [PubMed: 23227131]

- Kosorok, MR. Introduction to Empirical Processes and Semiparametric Inference. New York: Springer; 2008.
- Liang Y, Lu W, Ying Z. Joint modeling and analysis of longitudinal data with informative observation times. *Biometrics*. 2009; 65:377–384. [PubMed: 18759841]
- Lin H, Scharfstein DO, Rosenheck RA. Analysis of longitudinal data with irregular, outcome-dependent follow-up. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*. 2004; 66:791–813.
- Lin DY, Wei LJ, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*. 2000; 62:711–730.
- Lin DY, Ying Z. Semiparametric and nonparametric regression analysis of longitudinal data. *Journal of the American Statistical Association*. 2001; 96:103–126.
- Liu L, Huang X. Joint analysis of correlated repeated measures and recurrent events processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*. 2009; 58:65–81.
- Liu L, Huang X, O'Quigley J. Analysis of longitudinal data in the presence of informative observational times and a dependent terminal event, with application to medical cost data. *Biometrics*. 2008; 64:950–958. [PubMed: 18162110]
- Liu L, Wolfe RA, Huang X. Shared frailty models for recurrent events and a terminal event. *Biometrics*. 2004; 60:747–756. [PubMed: 15339298]
- Lunney JR, Lynn J, Foley DJ, Lipson S, Guralnik JM. Patterns of functional decline at the end of life. *Journal of the American Medical Association*. 2003; 289:2387–2392.
- Luo XH, Wang MC, Huang CY. A comparison of various rate functions of a recurrent event process in the presence of a terminal event. *Statistical Methods in Medical Research*. 2008; 19:167–182. [PubMed: 18445699]
- Neaton JD, Wentworth DN, Rhame F, Hogan C, Abrams DI, Deyton L. Considerations in choice of a clinical endpoint for AIDS clinical trials. *Statistics in Medicine*. 1994; 13:2107–2125. [PubMed: 7846414]
- Ramlau-Hansen H. Smoothing counting process intensities by means of kernel functions. *The Annals of Statistics*. 1983; 11:453–466.
- Ritov Y. Estimation in a linear regression model with censored data. *The Annals of Statistics*. 1990; 18:303–328.
- Song X, Davidian M, Tsiatis AA. A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics*. 2002; 58:742–753. [PubMed: 12495128]
- Song X, Wang CY. Semiparametric approaches for joint modeling of longitudinal and survival data with time varying coefficients. *Biometrics*. 2008; 64:557–566. [PubMed: 17725812]
- Sun J, Park DH, Sun L, Zhao X. Semiparametric regression analysis of longitudinal data with informative observation times. *Journal of the American Statistical Association*. 2005; 100:882–889.
- Sun J, Sun L, Liu D. Regression analysis of longitudinal data in the presence of informative observation and censoring times. *Journal of the American Statistical Association*. 2007; 102:1397–1406.
- Sun L, Song X, Zhou J, Liu L. Joint analysis of longitudinal data with informative observation times and a dependent terminal event. *Journal of the American Statistical Association*. 2012; 107:688–700.
- Tsiatis AA. Estimating regression parameters using linear rank tests for censored data. *The Annals of Statistics*. 1990; 18:354–372.
- Usvyat LA, Barth C, Bayh I, Etter M, von Gersdorff GD, Grassmann A, et al. Interdialytic weight gain, systolic blood pressure, serum albumin, and C-reactive protein levels change in chronic dialysis patients prior to death. *Kidney International*. 2013; 84:149–157. [PubMed: 23515055]
- Van Der Vaart, AW., Wellner, JA. Weak Convergence and Empirical Processes: With Applications to Statistics. New York: Springer; 1996.
- Vonesh EF, Greene T, Schluchter MD. Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine*. 2006; 25:143–163. [PubMed: 16025541]

- Wang MC, Brookmeyer R, Jewell NP. Statistical models for prevalent cohort data. *Biometrics*. 1993; 49:1–11. [PubMed: 8513095]
- Wang MC, Qin J, Chiang CT. Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association*. 2001; 96:1057–1065.
- Wilson RS, Beck TL, Bienias JL, Bennett DA. Terminal cognitive decline: Accelerated loss of cognition in the last years of life. *Psychosomatic Medicine*. 2007; 69:131–137. [PubMed: 17327212]
- Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997; 53:330–339. [PubMed: 9147598]
- Xu J, Zeger SL. Joint analysis of longitudinal data comprising repeated measures and times to events. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*. 2001; 50:375–387.
- Ye Y, Kalbfleisch JD, Schaubel DE. Semiparametric analysis of correlated recurrent and terminal events. *Biometrics*. 2007; 63:78–87. [PubMed: 17447932]
- Ying Z. A large sample study of rank estimation for censored regression data. *The Annals of Statistics*. 1993; 21:76–99.
- Zeng D, Cai J. A semiparametric additive rate model for recurrent events with an informative terminal event. *Biometrika*. 2010; 97:699–712. [PubMed: 23956458]

Figure 1.
Illustration of BIOCARD data for uncensored subjects in forward and backward time indices.

Table 1

Summary Statistics of the Simulation Studies

<i>n</i>	β_r			β_R			β_Y		
	<i>Bias</i>	<i>SE</i>	<i>CP</i>	<i>Bias</i>	<i>SE</i>	<i>CP</i>	<i>Bias</i>	<i>SE</i>	<i>CP</i>
Scenario 0:									
	$Z \sim B\left(1, \frac{1}{2}\right) V \sim N\left(0, \frac{1}{4}\right), \tau = 7, h(t, V) = 0.015, a(t) = t, \mu(V) = V, \sigma^2(V) = V^2, \log C \sim N[Z, (Z + 1)^2]$								
	$\beta_T = -0.5$								
<i>n</i> = 100	-0.004	0.121	0.944	-0.008	0.252	0.939	0.013	0.442	0.940
<i>n</i> = 200	-0.000	0.083	0.962	-0.010	0.171	0.944	0.027	0.317	0.952
	$\beta_T = 1$								
	$\beta_R = 1$								
<i>n</i> = 100	0.006	0.130	0.949	0.026	0.211	0.942	-0.001	0.367	0.941
<i>n</i> = 200	0.003	0.091	0.942	0.009	0.146	0.937	0.004	0.262	0.944
Scenario 1: $Z \sim U[0, 1], V \sim \text{Exp}(1), \tau = 7, h(t, V) = 0.01, a(t) = \sin(t), \mu(V) = V, \sigma^2(V) = V^2, C \sim U[0, 12] \text{ if } Z < 0.5 \text{ and } C \sim U[0, 15] \text{ if } Z \geq 0.5$									
	$\beta_T = 1$								
	$\beta_R = 0$								
<i>n</i> = 100	0.002	0.220	0.944	0.028	0.351	0.937	0.028	0.680	0.937
<i>n</i> = 200	0.002	0.144	0.962	0.010	0.237	0.946	0.008	0.472	0.941
	$\beta_T = 0.5$								
	$\beta_R = -0.5$								
<i>n</i> = 100	0.003	0.226	0.941	0.005	0.352	0.941	-0.009	0.715	0.925
<i>n</i> = 200	-0.002	0.147	0.949	-0.005	0.253	0.943	-0.002	0.502	0.935
Scenario 2:									
	$Z \sim B\left(1, \frac{1}{2}\right) V \sim N\left(0, \frac{1}{4}\right), \tau = 7, h(t, V) = \sqrt{3t V /50}, a(t) = t, \mu(V) = \sqrt{ V }, \sigma^2(V) = V , C \sim U[0, 12] \text{ if } Z = 1 \text{ and } C \sim U[0, 15] \text{ if } Z = 0$								
	$\beta_T = 1$								
	$\beta_R = -1$								
<i>n</i> = 100	0.002	0.111	0.937	-0.000	0.335	0.941	-0.021	0.470	0.938
<i>n</i> = 200	-0.006	0.077	0.943	-0.008	0.224	0.951	-0.028	0.334	0.952
	$\beta_T = -0.5$								
	$\beta_R = 0$								
<i>n</i> = 100	-0.003	0.107	0.944	-0.018	0.271	0.957	0.005	0.464	0.942
<i>n</i> = 200	0.003	0.072	0.959	-0.007	0.189	0.955	0.043	0.350	0.942

<i>n</i>	β_r				β_R				β_Y			
	<i>Bias</i>	<i>SE</i>	<i>CP</i>	<i>Bias</i>	<i>SE</i>	<i>CP</i>	<i>Bias</i>	<i>SE</i>	<i>Bias</i>	<i>SE</i>	<i>CP</i>	<i>CP</i>
Scenario 3: $Z_1 \sim B\left(1, \frac{1}{2}\right), Z_2 \sim U[0, 1], V \sim N\left(0, \frac{1}{4}\right), h(t, V) = \frac{ tV }{20}, \mathcal{A}(t) = t, \mu(V) = V, \sigma^2(V) = V^2, C \sim U[0, 12] \text{ if } Z_1 = 0 \text{ and } C \sim U[0, 15] \text{ if } Z_1 = 1$												
<i>n</i> = 100	$\beta_T = (1, 0.5)'$				$\beta_R = (-0.8, -0.5)'$				$\beta_Y = (2, 1)'$			
	0.002	0.111	0.938	−0.005	0.467	0.940	−0.032	0.446	0.936			
	−0.007	0.193	0.948	0.014	0.744	0.957	−0.036	0.757	0.950			
	0.005	0.076	0.952	0.030	0.320	0.941	−0.021	0.331	0.928			
<i>n</i> = 200	$\beta_T = (-0.3, 0)'$				$\beta_R = (-1, -0.1)'$				$\beta_Y = (-3, 0)'$			
	0.005	0.134	0.953	0.028	0.523	0.958	−0.020	0.550	0.942			
	−0.002	0.108	0.948	−0.014	0.416	0.950	−0.037	0.444	0.937			
	0.009	0.181	0.959	0.028	0.709	0.949	0.004	0.711	0.954			
<i>n</i> = 200	0.001	0.071	0.962	0.006	0.272	0.966	−0.016	0.305	0.942			
	0.008	0.131	0.949	0.017	0.493	0.945	−0.005	0.517	0.946			

Note: *Bias*, the empirical bias; *SE*, the empirical standard error; *CP*, the empirical coverage probability of 95% bootstrap confidence interval; $B\left(1, \frac{1}{2}\right)$, the Bernoulli distribution; $N(\mu, \sigma^2)$, the normal distribution; $U[1, 2]$, the Uniform distribution; $Exp(\lambda)$, the Exponential distribution.

Table 2

Summary of BIOCARD Data Analysis

		Coefficient	95% CI
Case 1	Right-censoring		
β_T	ApoE4+	-0.209	(-0.558, 0.167)
	Baseline age	-0.040	(-0.065, -0.022)
β_R	ApoE4+	-0.121	(-0.543, 0.321)
	Baseline age	-0.029	(-0.065, -0.005)
β_Y	ApoE4+	0.560	(-2.001, 3.166)
	Baseline age	0.161	(-0.028, 0.547)
Case 2	Left truncation and right-censoring		
β_T	ApoE4+	-0.044	(-0.125, 0.029)
β_R	ApoE4+	-0.035	(-0.250, 0.236)
β_Y	ApoE4+	0.831	(-2.998, 2.515)

Note: 95% CI, bootstrap percentile 95% confidence interval.

Table 3

Summary of CPCRA Data Analysis

	Coefficient	95% CI
β_T	0.179	(0.012, 0.328)
γ	0.156	(−0.025, 0.407)
β_Y	0.268	(−0.733, 0.831)

Note: 95% CI, bootstrap percentile 95% confidence interval.