4-13-2006

RECURRENT EVENT MODELS IN THE PRESENCE OF A TERMINAL EVENT: COMPARISON, INFERENCE AND DATA ANALYSIS

Xianghua Luo
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, xianghua@biostat.umn.edu

Mei-Cheng Wang
Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, mcwang@jhsph.edu

Recurrent Event Models in the Presence of a Terminal Event: Comparison, Inference and Data Analysis

Xianghua Luo *

Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, U.S.A.

and

Mei-Cheng Wang

SUMMARY. This article focuses on statistical implications of proportional rate models for recurrent event data in the presence of a terminal event. In such circumstances, various definitions of the recurrent rate function have been adopted in the proportional rate models. Although these rate functions have quite different interpretations, recognition of the differences has been lacking theoretically and practically. We compare three types of rate functions from both conceptual and quantitative perspectives; conclude that the inappropriate choice of a rate function may lead to misleading scientific conclusions. Simulations are conducted for comparisons of the focused models. Analysis of data from an AIDS clinical trial is presented to illustrate the analytical results.

KEY WORDS: Proportional rate model, Randomized clinical trial, Rate func-

* *email:* xluo@jhsph.edu

tion.

1. Introduction

In many longitudinal follow-up studies, recurrent events are recorded as outcome measurements where the occurrence of recurrent events could be stopped by a terminal event, such as death. For instance, recurrences of hospitalizations could be terminated by the death of a patient; repeated opportunistic infections could be terminated by death. The terminal event is usually not independent of the recurrent events. Hence, treating the terminal event as a part of an independent censoring mechanism is not generally appropriate. If the terminal event is of interest besides the recurrent events, a joint model for the recurrent and terminal events could be considered, otherwise the terminal event can be treated as a nuisance and explicit modelling of the terminal event can be avoided.

In the absence of a terminal event, the regression methods for recurrent events were studied in the important articles of Prentice, Williams and Peterson (1981), and Anderson and Gill (1982). Both articles studied intensity models which serve as proper predictive models in applications. However, to identify treatment effects or risk factors, marginal rate models would be more preferable by practitioners. Marginal rate models were considered by Pepe and Cai (1993), and Lin, Wei, Yang and Ying (2000), and independent censoring was still assumed as that in the intensity models. From the viewpoint of applications, the independent censoring assumption is reasonable in intensity models since censoring could be largely explained by event history and covariates, but more restrictive in the marginal rate models. With the intention to deal with dependent censoring in a marginal model, Wang,

Qin and Chiang (2001) proposed a nonparametric one-sample method and a semiparametric regression model for the recurrent events.

In the presence of a terminal event, one-sample methods, proposed by Cook and Lawless (1997), Ghosh and Lin (2000), and Wang, Qin and Chiang (2001), can be used to estimate the recurrent rate functions. Abundant regression approaches fall into several categories according to the inference of interest: Cook and Lawless (1997), Ghosh and Lin (2002), Huang and Wang (2004), Liu, Wolfe and Huang (2004), and Schaubel and Cai (2005) assumed proportional rate models for recurrent events; Ghosh and Lin (2003) proposed an accelerated failure time model for recurrent events, where the effect of covariates was directly on recurrent event times; Huang and Wang (2003) proposed joint models of the recurrent and terminal events, where the inference was focused on the frequency of recurrent events at the failure time of the terminal event. In the literature of the proportional rate models, we found that different definitions of rate function have been adopted for modelling recurrent event processes. However, implications of these rate models can be quite different or even conflicting with each other. In this article we present careful comparison of these proportional rate models and point out that inappropriate choice of a rate function may lead to misleading scientific results. We further provide guidance for choosing appropriate statistical models and interpreting analytical results. Our focus is on the rate functions and their subtle interrelationships throughout.

The rest of the article is organized as follows. In Section 2, the proportional rate models based on three different rate functions are studied in both conceptual and practical perspectives. Quantitative comparisons of the

focused models are presented in Section 3. A simulation study and the analysis of a data from an AIDS clinical trial are reported in Section 4. Some concluding remarks follow in Section 5.

2. Rate Functions and Proportional Rate Models

The proportional rate models are commonly used for analyzing recurrent event data in the presence or absence of an explicit terminal event. When a terminal event is present, various definitions of the rate function for a recurrent event process have been adopted in the proportional rate models. It is important to identify the distinctions of these rate functions, on which the interpretation of regression parameters is based. We will devote this section to introducing three distinct rate functions and their corresponding regression models.

2.1 *The Rate Function*

Let $N(t)$ denote the number of recurrent events occurring at or before time t, $t \geq 0$, and D, the time to the terminal event, where $N(t)$ and D are possibly correlated. The rate function, which is the occurrence rate of recurrent events over a time interval, is defined as

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(N(t+\Delta t) - N(t) > 0)}{\Delta t}, \quad t \in [0, \tau]. \tag{1}$$

Assuming that $\lim_{\Delta t \to 0^+} \Pr(N(t+\Delta t) - N(t) > 1) = 0$, we have $\lambda(t)dt = E[dN(t)]$. Further we define the cumulative rate function as $\Lambda(t) = \int_0^t \lambda(u)du$, which is equal to $E[N(t)]$. A non-parametric method for estimating the cumulative rate function was proposed by Wang, Qin and Chiang (2001).

This rate function is of great interest in randomized clinical trials because

of the fact that the population of this rate function is determined at time origin 0 for any $t \geq 0$. Since randomization is typically performed at time 0, the populations of the rate functions for different treatments are always comparable to each other at any $t \geq 0$. In the presence of a terminal event, e.g. death, we notice that the recurrent event process and the rate function after death are mathematically well defined, but latent in reality due to the fact that the recurrent event process after death does not exist. The rate function has a marginal interpretation, which is desirable for studying population-average treatment effects. However, by directly modelling the marginal rate function, it would be difficult to characterize the correlation between the recurrent event process and the terminal event. With the intention to deal with the dependence between the terminal and the recurrent events, Huang and Wang (2004) proposed a joint model. In their approach, conditioning on a latent variable Z, the recurrent and terminal events were modelled separately as two subject-specific models, with the correlation characterized by Z. The latent variable Z can be considered as, say, the latent health status of a patient. A brief review of the three assumptions (A1), (A2), and (A3) for this model is given as follows:

(**A1**) Let $\boldsymbol{X}$ be a $1 \times p$ vector of covariates and $\boldsymbol{x}$ be the realization of $\boldsymbol{X}$. There exists a nonnegative-valued latent variable Z so that, given $\boldsymbol{X} = \boldsymbol{x}$ and $Z = z$, the recurrent event process $N(\cdot)$ is a nonstationary Poisson process with intensity function

$$\lambda(t|\boldsymbol{x}, z) = z\lambda_0(t)\exp(\boldsymbol{x}\boldsymbol{\beta}), \quad 0 \leq t \leq \tau, \tag{2}$$

where $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters and the baseline intensity

function $\lambda_0(t)$ is a continuous function with the constraint $\Lambda_0(\tau) = \int_0^\tau \lambda_0(u)du = 1$. Research interest is within a fixed time interval $[0, \tau]$. The latent variable Z satisfies $E[Z|\boldsymbol{X}] = E[Z]$.

(**A2**) Given $(\boldsymbol{x}, z)$, the hazard function of D has the form

$$h(t|\boldsymbol{x}, z) = zh_0(t)\exp(\boldsymbol{x}\boldsymbol{\alpha}), \tag{3}$$

where $\boldsymbol{\alpha}$ is a $p\times 1$ vector of parameters and the baseline hazard function $h_0(t)$ is continuous.

(**A3**) Conditioning on $(\boldsymbol{x}, z)$, $(N(\cdot), D)$ are mutually independent.

Conditioning on z, the intensity function is also the rate function since a Poisson process is memoryless. When integrating out Z, one has the marginal rate function at time t for subjects with covariate $\boldsymbol{x}$,

$$\lambda(t|\boldsymbol{x}) = \mu_Z\lambda_0(t)\exp(\boldsymbol{x}\boldsymbol{\beta}), \tag{4}$$

where $\mu_Z = E[Z]$.

Assumption (A3) applies to $(N(t), D)$ for t before and after the time of the terminal event, D. Here, the model structure of the latent part, $\{(N(t), D) : t > D\}$, should be considered only as a "working assumption", which is assumed for ease of establishing statistical properties. In reality, to use model (A1-3), it is necessary to validate the model assumptions only for those $t \leq D$. Through the unobserved z and observed covariate $\boldsymbol{x}$, the correlation between a patient's recurrent event process and terminal event is explained. In other words, after controlling z and $\boldsymbol{x}$ on a patient, the recurrent event process and the terminal event are operating independently.

In this model, both the recurrent event model (in A1) and the terminal event model (in A2) have subject-specific interpretation for the regression coefficients. Also, the parameter β can be interpreted marginally, given the relationship in (4). Suppose X is a dichotomous treatment indicator, i.e. $X = 1$ indicates the treatment and $X = 0$ the control group, the interpretation of e^{β} would be the ratio of the occurrence rate of recurrent events for the treatment group to that for the control group at a patient's level. The interpretation of the regression coefficient for the recurrent events, β, is not affected by the terminal event. In contrast, we will see that the other two rate functions involve both the recurrent event process and the terminal event. Model (A1-3) will be used as our *core model*. Its relationship with models based on the other two rate functions will be discussed in Section 3.

2.2 *The Adjusted Rate Function*

The second rate function, termed as the adjusted rate function, is based on the observed counting process $N^*(t)$:

$$N^*(t) = \begin{cases} N(t) & \text{if } t < D, \\ N(D) & \text{if } t \geq D. \end{cases}$$

Equivalently, $N^*(t) = \int_0^t I(D \geq u) dN(u)$, where $I(\cdot)$ is an indicator function. The adjusted rate function, $\lambda^A(t)$, is hence defined as

$$\lambda^A(t)dt = E[dN^*(t)] = E\left[I(D \geq t)dN(t)\right]. \tag{5}$$

The corresponding cumulative rate function is $\Lambda^A(t) = E[N^*(t)]$. It follows from $N^*(t) \leq N(t)$ that $\lambda^A(t) \leq \lambda(t)$, $t \in [0, \tau]$. In some studies, the treatment effects are present in both the recurrent and terminal events under the core model (A1-3). However, the two effects can potentially result in the adjusted rate function in the opposite direction from a given rate function. For

example, patients in a later stage of AIDS suffer more frequent opportunistic infections than patients in an earlier stage, which corresponds to an increasing rate function in (1). However, after adjusting for the presence of deaths, we could observe a decreasing adjusted rate function. This might lead to a wrong impression that AIDS patients experience less frequent infections later than earlier. It can be seen from this example that the adjusted rate function is not appropriate for describing natural history of opportunistic infections. The adjusted rate function could, say, be useful for insurance companies to study the medical cost associated repeated hospitalizations, where the adjustment is needed because medical cost vanishes after one's death. From the standpoint of insurance companies, expense rather than patients' disease progression is the focus of interest.

The proportional rate model based on the adjusted rate function is

$$\lambda^A(t|\boldsymbol{x}) = \lambda_0^A(t)\exp(\boldsymbol{x}\boldsymbol{\beta}^A), \tag{6}$$

where $\lambda_0^A(t)$ is the baseline function and $\boldsymbol{\beta}^A$ is a $p \times 1$ vector of regression coefficients. See Ghosh and Lin (2002) for related work. Modeling the adjusted rate function based on the observed counting process is convenient since no assumption is needed for the recurrent event process after the terminal event. While the use of this model is relevant and interesting for some studies, in a clinical trial setting, the interpretation of the regression coefficients could be misleading because the adjusted rate function involves both recurrent and terminal events in a complicated manner. Consequently, a negative value of β^A in (6) does not necessarily imply a truly beneficial effect on recurrent events, in the same way, a positive value of β^A does not necessarily imply a harmful effect.

[Figure 1 about here.]

It is not difficult to find a situation in which the adjusted rate ratio does not retain the rate ratio. Suppose a new drug prolongs one's lifetime [Figure 1(a)] while it makes no difference in the occurrences of recurrent diseases, as compared with an existing drug [Figure 1(b)]. In this scenario, it is likely to observe more disease occurrences in patients receiving the new drug because they live longer than patients receiving the existing drug. This explains why the adjusted rate function is larger in the new drug group than that in the existing drug group [Figure 1(c)]. In this case, the new, life-saving drug is seen to be accompanied with higher frequencies of disease occurrences by the misleading result from comparing the adjusted rate functions.

Based on the previous discussion, it seems that, applying the adjusted rate model in (6) on clinical trials to study the treatment effect on the occurrences of recurrent events, is not generally appropriate. More quantitative discussion will be made in Section 3 to demonstrate the different situations where we can and cannot use the adjusted rate model.

2.3 *The Survivors' Rate Function*

Making inference on recurrent event processes for patients who are currently alive is of interest in many studies (Schaubel and Cai, 2005; Liu, Wolfe and Huang, 2004). The third rate function, $\lambda^S(t)$, termed as the survivors' rate function, is the recurrence rate among survivors:

$$\lambda^S(t)dt = E[dN(t)|D \geq t]. \tag{7}$$

The corresponding cumulative rate function among survivors is $\Lambda^S(t) = \int_0^t \lambda^S(u)du$. When the recurrent and terminal event processes are indepen-

dent, we have $\lambda^S(t) = \lambda(t)$. We note that $\lambda^S(t) = \lambda^A(t)/\Pr(D \geq t)$, which implies $\lambda^S(t) \geq \lambda^A(t)$. It should be noted that the population for defining the survivors' rate function changes with time. Explicitly, the defining population could only be randomized at baseline, i.e., $t = 0$. After time 0, this rate function is defined for unrandomly selected subpopulation, the survivors. Therefore, the survivors' rate function should be used with caution in randomized trials when the terminal event is not independent of the recurrent events.

The proportional rate model based on the survivors' rate function takes the form,

$$\lambda^S(t|\boldsymbol{x}) = \lambda_0^S(t) \exp(\boldsymbol{x}\boldsymbol{\beta}^S), \tag{8}$$

where $\lambda_0^S(t)$ is the baseline function and $\boldsymbol{\beta}^S$ is a $p \times 1$ vector of regression coefficients. Similar to the adjusted rate model, the survivors' rate model technically circumvents modeling recurrent event process after death. However, an informative terminal event can potentially confound the interpretation of $\boldsymbol{\beta}^S$. This is because, when the terminal event is informative for the recurrent event process, $\boldsymbol{\beta}^S$ carries information of the treatment effects from both the recurrent and terminal events. In contrast, from the standpoint of clinical investigation, the core model singles out the treatment effect on the recurrent diseases besides the treatment effect on survival. More details of model comparisons will be discussed in the next section.

A summary of the three focused rate functions and the corresponding proportional rate models is listed in Table 1. It should be noted that the statistical inferences of the three discussed models are all valid under their

own model assumptions. The focus of this paper is on the differences of their interpretations or implications.

[Table 1 about here.]

3. Quantitative Comparison of Models

We have seen that both the adjusted rate model and the survivors' rate model have their own limitations in a clinical trial setting, with or without the core model assumptions (A1-3). However, for ease of discussion, we use the core model as a frame of reference to investigate implications of the other two proportional rate models under the core model's assumptions. We will assume that X is a treatment indicator and the only covariate throughout this section.

3.1 *The Adjusted Rate Model under the Core Model Assumptions*

We first investigate the adjusted rate function, with the notion that the core model assumptions are correct. We have

$$\begin{aligned}
\lambda^A(t|x)dt &= E[I(D \geq t)dN(t)|x] \\
&= E[E[I(D \geq t)dN(t)|Z, X]|x] \\
&= E[\Pr(D \geq t|Z, X) \cdot E[dN(t)|Z, X]|x] \\
&= E\left[\exp\left\{-\int_0^t Zh_0(u)\exp(x\alpha)du\right\} \cdot Z\lambda_0(t)\exp(x\beta)dt \,\middle|\, x\right] \\
&= \lambda_0(t)\exp(x\beta)E[Z\exp\{-Z\exp(x\alpha)H_0(t)\}|x]dt, \qquad (9)
\end{aligned}$$

where $H_0(t) = \int_0^t h_0(u)du$ is the cumulative baseline hazard function for time to the terminal event in (A2). In general, the adjusted rates for different $\boldsymbol{x}$

values are not proportional to each other, i.e., rate ratios depend on t. The adjusted rate ratio for the treatment compared with the control group is

$$\frac{\lambda^A(t|x=1)}{\lambda^A(t|x=0)} = \exp\left[\beta + \log\left\{\frac{E[Z\exp\{-Ze^{\alpha}H_0(t)\}]}{E[Z\exp\{-ZH_0(t)\}]}\right\}\right], \tag{10}$$

which is not functionally independent of t. The adjusted rate functions for the treatment and control groups have a crossover at t whenever t satisfies

$$e^{\beta} = \frac{E[Z\exp(-ZH_0(t))]}{E[Z\exp(-Ze^{\alpha}H_0(t))]}.$$

The adjusted rate ratio in (10) is determined by four elements: β, α, $H_0(t)$, and the distribution of Z. It is obvious that $\lambda_0(t)$ does not affect the ratio whatsoever, though it does affect the absolute magnitude of the rate.

We next examine how the relationship of α and β affects the adjusted rate ratio leaving the other elements unchanged and assuming that $\Pr(Z > 0) > 0$ and $H_0(t) > 0$.

Scenario I: $\beta = 0$ **and** $\alpha \neq 0$

In this situation, the treatment does not affect the recurrent events, but it affects the terminal event. If $\alpha < 0$, $\lambda^A(t|x=1) > \lambda^A(t|x=0)$; if $\alpha > 0$, $\lambda^A(t|x=1) < \lambda^A(t|x=0)$. In either case, by comparing the adjusted rates, we could "observe" an effect of X on the recurrent events which actually does not exist. As we have seen, different directions of the effect on the terminal event yield false effects in different directions on the recurrent events. Modeling the adjusted rate is not appropriate for examining the treatment effect on the recurrent events in this scenario.

Scenario II: $\alpha\beta > 0$

This means that the treatment effects on the recurrent and terminal events have the same direction. For example, when the recurrent and terminal events are both adverse events, $\alpha\beta > 0$ means the treatment is either beneficial ($\alpha < 0, \beta < 0$) or harmful ($\alpha > 0, \beta > 0$) for both events. In such circumstances, β and $\log\{E[Z\exp\{-Ze^{\alpha}H_0(t)\}]/E[Z\exp\{-ZH_0(t)\}]\}$ have opposite signs, which means that the existence of treatment effect on the terminal event ($\alpha \neq 0$) attenuates the adjusted rate ratio from the underlying rate ratio. The attenuation can lead to adjusted rate ratio in an opposite direction from the rate ratio. In addition, considering the nonproportionality and possible crossovers between the adjusted rate functions, the adjusted rate model is not applicable in this scenario.

Scenario III: $\alpha\beta < 0$

Consider the situation that the treatment effects on the recurrent and terminal events are in opposite directions. For example, a drug can decrease the frequency of repeated opportunistic infections in AIDS patients, but increase the risk of death at the same time. We notice that β and $\log\{E[Z\exp\{-Ze^{\alpha}H_0(t)\}]/E[Z\exp\{-ZH_0(t)\}]\}$ have the same sign here, which means no crossover could occur between $\lambda^A(t|x=0)$ and $\lambda^A(t|x=1)$. However, α inflates the adjusted rate ratio from the underlying rate ratio. Hence, the adjusted rate model does not apply in this scenario.

Scenario IV: $\beta \neq 0$ **and** $\alpha = 0$

In this case, the adjusted rate degenerates to $\lambda_0(t)E[Z\exp\{-ZH_0(t)\}]\exp(x\beta)$. Define the product of the first two terms as $\lambda_0^A(t) = \lambda_0(t)E[Z\exp\{-ZH_0(t)\}]$,

which becomes the new baseline rate function, then the adjusted rate function has the form $\lambda^A(t|x) = \lambda_0^A(t)\exp(x\beta)$. We can see that e^β could either be interpreted as the rate ratio coming from the core model, or as the adjusted rate ratio. This means that the proportional rate model can be applied to correctly identify the treatment effect on the recurrent events even though the absolute magnitude of $\lambda^A(t|x)$ is biased from $\lambda(t|x)$. The bias term is $\mu_Z^{-1}E[Z\exp\{-ZH_0(t)\}]$, since $\lambda^A(t|x) = \lambda(t|x)\mu_Z^{-1}E[Z\exp\{-ZH_0(t)\}]$, where $\lambda(t|x) = \mu_Z\lambda_0(t)\exp(x\beta)$.

Scenario V: $N(t)$ and D are independent for given x

Now we consider the situation where the recurrent and terminal event processes are conditionally independent within each treatment group. Under the core model's assumptions, the adjusted rate for given x is $\lambda_0(t)\exp(x\beta)\mu_Z$ $\exp\{-\mu_Z\exp(x\alpha)H_0(t)\}$ and the adjusted rate ratio, $\lambda^A(t|x=1)/\lambda^A(t|x=0) = \exp\{\beta - \mu_Z H_0(t)(e^\alpha - 1)\}$. We find that under the conditional independence condition, the adjusted rate ratio could be attenuated or inflated from the underlying rate ratio when different combinations of α and β present, as we discussed in Scenarios I-IV. After adjusting for death, the adjusted functions distort the underlying rate functions and may be misleading when it is used to examine the treatment effect on the recurrent events, no matter whether the recurrent and terminal event processes are independent or not.

3.2 *The Survivors' Rate Model under the Core Model Assumptions*

Assuming that the core model in (A1-3) is correct, the survivors' rate function is:

$$
\begin{aligned}
\lambda^S(t|x)dt &= E[dN(t)|D \geq t, x] \\
&= E[E[dN(t)|D \geq t, Z, X]|D \geq t, x] \\
&= E[E[dN(t)|Z, X]|D \geq t, x] \\
&= E[Z\lambda_0(t)\exp(x\beta)dt|D \geq t, x] \\
&= \lambda_0(t)\exp(x\beta)E[Z|D \geq t, x]dt, \qquad (11)
\end{aligned}
$$

where

$$
\begin{aligned}
E[Z|D \geq t, x] &= \frac{E[Z\Pr(D \geq t|Z, X)|x]}{\Pr(D \geq t|x)} \\
&= \frac{E[Z\exp\{-Z\exp(x\alpha)H_0(t)\}|x]}{E[\exp\{-Z\exp(x\alpha)H_0(t)\}|x]}. \qquad (12)
\end{aligned}
$$

The third equality in (11) follows from Assumption (A3). We prove in the Appendix that $E[Z|D \geq t, x = 1] \geq E[Z|D \geq t, x = 0]$ when $\alpha < 0$; $E[Z|D \geq t, x = 1] \leq E[Z|D \geq t, x = 0]$ when $\alpha > 0$. The survivors' rate ratio for the treatment group to the control group is

$$
\frac{\lambda^S(t|x = 1)}{\lambda^S(t|x = 0)} = \exp\left[\beta + \log\left\{\frac{E[Z|D \geq t, x = 1]}{E[Z|D \geq t, x = 0]}\right\}\right],
$$

which is not functionally independent of t. The survivors' rate functions for the two treatment groups have a crossover at t whenever t satisfies

$$
e^\beta = \frac{E[Z|D \geq t, x = 0]}{E[Z|D \geq t, x = 1]}.
$$

We now examine how different combinations of α and β (in Scenarios I-IV) affect the survivors' rate ratio when the recurrent and terminal event processes are possibly correlated within each treatment group.

Scenario I: $\beta = 0$ **and** $\alpha \neq 0$

If $\alpha < 0$ and $\beta = 0$, $\lambda^S(t|x = 1) > \lambda^S(t|x = 0)$; and if $\alpha > 0$ and $\beta = 0$, $\lambda^S(t|x = 1) < \lambda^S(t|x = 0)$. Through the comparison of the survivors' rate functions, a false treatment effect on the recurrent events would be observed, it, however, does not actually exist.

Scenario II: $\alpha\beta > 0$

When $\alpha < 0$ and $\beta < 0$, we have proved that $E[Z|D \geq t, x = 1] \geq E[Z|D \geq t, x = 0]$. This is intuitive, because the treatment prolongs patients' lifetime so that the survivors at time t in the treatment group are more fragile than the survivors in the control group. By "more fragile" we mean a larger expected value of the latent variable Z. Hence, $\log\{E[Z|D \geq t, x = 1]/E[Z|D \geq t, x = 0]\} \geq 0$. Similarly, when $\alpha > 0$ and $\beta > 0$, $\log\{E[Z|D \geq t, x = 1]/E[Z|D \geq t, x = 0]\} \leq 0$. In either case, β and $\log\{E[Z|D \geq t, x = 1]/E[Z|D \geq t, x = 0]\}$ have opposite signs, which means that the survivors' rate ratio is attenuated from the underlying rate ratio. The attenuation can lead to a survivors' rate ratio in an opposite direction from the underlying rate ratio. In addition, the survivors' rate functions for the treatment and control groups are not proportional. The survivors' rate model is not applicable in this scenario.

Scenario III: $\alpha\beta < 0$

By a similar argument as that in Scenario II, it can be seen that β and $\log\{E[Z|D \geq t, x = 1]/E[Z|D \geq t, x = 0]\}$ have the same sign in this scenario. Accordingly the survivors' rate ratio always has the same direction as the rate ratio. However, the existence of the treatment effect on the

terminal event, α, inflates the survivors' rate ratio from the underlying rate ratio, and also causes the nonproportionality of the survivors' rates between the two treatment groups. Again, the proportional rate model using the survivors' rate function does not work in this scenario.

Scenario IV: $\beta \neq 0$ **and** $\alpha = 0$

When the treatment has no effect on the terminal event, it can be seen from (12) that

$$E[Z|D \geq t, x = 1] = E[Z|D \geq t, x = 0] = \frac{E[Z \exp\{-ZH_0(t)\}]}{E[\exp\{-ZH_0(t)\}]}.$$

It means that the survivors at time t in the treatment and the control groups are comparable, in terms of the average latent health status. Hence, $\lambda^S(t|x = 1)/\lambda^S(t|x = 0) = e^\beta$. Here, e^β could be seen either as the rate ratio from the core model, or as the survivors' rate ratio. Consequently, the survivors' rate model can be applied to correctly identify the treatment effect on the recurrent events in this scenario. Here $\lambda^S(t)$ is different from $\lambda(t)$ by a factor, $\mu_Z^{-1} E[Z|D \geq t]$.

Scenario V: $N(t)$ **and** D **are independent for given** x

When the recurrent and terminal event processes are independent within each treatment group, the survivors' rate function $\lambda^S(t|x)$ equals the rate function $\lambda(t|x)$. In such situation, modelling the survivors' rate is equivalent to modelling the rate function. It should be noted that this equivalence does not rely on the assumptions of the core model.

4. Simulations and A Data Analysis

4.1 *Simulations*

We examine the appropriateness of the proportional rate model for the adjusted rate and the survivors' rate under the core model. Assume Z follows a gamma distribution with mean 10 and variance 50, $\lambda_0(t) = 1/10$, and $h_0(t) = t/400$. X is the treatment indicator, where $X = 1$ corresponds to treatment and $X = 0$ control group. The adjusted rate function has the explicit form

$$\lambda^A(t|x) = \frac{1}{125}\exp(x\beta)\left[\frac{1}{800}t^2\exp(x\alpha)+\frac{1}{5}\right]^{-3},$$

and the survivors' rate function is

$$\lambda^S(t|x) = \frac{1}{5}\exp(x\beta)\left[\frac{1}{800}t^2\exp(x\alpha)+\frac{1}{5}\right]^{-1}.$$

[Figure 2 about here.]

[Figure 3 about here.]

Figures 2 and 3 give the adjusted rate functions and the survivors' rate functions with different combinations of (α, β) respectively, with their corresponding log rate ratios. The simulation results confirm the theoretical comparisons we discussed in the previous sections. We also observe that, when the magnitude of the treatment effect on the terminal event is smaller than that on the recurrent events, the violation of proportionality in the adjusted or the survivors' rate functions is less sever, i.e., the adjusted or the survivors' rate ratio is closer to a constant. When the treatment has no effect on the terminal event, the β-values of the three discussed proportional rate

models all represent the same quantity one is trying to estimate, log rate ratio, which is the treatment effect on the recurrent events.

4.2 *Data Analysis*

Now we compare three rate functions using the data from the Terry Beirn Community Programs for Clinical Research on AIDS (CPCRA) didanosine/zalcitabine trial. This clinical trial aimed to compare two treatments, didanosine (ddI) and zalcitabine (ddC), for patients who previously received zidovudine and had 300 or fewer CD4 cells per cubic millimeter or a diagnosis of AIDS (Abrams et al., 1994). A focused interest was the comparison of the survival time and the clinical disease progression for the two treatment arms. Clinical disease progression was defined as the occurrence of certain recurrent diseases from a class of AIDS-defining conditions and some opportunistic diseases (Abrams et al., 1994). This randomized trial assigned 230 patients to receive ddI and 237 to receive ddC. During the follow-up time with the median length 16 months, 100 deaths were recorded in the ddI group and 88 in the ddC group. One hundred twenty of the patients taking ddI experienced a total count of 172 recurrent diseases and 115 taking ddC experienced 191 recurrent diseases.

To compare the survival for the two treatments we plot the Kaplan-Meier estimate of the survival function for both ddI and ddC groups in Figure 4. It confirms that ddC may have provided survival advantage over ddI (Abrams et al., 1994). However, during the first 6 months of follow-up, the survival of patients in the two treatment groups was virtually the same.

In order to examine the treatment effect on clinical disease progression, the following three one-sample methods, the estimation of the cumulative

rate function (Wang et al., 2001), the estimation of the adjusted cumulative rate function (Ghosh and Lin, 2000; Cook and Lawless, 2002), and the estimation of the survivors' cumulative rate function (Cook and Lawless, 1997) are applied. The resulting estimates for both treatment arms are shown in Figure 5.

The estimates of the cumulative rate function in Figure 5(a) show that the estimated cumulative rate of the recurrent diseases for both treatment arms were not significantly different in the middle of the follow-up period (month 8 to month 12). Before month 8, the ddC group had a slightly bigger cumulative rate than the ddI group and after month 12, the ddC group had a smaller cumulative rate than the ddI group. The 95% pointwise Bootstrap confidence bands for the two treatments (not shown) were explored, which overlapped in the first 16 months of follow up, however, showed a significant difference after the 16th month. Therefore, ddC was shown to be more efficacious than, or at least as efficacious as, ddI in delaying disease progression since the 8th month after randomization.

Now we examine the adjusted cumulative rate functions for both treatments in Figure 5(b) and the the survivors' cumulative rate functions in Figure 5(c). It is found that during the first 6 months of follow-up, all three kinds of function estimate give a consistent result that the cumulative rate (or the adjusted cumulative rate, or the survivors' cumulative rate) for the ddC group is slightly bigger than for the ddI group. This is because the two treatments didn't work differently on survival during the first 6 months after randomization (see Figure 4). A remarkable difference between the estimated cumulative rate functions and the estimated adjusted or survivors'

cumulative rate functions appeared after month 6, when both the adjusted and the survivors' cumulative rates showed a "beneficial" effect of ddI over ddC. This confirms our discussion made earlier that the survival difference between the two treatments can attenuate the adjusted and survivors' rate ratio from the underlying rate ratio, or even reverse their directions from the underlying rate ratio. It should be noticed that both the estimated adjusted and survivors' cumulative rate functions were smaller than the estimated cumulative rate function, with the estimated adjusted cumulative rate function even smaller than the survivors' one.

[Figure 4 about here.]

[Figure 5 about here.]

5. Discussion

In clinical trials, it is not uncommon that the follow-up of a patient is stopped by the occurrence of a terminal event. In this situation the assessment of a treatment or a risk factor using recurrent events is likely to be confounded by the presence of the terminal event. Similar situations exist in other data settings, such as competing risk data, where the occurrence of one type of event can preclude the occurrence of the other types of events. In these circumstances, choosing appropriate models is necessary and critical for the correct interpretation of the data. In this paper, we compared three proportional rate models for recurrent event data in the presence of a terminal event. The behaviors of the discussed models, as well as the appropriateness of assuming these models, under different situations were investigated. We showed that

the inappropriate choice of a model may lead to misleading scientific results in various situations.

Acknowledgements

The authors would like to thank Drs Zhiqing Tan, Constantine Frangakis, Qian-Li Xue, Howard Chilcoat, and Larry Wissow at the Johns Hopkins University for their helpful suggestions; Dr. Chiung-Yu Huang at the National Institutes of Health for many detailed suggestions and comments; and Yijian Huang for generously sharing the data used in writing this manuscript.

References

Abrams, D. I., Goldman, A. I., Launer, C., Korvick, J. A., Neaton, J. D., Crane, L. R., Grodesky, M., Wakefield, S., Muth, K., Kornegay, S., Cohn, D. L., Harris, A., Luskin-Hawk, R., Markowitz, N., Sampson, J., Thompson, M., Deyton, L. and the Terry Beirn Community Programs for Clinical Research (1994). A comparative trial of didanosine or zalcitabine after treatment with zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine* **330**, 657–662.

Anderson, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. *Annals of Statistics* **10**, 1100–1120.

Cook, R. J. and Lawless, J. F. (1997). Marginal analysis of recurrent events and a terminating event. *Statistics in Medicine* **16**, 911–924.

Cook, R. J. and Lawless, J. F. (2002). Analysis of repeated events. *Statistical Methods in Medical Research* **11**, 141–166.

Ghosh, D. and Lin, D. Y. (2000). Nonparametric analysis of recurrent events and death. *Biometrics* **56**, 554–562.

Ghosh, D. and Lin, D. Y. (2002). Marginal regression models for recurrent and terminal events. *Statistica Sinica* **12**, 663–688.

Ghosh, D. and Lin, D. Y. (2003). Semiparametric analysis of recurrent events data in the presence of dependent censoring. *Biometrics* **59**, 877–885.

Huang, C.-Y. and Wang, M.-C. (2004). Joint modeling and estimation for recurrent event processes and failure time data. *Journal of American Statistical Association* **99**, 1153–1165.

Huang, Y. and Wang, M.-C. (2003). Frequency of recurrent events at failure time: modeling and inference. *Journal of American Statistical Association* **98**, 663–670.

Lin, D. Y., Wei, L. J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *Journal of Royal Statistical Society, Series B* **62**, 711–730.

Liu, L., Wolfe, R. A. and Huang, X. L. (2004). Shared frailty models for recurrent events and a terminal event. *Biometrics* **60**, 747–756.

Pepe, M. S. and Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of American Statistical Association* **88**, 811–820.

Prentice, R. L., Williams, B. J. and Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika* **68**, 373–379.

Schaubel, D. E. and Cai, J. (2005). Analysis of clustered recurrent event data with application to hospitalization rates among renal failure patients. *Biostatistics* **6**, 404–419.

Wang, M.-C., Qin, J. and Chiang, C. T. (2001). Analyzing recurrent event data with informative censoring. *Journal of American Statistical Association* **96**, 1057–1065.

Appendix

We now prove that for a univariate real-valued covariate X, $E[Z|D \geq t, x]$ is monotonic in x under the core model assumptions. We first consider the partial derivative of $E[Z|D \geq t, x]$ with respect to x:

$$\begin{aligned}
& \frac{\partial}{\partial x} E[Z|D \geq t, x] \\
= \; & \frac{\partial}{\partial x} \left(\frac{E[Z \exp\{-Z \exp(x\alpha) H_0(t)\}|x]}{E[\exp\{-Z \exp(x\alpha) H_0(t)\}|x]} \right) \\
= \; & \alpha \exp(x\alpha) H_0(t) \{ E^2[Z \exp\{-Z \exp(x\alpha) H_0(t)\}] \\
& - E[Z^2 \exp\{-Z \exp(x\alpha) H_0(t)\}] E[\exp\{-Z \exp(x\alpha) H_0(t)\}] \}.
\end{aligned}$$

By Cauchy-Schwarz inequality we know that

$$\begin{aligned}
& E^2[Z \exp\{-Z \exp(x\alpha) H_0(t)\}] \\
= \; & E^2[Z \exp\{-Z \exp(x\alpha) H_0(t)/2\} \cdot \exp\{-Z \exp(x\alpha) H_0(t)/2\}] \\
\leq \; & E[Z^2 \exp\{-Z \exp(x\alpha) H_0(t)\}] E[\exp\{-Z \exp(x\alpha) H_0(t)\}].
\end{aligned}$$

Therefore, when $\alpha < 0$, $\frac{\partial}{\partial x} E[Z|D \geq t, x] \geq 0$, which means that $E[Z|D \geq t, x]$ is non-decreasing in x. Similarly, when $\alpha > 0$, $\frac{\partial}{\partial x} E[Z|D \geq t, x] \leq 0$, which means that $E[Z|D \geq t, x]$ is non-increasing in x. Obviously, $E[Z|D \geq t, x = 1] \geq E[Z|D \geq t, x = 0]$ when $\alpha < 0$; $E[Z|D \geq t, x = 1] \leq E[Z|D \geq t, x = 0]$ when $\alpha > 0$.

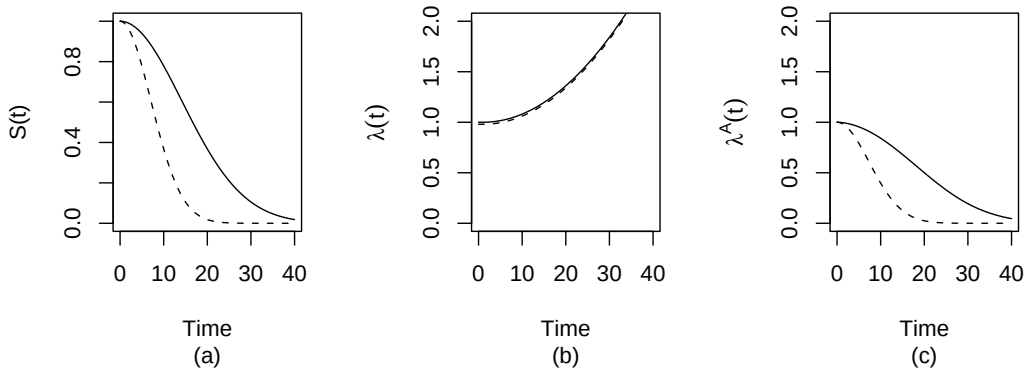

Figure 1. Misleading result from adjusted rate functions. (a) Survival functions for death; (b) rate functions for recurrent diseases; and (c) adjusted rate functions (——, the new drug; - - - , the existing drug).

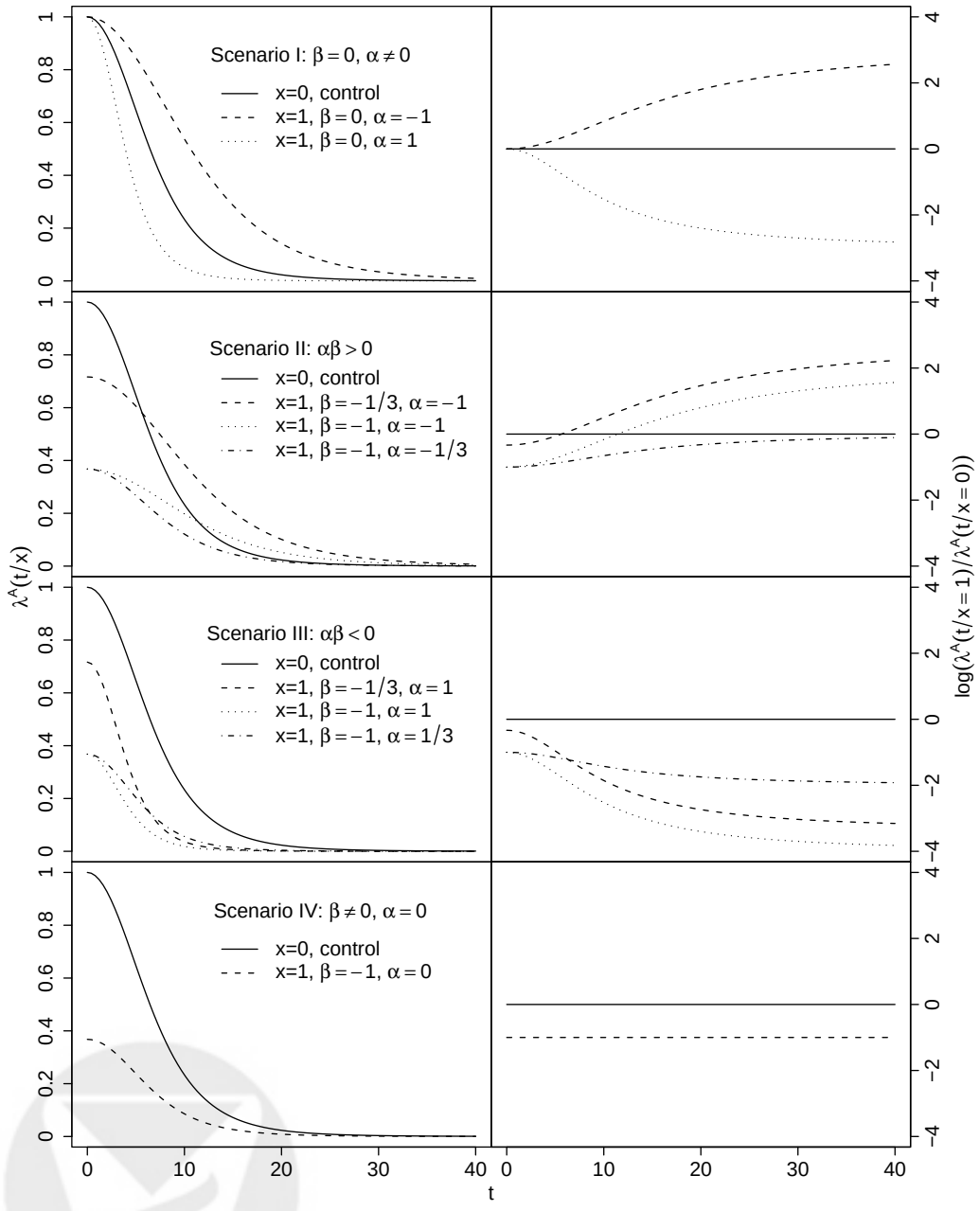

Figure 2. Adjusted rate and logarithm of adjusted rate ratio, $\lambda_0(t) = 1/10$, $h_0(t) = t/400$, $Z \sim$ gamma (mean=10, variance=50).

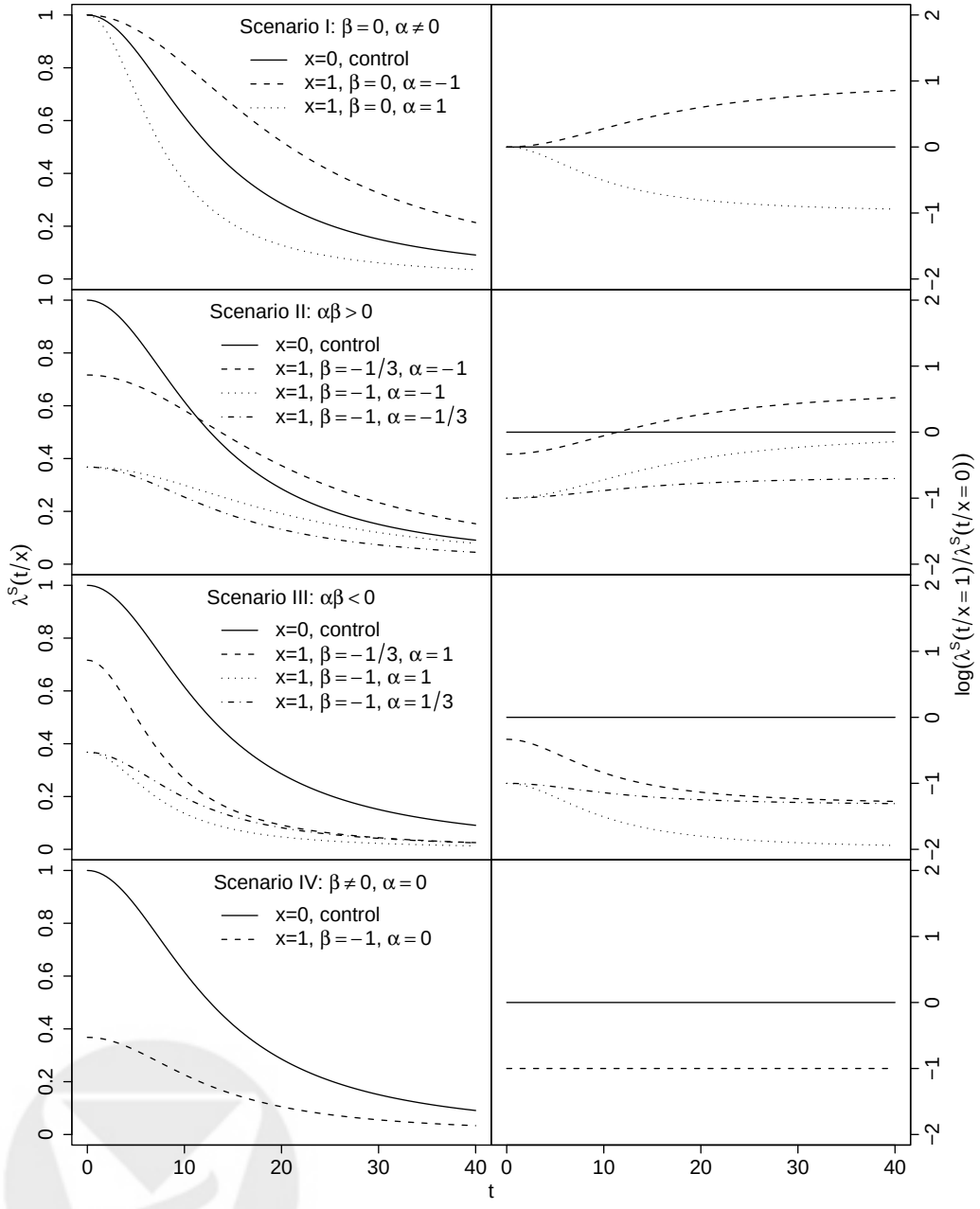

Figure 3. Survivors' rate and logarithm of survivors' rate ratio, $\lambda_0(t) = 1/10$, $h_0(t) = t/400$, $Z \sim$ gamma (mean=10, variance=50).

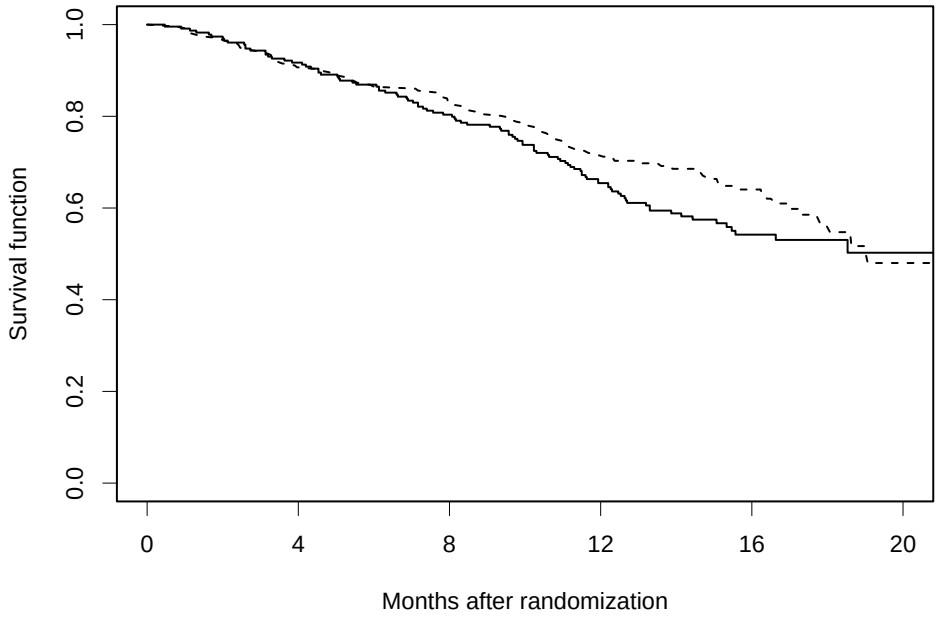

Figure 4. Kaplan-Meier estimates of the survival function for death from CPCRA ddI/ddC data (——, ddI; – – –, ddC).

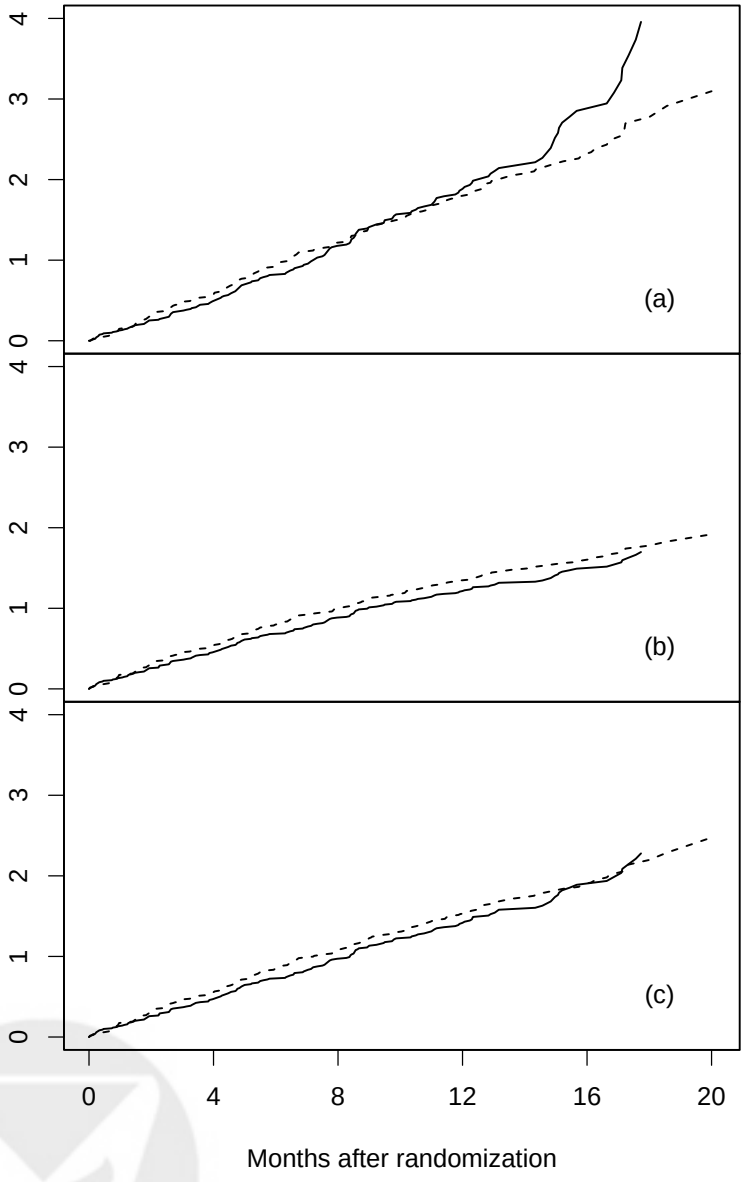

Figure 5. Estimates of three types of cumulative rate function for recurrent events from CPCRA ddI/ddC data. (a) Cumulative rate function; (b) adjusted cumulative rate function; and (c) survivors' cumulative rate function (——, ddI; – – –, ddC).

Table 1

Rate functions and proportional rate models

	Definition	Model
Rate	$\lambda(t)dt = E[dN(t)]$	$\lambda(t\|\boldsymbol{x}, z) = z\lambda_0(t)\exp(\boldsymbol{x\beta})$
		$h(t\|\boldsymbol{x}, z) = zh_0(t)\exp(\boldsymbol{x\alpha})$
Adjusted rate	$\lambda^A(t)dt = E\left[I(D \geq t)dN(t)\right]$	$\lambda^A(t\|\boldsymbol{x}) = \lambda_0^A(t)\exp(\boldsymbol{x\beta}^A)$
Survivors' rate	$\lambda^S(t)dt = E[dN(t)\|D \geq t]$	$\lambda^S(t\|\boldsymbol{x}) = \lambda_0^S(t)\exp(\boldsymbol{x\beta}^S)$