*University of Michigan School of Public Health*
The University of Michigan Department of Biostatistics Working Paper Series

*Year* 2006 *Paper* 57

# Semiparametric Analysis for Correlated Recurrent and Terminal Events

Yining Ye* Jack Kalbfleisch†
Doug E. Schaubel‡

*University of Michigan Biostatistics, yey@umich.edu
†University of Michigan, jdkalbfl@umich.edu
‡University of Michigan, deschau@umich.edu

# Semiparametric Analysis for Correlated Recurrent and Terminal Events

Yining Ye, Jack Kalbfleisch, and Doug E. Schaubel

**Abstract**

In clinical and observational studies, recurrent event data (e.g. hospitalization) with a terminal event (e.g. death) are often encountered. In many instances, the terminal event is strongly correlated with the recurrent event process. In this article, we propose a semiparametric method to jointly model the recurrent and terminal event processes. The dependence is modeled by a shared gamma frailty that is included in both the recurrent event rate and terminal event hazard function. Marginal models are used to estimate the regression effects on the terminal and recurrent event processes and a Poisson model is used to estimate the dispersion of the frailty variable. A sandwich estimator is used to achieve additional robustness. An analysis of hospitalization data for patients in the peritoneal dialysis study is presented to illustrate the proposed method.

# Semiparametric Analysis of Correlated Recurrent and Terminal Events

Yining Ye

Email: yey@umich.edu
Office: (734)615-3665 Fax: (734)763-2215

John D. Kalbfleisch

Email: jdkalbfl@umich.edu
Office: (734)615-7067 Fax: (734)615-7068

Douglas E. Schaubel

Email: deschau@umich.edu
Office: (734)615-9825 Fax: (734)763-2215

Department of Biostatistics
University of Michigan
1420 Washington Heights
Ann Arbor, Michigan, 48109-2029

*March 19, 2006*

# Semiparametric Analysis for Correlated Recurrent and Terminal Events

SUMMARY. In clinical and observational studies, recurrent event data (e.g. hospitalization) with a terminal event (e.g. death) are often encountered. In many instances, the terminal event is strongly correlated with the recurrent event process. In this article, we propose a semiparametric method to jointly model the recurrent and terminal event processes. The dependence is modeled by a shared gamma frailty that is included in both the recurrent event rate and terminal event hazard function. Marginal models are used to estimate the regression effects on the terminal and recurrent event processes and a Poisson model is used to estimate the dispersion of the frailty variable. A sandwich estimator is used to achieve additional robustness. An analysis of hospitalization data for patients in the peritoneal dialysis study is presented to illustrate the proposed method.

KEY WORDS: Frailty; Proportional hazard model; Rate function; Recurrent event; Survival analysis.

## 1. Introduction

Data on recurrent events frequently arise in clinical and observational studies. Examples include repeated hospitalizations, the occurrence of new tumors in patients with superficial bladder cancer and the occurrences of opportunistic infections in HIV-infected subjects. Various methods have been considered for the analysis of recurrent events. These methods include the complete intensity approach (e.g. Prentice, Williams and Peterson, 1981) and the marginal rate approach (e.g. Pepe and Cai, 1993; Lawless and Nadeau, 1995; Lin, Wei, Yang, and Ying, 2000). In these approaches, it is assumed that, conditional on the covariates in the model, the censoring is independent of the recurrent events. In many instances, however, there exists a terminal event, death for example, which precludes the occurrence of further events. Further, it is often the case that the terminal event is strongly correlated with the recurrent event process. More explicitly, if the rate of the recurrent event is unusually high (low) in an individual, that individual is also subject to increased (decreased) rate of death.

Methods of analysis of repeated events in the presence of a terminal event can also be classified into two categories. There are analyses that focus on the marginal rates of the recurrent and terminal events and complete intensity approaches in which frailties are used to account for the correlation between the rates of recurrent and terminal events.

Marginal models have been considered by several authors. In these, the rate functions are not taken to be complete intensity functions but rather correspond to average rates that would arise across the population (e.g. Ghosh and Lin, 2002). The correlation between the recurrent event process and the terminal event is left unspecified in these models. Frailty models or shared random effects models specify the dependence between the recurrent events and the terminal event by allowing a common frailty variable to have a multiplicative effect on their respective rates. Thus, they assume that the complete intensity of the recurrent events and the terminal event is fully specified by the observed covariates and the unobserved frailty

(e.g. Wang, Qin and Chiang, 2001; Huang and Wang, 2004; Liu, Wolfe and Huang, 2004). In all of the frailty models, it is assumed that given the frailty, the recurrent event process is a nonhomogeneous Poisson process and this plays a central role in all aspects of the estimation. It is to be expected, therefore, that the estimation procedures will be sensitive to deviations from the Poisson assumption.

We propose a joint semiparametric model in which the correlation between the terminal event and the recurrence event is incorporated through the frailty. Our model for the event rates has the spirit of a marginal model, however, in that it is conditional only on the covariates (and the frailty) and not on the previous history of the process. The estimation of the regression coefficients is based on the estimating functions for marginal rate models. Different from the marginal rate model proposed previously, the proposed method provides a way to estimate the degree of dependence between the two processes.

The remaining of the article is organized as follows. In Section 2, the model is specified. A series of estimating equations are specified to estimate the parameters in the models and numerical methods are described. In Section 3, the proposed approach is compared with the method of Huang and Wang (2004). Section 4 gives results of some simulation studies and the method is applied to the data from a prospective study of peritoneal dialysis patients in Section 5. The paper concludes with some discussion in Section 6.

## 2. Model

Let $C_i$ and $D_i$ be the censoring and death (or terminal event) time and $T_{ik}$ be the $k$th recurrent event time for the $i$th subject, $i = 1, 2, \ldots, n$; $k = 1, \ldots, m_i$. Let $N_i^{R*}(t) = \int_0^t dN_i^{R*}(u)$ be the actual number of recurrent events in the time interval $(0, t]$ for the $i$th subject and $N_i^{D*}(t) = I(D_i \leq t) = \int_0^t dN_i^{D*}(u)$. Let $X_i = min(C_i, D_i, \tau)$ and let $Y_i(t) = I(X_i \geq t)$ be the at-risk indicator, where $\tau$ is the study ending time. Two processes are observed during the time interval $[0, \min(X_i, \tau)]$, namely $N_i^R(t) = \int_0^t Y_i(u) dN_i^{R*}(u)$ and $N_i^D(t) = \int_0^t Y_i(u) dN_i^{D*}(u)$,

where $N_i^R(t)$ and $N_i^D(t)$ are the observed numbers of recurrent events and deaths respectively. The observed data for subject $i$ at time t is denoted by $O_i(t) = \{Y_i(u), N_{Ri}(u), N_{Di}(u), 0 \leq u \leq t\}$. Let $\widetilde{Z}(t) = \{\widetilde{Z}(u), 0 \leq u < t\}$ be the covariate history for an individual and $Z(t)^T = \{Z_1(t), Z_2(t), \ldots\}$ comprise functions of $\widetilde{Z}(t)$. For simplicity, we consider that $Z$ is time-independent, but the proposed joint model can easily incorporate time-dependent covariates.

We consider a (partial) marginal rate of the recurrent event given $D = s$ and $\gamma$, which is defined as $d\Lambda_R(t|\gamma) = E[dN^{R*}(t)|Z, D = s, \gamma]$, $s \geq t$. It is the average rate of the recurrent events at time $t$ associated with $Z$ for those individuals with frailty $\gamma$ and whose survival time is $s$, where $s \geq t$. Note that $d\Lambda_R(t)$ may depend on $Z$ and the frailty $\gamma$, but does not depend on death time $s$. This in effect assumes that $\gamma$ accounts for the correlation between the recurrent events and death. Our method explicitly acknowledges the fact that death stops further recurrent events in that, given $t > D$, $dN^{R*}(t)$ takes the value 0.

The joint semi-parametric model that we consider can be expressed as,

$$d\Lambda_R(t|\gamma) = \gamma \exp(\beta^T Z) d\Lambda_{0R}(t), \tag{1}$$

$$d\Lambda_D(t|\gamma) = \gamma \exp(\alpha^T Z) d\Lambda_{0D}(t), \tag{2}$$

where $d\Lambda_D(t|\gamma) = P(dN^{D*}(t) = 1|\ D \geq t, \gamma, Z)$ is the hazard function for $D$ and $d\Lambda_{0D}(t)$ and $d\Lambda_{0R}(t)$ are the unspecified baseline hazard function for death and the baseline recurrent event rate respectively. For convenience, we assume that the frailty $\gamma$ has a gamma distribution with mean 1, variance $\theta$, and density $f_\theta(\gamma) = \frac{1}{\Gamma(1/\theta)\theta^{1/\theta}} \exp(-\gamma/\theta)\gamma^{1/\theta}$. As is the usual convention for frailty models, the mean $E[\gamma] = 1$ is fixed for identifiability an the distribution of $\gamma$ is assumed to be independent of $Z$. It should be noted that the joint model can handle different covariate vectors in the recurrent event and death rate model by fixing the appropriate elements of $\alpha$ and $\beta$ to 0. The above model can also be generalized to allow different effects of frailty on the recurrent event process and death as in Liu et al. (2004).

The following additional assumptions are made for the joint model:

1. Censoring is independent. Thus the distribution of censoring time $C$ may depend on $Z$ but not on $\gamma$, $N^{R*}(.)$ or $N^{D*}(.)$; i.e.,

$$\lim_{h\to 0}\frac{1}{h}P(t\le C<t+h|Z,N^{R*}(u),N^{D*}(u),\gamma,0<u<t)=\lim_{h\to 0}\frac{1}{h}P(t\le C<t+h|Z).$$

2. The recurrent event process and death process are continuous. As such, the recurrent event and death cannot happen at the same time.

3. For the purpose of estimating the distribution of $\gamma$, we assume that given $Z$ and $\gamma$, the recurrent event process $N^{R*}(\cdot)$ before death follows a non-stationary Poisson process with intensity function $\gamma\exp(\beta^T Z)d\Lambda_{0R}(t)$.

If the frailty, $\gamma$, is known, the estimating equations for $\alpha$ and $\beta$ are as discussed in Lin et al (2000) and are identical to those that arise from the usual partial likelihood (Cox, 1972). However, $\gamma_i$ is not observed. Therefore we consider an induced marginal model for $\alpha$ and $\beta$,

$$\begin{aligned} d\Lambda_R(t) &= E[dN^{R*}(t)|Z,D\ge t],\\ d\Lambda_D(t) &= E[dN^{D*}(t)|Z,D\ge t]. \end{aligned}$$

Taking the conditional expectation of (1) and (2) given $Z$ and $D\ge t$, we obtain,

$$d\Lambda_R(t)=w(t)\exp(\beta^T Z)d\Lambda_{0R}(t), \tag{3}$$

$$d\Lambda_D(t)=w(t)\exp(\alpha^T Z)d\Lambda_{0D}(t), \tag{4}$$

where $w(t)=E[\gamma|D\ge t,Z]=\{1+\Lambda_{0D}(u)\exp(\alpha^T Z)\theta\}^{-1}$ under the assumed gamma distribution for $\gamma$. Given $w(t)$, the models (3) and (4) have a standard proportional rate/hazard form. Estimating equations for $\alpha$ and $\beta$ can be obtained by taking the first derivatives of the pseudo partial likelihood arising from (3) and (4), treating $w(t)$ as a known function. To

estimate $\Lambda_{0D}$ and $\Lambda_{0R}$, we use the Nelson-Aalen type estimators. In order to estimate $\theta$, we use likelihood methods and introduce the assumption that conditional on $\gamma$, the recurrent event process follows a nonhomogeneous Poisson process with intensity $d\Lambda_R(t|\gamma)$. Let $\delta_{ik}$ be the indicator of the recurrent event at time $t_{ik}$. The likelihood based on $\gamma_i$ and the observed data $O_i(\tau)$ is,

$$\begin{aligned} L(O_i(\tau),\gamma_i) &= \{\gamma_i \exp(\alpha^T Z_i) d\Lambda_{0D}(X_i)\}^{\Delta_i} \times \exp\{-\gamma_i d_i\} \\ &\quad \times \exp\{-\gamma_i r_i\} \times \prod_{k=1}^{m_i} \{\gamma_i \exp(\beta^T Z_i) d\Lambda_{0R}(t_{ik})\} f_\theta(\gamma_i). \end{aligned}$$

where $\Delta_i = I(D_i \leq min(C_i, \tau))$, $r_i = \int_0^\infty Y_i(u) \exp(\beta^T Z_i) d\Lambda_{0R}(u)$, $d_i = \int_0^\infty Y_i(u) \exp(\alpha^T Z_i) d\Lambda_{0D}(u)$ and $m_i$ is the number of recurrent events experienced by the $i$th subject.

Integrating over $\gamma_i$ and taking a product over $i$ gives the likelihood,

$$L(\theta) = \prod_{i=1}^{n} L(O_i(\tau)) \propto \prod_{i=1}^{n} \frac{\Gamma(m_i + \Delta_i + 1/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}(r_i + d_i + 1/\theta)^{(c_i+1/\theta)}}. \tag{5}$$

Differentiating the logarithm of (5) with respect to $\theta$ gives the estimating equation for $\theta$. As noted earlier, the Poisson assumption is only utilized in the estimating equation for $\theta$ but not directly in the estimation of $\alpha$ and $\beta$.

Let $\eta = (\beta, \alpha, \theta, d\Lambda_{0D}, d\Lambda_{0R})$ and for a parameter $\phi$ (e.g. $\phi = \alpha$), define

$$S_1^{(k)}(\phi, t) = n^{-1} \sum_{i=1}^{n} Y_i(t) w_i(t) Z_i^{\otimes k} \exp(\phi^T Z_i),$$

(k=0,1,2), where $a^{\otimes 0} = 1$, $a^{\otimes 1} = a$ and $a^{\otimes 2} = aa^T$. Further, let $t_{R1}, \ldots, t_{Rm}$ be the ordered distinct recurrent event times and $t_{D1}, \ldots, t_{Df}$ be the ordered failure times. Estimates of the intensities are discrete with jumps at the distinct event times. We let $\lambda_{0R} = (\lambda_{0R1}, \lambda_{0R2}, \ldots, \lambda_{0Rm})^T$ and $\lambda_{0D} = (\lambda_{0D1}, \lambda_{0D2}, \ldots, \lambda_{0Df})^T$, where $\lambda_{0Rj} = d\Lambda_{0R}(t_{Rj})$, $j = 1, \ldots, m$ and $\lambda_{0Dj} = d\Lambda_{0D}(t_{Dj})$, $j = 1, \ldots, f$. Define $d_{Rj}$ and $d_{Dj}$ as the number of recurrent events at $t_{Rj}$ and the number of death at $t_{Dj}$ respectively. Note that the ties are being handled using the Breslow approximation (Kalbfleisch and Prentice, 2002, Section 4.2.3). The

unbiased estimating equations are $U(\eta) = (U_1^T, U_2^T, U_3, U_4^T, U_5^T)^T = 0$, where the components of $U$ respectively correspond to $\beta$, $\alpha$, $\theta$, $\lambda_{0D}$ and $\lambda_{0R}$. We have

$$\begin{aligned} U_1 &= \sum_{i=1}^{n} \int_0^\infty \left\{ Z_i - \frac{S_1^{(1)}(\beta, u)}{S_1^{(0)}(\beta, u)} \right\} dN_i^R(u), \\ U_2 &= \sum_{i=1}^{n} \int_0^\infty \left\{ Z_i - \frac{S_1^{(1)}(\alpha, u)}{S_1^{(0)}(\alpha, u)} \right\} dN_i^D(u). \\ U_3 &= \frac{\partial \log L(\theta)}{\partial \theta}. \end{aligned}$$

Finally, the $j$th elements of $U_4$ and $U_5$ are,

$$U_{4j} = d_{Dj} - nS_1^{(0)}(\alpha, t_{Dj})\lambda_{0Dj}, \quad j = 1, \ldots f. \tag{6}$$

$$U_{5j} = d_{Rj} - nS_1^{(0)}(\beta, t_{Rj})\lambda_{0Rj}, \quad j = 1, \ldots m \tag{7}$$

Our numerical approach to solve $U(\eta) = 0$ converges quickly and can be summarized as follows:

1. Let $\theta^{(0)}$, $\alpha^{(0)}$ and $\Lambda_{0D}^{(0)}(u)$ be initial estimates. Typically, we can set $\theta^{(0)} = 1$, $\alpha^{(0)} = 0$ and let $\Lambda_{0D}^{(0)}(u)$ be the Nelson-Aalen type estimate of the cumulative death hazard for the sample.

2. Let $w_i^{(0)}(u) = w_i(u; \Lambda_{0D}^{(0)}, \alpha^{(0)}, \theta^{(0)})$.

3. Replace $w_i(u)$ with $w_i^{(0)}(u)$ in $U_1$, $U_2$, $U_4$ and $U_5$ and solve the resulting equations $U_1 = 0$, $U_2 = 0$, $U_4 = 0$ and $U_5 = 0$ respectively for updated estimates $\beta^{(1)}$, $\alpha^{(1)}$, $\lambda_{0D}^{(1)}$ and $\lambda_{0R}^{(1)}$.

4. Given $\alpha^{(1)}$, $\beta^{(1)}$, $\lambda_{0D}^{(1)}$ and $\lambda_{0R}^{(1)}$, update estimate of $\theta$ to $\theta^{(1)}$ from $U_3(\theta)$.

5. Replace $\theta^{(0)}$, $\alpha^{(0)}$, $\Lambda_{0D}^{(0)}(u)$ with $\theta^{(1)}$, $\alpha^{(1)}$ and $\Lambda_{0D}^{(1)}(u)$. Repeat step (2) to (4) until the estimates of $\theta$, $\alpha$ and $\beta$ converge.

In order to establish the asymptotic results for this approach, it seems that we shall need at a minimum the following four conditions for $i = 1, \ldots, n$:

- $\{N_i^R(.), N_i^D(.), Z_i(.), \gamma_i\}$ are independently and identically distributed.
- $Pr(C_i \geq \tau) > 0$.
- $N_i^R(\tau)$ is bounded by a constant.
- $A = -n^{-1}\partial U(\eta)/\partial \eta^T$ is positive definite with probability 1.

These regularity conditions are similar to those of Lin et al. (2000). Under these conditions, the proposed procedure should lead to consistent estimation of all parameters $(\alpha, \beta, \theta, \Lambda_{0D}(u), \Lambda_{0R}(u), u < \tau)$ and the profiled scores for $\alpha, \beta, \theta$ should be asymptotically normal. Following the approach of Lin et al (2000), it can be seen that the components of $U_1, U_2$ and $U_3$ are asymptotically uncorrelated random variables, and by arguments developed there, a central limit theorem would apply. These estimating equations, however, also contain the functions $\Lambda_{0D}$ and $\Lambda_{0R}$ which are estimated using the Nelson-Aalen type estimators as in equations (7) and (6). Uncertainty in these estimates would need to be accounted for in the asymptotic results for $\hat{\beta}$, $\hat{\alpha}$ and $\hat{\theta}$.

Following Murphy (1995), we consider a discrete version of the baseline hazard and rate functions with jumps only at the distinct event times. Let $A(\eta)$ be defined as above and let $\Sigma(\eta) = n^{-1}\sum_{i=1}^n U_i(\eta)^{\otimes 2}$. Let $\hat{\eta}$ be the estimate of $\eta$ and let $\hat{A} = A(\hat{\eta})$ and $\hat{\Sigma} = \Sigma(\hat{\eta})$. Analogous to the results of Murphy (1995) and Parner (1998), we expect that the asymptotic distribution of $\hat{\alpha}, \hat{\beta}$ and $\hat{\theta}$ should be asymptotically normal with covariance estimated by the appropriate submatrix of $\hat{A}^{-1}\hat{\Sigma}(\hat{A}^{-1})^T$. By using the sandwich estimator our estimation should be robust to deviations from the Poisson process assumption and also should account for possible correlations induced by only making marginal assumptions on the death and recurrent event rates. Additional work is needed in developing a full asymptotic treatment of this approach.

The dimension of $A$ will increase as the sample size increases, which might lead to calculation difficulties for large samples. However, it is possible to simplify the calculation so that we need only numerically invert a matrix of smaller dimension. Let $\eta_1^T = (\beta^T, \alpha^T, \theta, \lambda_{0D}^T)$, $\eta_2 = \lambda_{0R}$, $U^{(1)}(\eta) = (U_1^T, U_2^T, U_3, U_4^T)^T$ and $U^{(2)}(\eta) = U_5$, and write

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = n^{-1} \begin{pmatrix} -\frac{\partial U^{(1)}(\eta)}{\partial \eta_1^T} & -\frac{\partial U^{(1)}(\eta)}{\partial \eta_2^T} \\ -\frac{\partial U^{(2)}(\eta)}{\partial \eta_1^T} & -\frac{\partial U^{(2)}(\eta)}{\partial \eta_2^T} \end{pmatrix}.$$

The dimension of $A$ is $2p + 1 + f + m$, which may be large as the sample size increases. The direct numerical inversion may be time consuming. Since $A_{22}$ is a diagonal $(m \times m)$ matrix, however, calculation is simplifed by noting that

$$A^{-1} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} J^{-1} & -J^{-1}F_2 \\ -F_1 J^{-1} & A_{22}^{-1} + F_1 J^{-1} F_2 \end{pmatrix}$$

where $F_1 = A_{22}^{-1} A_{21}$, $F_2 = A_{12} A_{22}^{-1}$ and $J = A_{11} - A_{12} F_1$. It follows that only a matrix of dimension $2p + 1 + f$ need to be inverted directly. One could also let $\eta_1 = (\beta, \alpha, \theta)$ and $\eta_2 = (\lambda_{0D}, \lambda_{0R})$. In this case, a matrix of dimension $2p + 1$ need to be inverted directly in addition to the relatively straightforward inversion of an upper triangular matrix of dimension $m + f$ corresponding to the partial derivatives, $-\partial U^{(2)}(\eta)/\partial \eta_2^T$. Another possible approach is to use bootstrapping methods to estimate the standard errors of the estimators so that no matrix inversion is needed. Finally, in very large samples, a piece-wise constant model for the baseline hazard and rate functions with a fixed number of jump points could be fitted to avoid computational difficulties.

## 3. Comparison

Wang et al. (2001) considered the analysis of recurrent events in a case where the censoring may be dependent. Let $C_{di}$ be the dependent censoring time for subject $i$ $(i = 1, 2, \ldots, n)$ and let $\gamma^\dagger$ be a nonnegative latent frailty type variable with mean $\mu$. No parametric assumption is made on the distribution of $\gamma^\dagger$. Conditional on $\gamma^\dagger$ and $Z$, it is assumed that $N^{R*}(t)$ is a

nonhomogeneous Poisson process with intensity,

$$d\Lambda_R(t) = \gamma^\dagger d\Lambda_{0R}^\dagger(t) \exp(\beta^T Z), \tag{8}$$

where $d\Lambda_{0R}^\dagger(t)$ is the continuous baseline intensity function with $\int_0^\tau d\Lambda_{0R}^\dagger(u) = 1$ and $\beta$ measures the covariate effect on the average rate of the recurrent event. Their crucial assumption is that conditional on $\gamma^\dagger$, $C_d$ is independent of $N^{R*}(t)$.

The estimation procedure of Wang et al. (2001) relies on the Poisson assumption. Specifically, it is noted that given $(\gamma_i^\dagger, C_{di}, Z_i, m_i)$, the observed times $(T_{i1}, T_{i2}, \ldots, T_{im_i})$ are the order statistics of a set of independent and identically distributed random variables with density function $\pi_i(t) = \lambda_{0R}^\dagger(t)/\Lambda_{0R}^\dagger(X_i)$, $0 \le t \le X_i$. Here, $m_i$ is the number of events occurring before $X_i = \min(C_{di}, \tau)$.

Note that the conditional density $\pi_i(t)$ does not depend on $\gamma_i^\dagger$ or $z_i$ and $\pi_i(t)$ is a truncated density function of $\lambda_{0R}^\dagger(t)$. The cumulative distribution of $\lambda_{0R}^\dagger(t)$, $\Lambda_{0R}^\dagger(t)$ can be estimated by a nonparametric maximum likelihood estimator, which has a simple product-limit form,

$$\hat{\Lambda}_{0R}^\dagger(t) = \prod_{t_{Rj} > t} \left(1 - \frac{d_{(Rj)}}{R_{(j)}}\right), \tag{9}$$

where $R_{(j)}$ is the total number of events with event time and observed terminating time satisfying $\{t_{ik} \le t_{Rj} \le X_i\}$, $k = 1, \ldots, m_i$ and $j = 1, \ldots, m$. Therefore the estimating equation of $\beta$ can be formed by applying the information obtained from $\Lambda_{0R}^\dagger(t)$. The class of estimating equation is defined as,

$$n^{-1} \sum_{i=1}^{n} \bar{Z}_i^T \{m_i \Lambda_{0R}^\dagger(X_i)^{-1} - \exp(\beta_a^T \bar{Z}_i)\} = 0,$$

where $\bar{Z}_i = (1, Z_i)^T$ and the augmented $\beta_a^T = \{\ln(\mu), \beta^T\}$.

Huang and Wang (2004) extended the method to incorporate situations where one aspect of informative censoring is associated with a terminal event (e.g. death). By adding a model

for the intensity of the death process to (8), their joint complete intensity model can be expressed as,

$$d\Lambda_R(t) = \gamma^\dagger d\Lambda_{0R}^\dagger(t)\exp(\beta^T Z),$$
$$d\Lambda_D(t) = \gamma^\dagger d\Lambda_{0D}^\dagger(t)\exp(\alpha^T Z), \tag{10}$$

where $\Lambda_{0D}^\dagger$ is the baseline cumulative hazard. Thus, they assume that conditional on $Z$ and $\gamma^\dagger$, $(N^{R*}(t),C_d,D)$ are mutually independent and $N^{R*}(t)$ is a nonhomogeneous Poisson process. Note that this model is based on an assumed latent process of recurrent events that continues past the death time $D$ so that the methods of Wang et al. (2001) can be directly applied to obtain an estimate of $\beta$. It is then proposed to estimate $\gamma_i^\dagger$ by,

$$\hat{\gamma}_i^\dagger = \frac{m_i}{\hat{\Lambda}_{0R}^\dagger(X_i)\exp\hat{\beta}^T Z_i},$$

and plugging $\hat{\gamma}_i^\dagger$ into the score function from (10), $\alpha$ can be estimated. Empirical process theory was used to study the large-sample properties of $\hat{\alpha}$ and $\hat{\Lambda}_{0D}^\dagger(t)$.

The differences from our suggested model (MR model) and this non-parametric frailty approach (NPF model) proposed by Huang and Wang (2004) can be summarized as follows:

1. The NPF model assumes that conditional on the frailty variable $\gamma^\dagger$, the recurrent event process is independent of the death process. In the MR model, we recognize the fact that death stops further recurrent events and the marginal rate is defined as $d\Lambda_R(t) = E[dN^{R*}(t)|Z, D = s, \gamma]$, for $s \geq t$, which incorporates a kind of conditional independence. This gives the rate of recurrent events among individuals who are alive at time $t$. The recurrent event process is not independent of the death process even conditional on the frailty. The NPF model could be redefined in a similar manner to avoid the need for assuming a latent recurrent event process.

2. The independent censoring assumption is relaxed via the use of frailty in the NPF model. The assumption, however is required for the MR model. It can be relaxed,

but it requires modeling of the censoring distribution or the use of an inverse weighting method to adjust for the dependence.

3. For both models, the frailty is assumed to act as a multiplicative factor on both the hazard and the rate functions and thus induces the dependence between the recurrent event process and the death process. The frailty distribution is left nonparametric in the NPF model whereas it is modeled in the MR model. As a consequence,

   - Direct inferences about the relationship between the recurrent event process and the death process are not made in the NPF model. In the MR model, the correlation is modeled by the parameters of the frailty distribution, for example the variance $\theta$ in the gamma frailty model. Let $r_1(t) = E[N_i^{R*}(t)|Y_i(t) = 1, Zi, dN_i^{D*}(t) = 1]$ and $r_2(t) = E[N_i^{R*}(t)|Y_i(t) = 1, Zi, dN_i^{D*}(t) = 0]$. Then $\frac{r_1(t)}{r_2(t)} = \theta + 1$. For instance, if $\theta = 1$, the expected number of recurrent events in $(0, t]$ for individuals who die at time t is twice the expected number for individuals with identical covariates who are known to survive until time t. On the other hand, if $\theta = 0$, the expected number of recurrent events would be the same no matter whether the individual is known to die at time t or survive until time t. In this way, $\theta = 0$ indicates the independence of the two processes.

   - The parametric assumption made in the MR model may not be robust to the misspecification of the frailty distribution though the efficiency should be increased. We performed some simulations in Section 5 to assess the performance of the proposed estimators when the gamma distribution is correct and when it is misspecified.

4. Time-dependent covariates are not allowed in the NPF model, but can be easily handled by the MR model.

5. Though the correlation between the recurrent event and the death processes is modeled by the frailty, the NPF approach uses separate procedures for the estimation. Information from the death process is not used in the estimation of $\beta$, and extra variation is brought to the estimation of $\alpha$ by $\hat{\gamma}^{\dagger}$. The MR model takes advantage of the assumed correlation and more efficiency should be expected. On the other hand, because the estimating equations in the MR model are more complicated in the way that the nonparametric and parametric components interact, asymptotic properties are more difficult to establish.

6. The Poisson assumption is made in both models for the recurrent event process. It is the key to estimation of the baseline rate the NPF model, whereas in the MR model, the assumption is only applied for updating $\theta$. Thus, it is expected that the MR could be be more robust to the departures from the Poisson assumption.

7. The nonparametric estimator of the baseline rate function in the NPF model (9) does not use the assumption that $\Lambda_R(t) = \gamma d\Lambda_{0R}(t) \exp(\beta^T Z)$ as the Nelson-Aalen estimator does in the MR model. As a consequence, we expect the NPF model to be less efficient than the MR model.

## 4. Simulation Study

Simulations were carried out to evaluate our proposed method and to compare the MR model to the NPF model. One single binary covariate, $Z$, was generated taking value 1 or 0 with probability 0.5. The censoring time was taken to follow a continuous uniform distribution on $[1, 10]$. Given the frailty $\gamma$ and the covariate $Z$, a subject's recurrent event process was a nonhomogeneous Poisson process with the corresponding intensity function $d\Lambda_R(t) = \gamma \exp(\beta Z)dt$. Similarly, the terminal event time was generated from an exponential distribution with hazard $d\Lambda_D(t) = 0.2\gamma \exp(\alpha Z)dt$. Thus $\Lambda_{0R}(t) = t$ and $\Lambda_{0D}(t) = 0.2t$.

Simulations were carried out for the settings described in Table 1. In all settings, except $I_e$ and $I_f$, $\gamma$ follows a gamma distribution with unit mean and variance $\theta$. In setting $I_e$, $\gamma$ follows a lognormal distribution with unit mean and variance 0.65; in setting $I_f$, $\gamma$ is generated as one tenth of a Poisson variable with mean 10. This has a variance of 0.10, which is close to independence. It is the same as one of the settings in the simulation study of Huang and Wang (2004). We increased the variance of the Poisson frailty on the suggestion of a reviewer, and obtained similar results with respect to bias (simulation not shown).

Table 2 presents results from the MR model and the NPF model for settings $I_a$, $I_b$, $I_e$ and $I_f$. In the first two settings, the frailty distribution is correctly specified for the MR model. In settings $I_e$ and $I_f$, the gamma frailty distribution is misspecified and the goal is to compare the results from the two models when the parametric assumption for $\gamma$ is violated in the MR model. The empirical bias and the empirical standard deviation of the estimators for the four settings are shown. The simulation study is based on 1000 simulated samples. Also in setting $I_a$ and $I_b$, the estimators of $\alpha$ and $\beta$ from the MR and NPF model both perform well in that the empirical bias is small for the both models. There seems to be some small bias in the estimation of $\theta$ in the MR model. The empirical standard errors for the estimates from the MR model are smaller than those from the NPF model suggesting that the MR model is a more efficient approach as expected. There is no evidence of bias in the estimation of $\alpha$ and $\beta$ for either model in any of the cases considered and, in particular, for the MR model when $\gamma$ does not follow a gamma distribution. This lack of sensitivity to mispecificaton of the frailty model is consistent with the simulation studies carried out by Glidden and Vittinghoff (2004) for frailty models for clustered survival data.

We also carried out a number of simulations to assess the performance of the proposed sandwich estimators. In this case, we considered different sample sizes ($n$=100 and 200), different coefficient values and different distributions for $\gamma$. The results are shown in the Tables

3 and 4. It can be seen that the variance estimates are accurate and the associated coverage probabilities are close to the nominal level of 0.95 for $\alpha$ and $\beta$. The coverage probabilities for $\theta$ seem to be slightly lower than the nominal level. When the frailty does not follow a gamma distribution as in $I_e$ and $I_f$, the variance estimators are very close to the empirical ones and the coverage probabilities of the intervals for $\alpha$ and $\beta$ still are close to the nominal level.

## 5. Application

We now fit the proposed MR model to the CANUSA study (Canada-U.S.A. peritoneal dialysis study group, 1998), a prospective cohort study of end-stage renal disease patients receiving peritoneal dialysis in Canada and the USA. Patients were enrolled and followed between September 1, 1990 and December 31, 1992. The recurrent event of interest is hospitalization and the terminal event is the failure of peritoneal dialysis, which occurs at the minimum of the time until death, technique failure or withdrawal from peritoneal dialysis.

A total of 680 patients were enrolled in this study; forty-two percent were female, 82% were Caucasian and the average age was 54. The number of hospitalizations per patient ranged from 0 to 23 with an average of about 1.7. About two-thirds of the patients were hospitalized at least once. Kidney transplantation was performed on 19.1% of the patients and was considered as random censoring. It is probably reasonable to treat kidney transplantation as random censoring since patients are not prioritized on the transplantation waiting list according to their disease severity. By the end of the study, 50% of the patients experienced the terminal event.

The covariates of interest include country (USA or Canada), age, gender, race, the causes of renal failure (polycystic kidney disease, diabetes, renal vascular, glomerulonephritis and other causes), baseline renal clearance measure, non-protein catabolic rate, percent lean body mass, serum albumin, subjective global assessment, cardiovascular disease and Karnofsky score.

The results of the analysis are summarized in Table 5. The frailty parameter was estimated

to be $\hat{\theta} = 0.990$ with an estimated standard error 0.12 ($P < 0.001$). According to this estimate, a patient who is known to fail from peritoneal dialysis at time t is expected to have almost twice as many hospitalizations for a patient who hasn't failed by time t. As one might expect, therefore, the rate of hospitalization is highly associated with the rate of failure from peritoneal dialysis. That is, patients with a high (low) hospitalization rate tend to have a larger (smaller) chance of failure from peritoneal dialysis.

After adjusting for the other covariates, the USA patients tend to have a higher rate of hospitalization than the Canadian patients ($P < 0.05$). However, no difference is found with respect to the failure rate. Percent lean body mass has a significant effect on both the failure rate and hospitalization rate; the higher percent of a patient's lean body mass is, the lower failure rate and hospitalization rate. Female patients have a lower failure rate, but gender is not found to be related to the hospitalization rate. In addition, patients whose renal failure is caused by diabetes, renal vascular disease failure or other have a higher rate of hospitalization than the patients whose renal failure is from glomerulonephritis; the cause of renal failure, however, does not seem to affect the failure rate. Having a high Karnofsky score decreases the estimated hospitalization rate but surprisingly not the failure rate.

We also carried out a naive analysis of the hospitalization rates, treating failure from peritoneal dialysis as a form of independent censoring. We fitted a marginal rate model and obtained robust sandwich type estimators as in Lin et al. (2000). The failure process was treated independently and analyzed using an ordinary Cox model. The results are shown in Table 6. Compared to the results in Table 5, and the coefficient the parameter estimates are smaller in magnitude under the naive model, which doesn't account for the dependence between the two processes. This attenuation seems to be the result of a positive correlation between the processes and the fact that the effect of each covariate on the death and hospitalization processes is in the same direction.

## 6. Discussion

In this article, we have developed and analyzed a shared frailty model for the recurrent events in the presence of a terminal event. The model is similar to the nonparametric frailty model proposed by Huang and Wang (2004) and the analysis leads to a notable increase of efficiency. Though a parametric assumption for the frailty is made in the MR model, simulation studies suggest that the model is robust to deviations from that assumption, at least in those cases considered. Time-dependent covariates can be easily handled in the model and the analysis we propose. Thus, departures from proportional hazards could be incorporated by introducing interactions with time. One advantage of our method is that the degree of association between the recurrent and terminal event processes can be estimated through the estimation of the variance of $\gamma$. The empirical variance of the $\hat{\gamma}_w$ in the NPF model would tend to over-estimate the frailty variance since it would incorporate both the frailty variance and the variation due to the underlying recurrent event process.

Liu et al. (2004) carried out maximum likelihood estimation in their frailty model by assuming that the recurrent events follow a nonhomogeneous Poisson process conditional on the frailty. A Monte Carlo EM algorithm with a Metropolis-Hasting sampler in the E-step is adapted to obtain the maximum likelihood estimator. The frailty effect is allowed to be different for the two processes and time-dependent covariates can be incorporated. As is often the case, however, the EM algorithm is slow to converge and the method is computationally much more intensive than the method proposed here. The estimation method of Liu et al. (2004) is based on a complete intensity model for recurrent events, and may therefore be expected to be sensitive to departures from this assumption. On the other hand, the proposed estimating equations combined with the use of the sandwich estimator should make our method more robust. Finally, the computational procedure converges relatively fast with the MR method, which makes bootstrapping a practical option for standard error estimation. Liu et al. (2004) also allow for different but related frailty effects on the recurrent and

the terminal event processes. Our methods could be similarly generalized to fit their model although some numerical integration methods would likely be needed. Alternative models that allow separate frailties could also be investigated.

In this paper, we have assumed nonparametric forms of $\Lambda_{0R}$ and $\Lambda_{0D}$. The large sample properties are therefore difficult to verify fully. Murphy (1995) and Parner (1998) studied the asymptotic properties of the shared gamma frailty model. They provide a general approach which could possibly establish the asymptotic properties of our proposed parameter estimators, but detailed arguments are still to be given. Simulation results suggest, however, that the proposed variance estimators are accurate and we expect the proposed method to be valid in many practical settings.

The estimation of $\theta$ in the proposed method requires the assumption that conditional on frailty, the recurrent event follows a nonhomogeneous Poisson process. It would be desirable to develop an estimation procedure which can relax this assumption.

## Acknowledgements

The authors wish to thank David Churchill and the Canada-U.S.A Peritoneal Dialysis Study Group for permission to analyze the CANUSA data.

## References

Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993). *Statistical models based on counting processes.* New York: Springer-Verlag.

Andersen, P. K., Klein, J., Knudsen, K. and Palacios, R. (1997). Estimation of variance in Cox's regression model with shared gamma frailties. *Biometrics* **53**, 1475-1484.

Canada-U.S.A. Peritoneal Dialysis Study Group. (1996). Adequacy of dialysis and nutrition in continuous peritoneal dialysis: Correlation with clinical outcomes. *Journal of the American Society of Nephrology* **7**, 198-207

Cook, R. J., Lawless, J. F. and Nadeau, J. C. (1996). Robust tests for treatment comparisons based on recurrent event responses. *Biometrics* **52**, 732-739.

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

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

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187-120.

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 event. *Statistica Sinica* **12**, 663-688.

Glidden DV and Vittinghoff E. (2004). Modelling clustered survival data from multicentre clinical trials. *Statistics in Medicine* **23**,369-88.

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

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

Lawless, J. F. and Nadeau, J. C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics* **37**, 18-68.

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

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

Murphy, S. (1995). Asymptotic theory for the frailty model. *Annals of Statistics* **23**, 182-198.

Parner, E. (1998). Asymptotic theory for the correlated Gamma-frailty model. *Annals of Statistics* **28**, 183-214.

Pepe M. S. and Cai J. (1993). Some graphical displays and marginal regression analyses of recurrent events and a terminating event. *Journal of the American Statistical Association* **16**, 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.

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

## Appendix

The partial derivatives of the joint estimating equation $U(\eta)$ are listed in this section. For a parameter $\phi$ (e.g. $\phi = \alpha$), we define $S_1^{(k)}(\phi, t) = n^{-1} \sum_{i=1}^n Y_i(t) w_i(t) Z_i^{\otimes k} \exp(\phi^T Z_i)$ and $S_2^{(k)}(\phi, t) = n^{-1} \sum_{i=1}^n Y_i(t) w_i(t)^2 Z_i^{\otimes k} \exp(\phi^T Z_i)$.

The partial derivatives of $U_1(\beta)$ are,

$$\begin{aligned}
\frac{\partial U_1(\beta)}{\partial \beta} &= \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{-S_1^{(2)}(\beta,u)}{S_1^{(0)}(\beta,u)} + \left( \frac{S_1^{(1)}(\beta,u)}{S_1^{(0)}(\beta,u)} \right)^{\otimes 2} \right\} dN_i^R(u), \\
\frac{\partial U_1(\beta)}{\partial \alpha} &= \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{S_2^{(2)}(\alpha+\beta,u)}{S_1^{(0)}(\beta,u)} - \frac{S_1^{(1)}(\beta,u)(S_2^{(1)}(\alpha+\beta,u))^T}{S_1^{(0)}(\beta,u)^2} \right\} \Lambda_{0D}(u)\theta dN_i^R(u), \\
\frac{\partial U_1(\beta)}{\partial \theta} &= \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{S_2^{(1)}(\alpha+\beta,u)}{S_1^{(0)}(\beta,u)} - \frac{S_1^{(1)}(\beta,u)S_2^{(0)}(\alpha+\beta,u)}{S_1^{(0)}(\beta,u)^2} \right\} \Lambda_{0D}(u) dN_i^R(u), \\
\frac{\partial U_1(\beta)}{\partial \lambda_{0Dj}} &= \sum_{i=1}^{n} \int_{t_{Dj}}^{\tau} \left\{ \frac{S_2^{(1)}(\alpha+\beta,u)}{S_1^{(0)}(\beta,u)} - \frac{S_1^{(1)}(\beta,u)S_2^{(0)}(\alpha+\beta,u)}{S_1^{(0)}(\beta,u)^2} \right\} \theta dN_i^R(u), \quad j = 1,\ldots,f.
\end{aligned}$$

The partial derivatives of $U_2(\alpha)$ are,

$$\begin{aligned}
\frac{\partial U_2(\alpha)}{\partial \alpha} &= \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{S_1^{(1)}(\alpha,u)^{\otimes 2} - \theta\Lambda_{0D}(u)S_1^{(1)}(\alpha,u)(S_2^{(1)}(2\alpha,u))^T}{S_1^{(0)}(\alpha,u)^2} \right. \\
&\quad \left. - \frac{S_1^{(2)}(\alpha,u) - \theta\Lambda_{0D}(u)S_2^{(2)}(2\alpha,u)}{S_1^{(0)}(\alpha,u)} \right\} dN_i^D(u), \\
\frac{\partial U_2(\alpha)}{\partial \theta} &= \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{S_2^{(1)}(2\alpha,u)}{S_1^{(0)}(\alpha,u)} - \frac{S_1^{(1)}(\alpha,u)S_2^{(0)}(2\alpha,u)}{S_1^{(0)}(\alpha,u)^2} \right\} \Lambda_{0D}(u) dN_i^D(u), \\
\frac{\partial U_2(\alpha)}{\partial \lambda_{0Dj}} &= \sum_{i=1}^{n} \int_{t_{Dj}}^{\tau} \left\{ \frac{S_2^{(1)}(2\alpha,u)}{S_1^{(0)}(\alpha,u)} - \frac{S_1^{(1)}(\alpha,u)S_2^{(0)}(2\alpha,u)}{S_1^{(0)}(\alpha,u)^2} \right\} \theta dN_i^D(u), \quad j = 1,\ldots f.
\end{aligned}$$

The partial derivatives of $U_3(\theta)$ are,

$$\begin{aligned}
\frac{\partial U_3(\theta)}{\partial \theta} &= \sum_{i=1}^{n} \frac{\partial^2 \log(L(O_i(\tau); Z_i)}{\partial^2 \theta}, \\
\frac{\partial U_3(\theta)}{\partial \alpha} &= \sum_{i=1}^{n} \frac{d_i Z_i}{r_i\theta^2 + d_i\theta^2 + \theta} - \frac{(m_i + \Delta_i + 1/\theta)\theta^2 d_i Z_i}{(\theta^2 r_i + \theta^2 d_i + \theta)^2}, \\
\frac{\partial U_3(\theta)}{\partial \beta} &= \sum_{i=1}^{n} \frac{r_i Z_i}{r_i\theta^2 + d_i\theta^2 + \theta} - \frac{(m_i + \Delta_i + 1/\theta)\theta^2 r_i Z_i}{(\theta^2 r_i + \theta^2 d_i(X_i) + \theta)^2}, \\
\frac{\partial U_3(\theta)}{\partial \lambda_{0Rj}} &= \sum_{i=1}^{n} Y_i(t_{Rj}) \left\{ \frac{\exp(\beta^T Z_i)}{\theta^2 r_i + \theta^2 d_i + \theta} - \frac{(m_i + \Delta_i + 1/\theta)\theta^2 \exp(\beta^T Z_i)}{(\theta^2 r_i + \theta^2 d_i + \theta)^2} \right\}, \quad j = 1,\ldots,m, \\
\frac{\partial U_3(\theta)}{\partial \lambda_{0Dj}} &= \sum_{i=1}^{n} Y_i(t_{Dj}) \left\{ \frac{\exp(\alpha^T Z_i)}{\theta^2 r_i + \theta^2 d_i + \theta} - \frac{(m_i + \Delta_i + 1/\theta)\theta^2 \exp(\alpha^T Z_i)}{(\theta^2 r_i + \theta^2 d_i + \theta)^2} \right\}, \quad j = 1,\ldots,f.
\end{aligned}$$

*The partial derivatives of the jth element of* $U_4(\lambda_{0R})$, $j = 1,\ldots,m$ *and* $U_5(\lambda_{0D})$, $j = 1,\ldots,f$ *are straightforward and are not shown here.*

**Table 1: Settings for the Simulation Study**

| | Setting I: n=200 | | | | | | Setting II: n=100 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Settings | $I_a$ | $I_b$ | $I_c$ | $I_d$ | $I_e$ | $I_f$ | $II_a$ | $II_b$ | $II_c$ | $II_d$ |
| $\alpha$ | 0.5 | 0.5 | 0 | 0 | 0.5 | 0.5 | 0.5 | 0.5 | 0 | 0 |
| $\beta$ | 0.5 | 0.5 | 0 | 0 | 0.5 | 0.5 | 0.5 | 0.5 | 0 | 0 |
| $\theta$ | 0.5 | 1 | 0.5 | 1 | NA | NA | 0.5 | 1 | 0.5 | 1 |
| $E[m_i]$ | 3.05 | 2.73 | 2.72 | 2.45 | 3.05 | 3.38 | 3.06 | 2.75 | 2.73 | 2.47 |
| $E[\Delta_i]$ | 61.2% | 54.5% | 54.3% | 49% | 61.1% | 67.4% | 61.2% | 54.8 % | 54.3% | 49% |

$E[m_i]$: average number of recurrent events per subject;

$E[\Delta_i]$: average percentage of subjects who experience the terminal event

**Table 2 : Comparison of MR Model with NPF Model Based on 1000 Simulated Samples**

| | Setting $I_a$: $\gamma \sim \Gamma, \theta = 0.5$ | | | | Setting $I_b$: $\gamma \sim \Gamma, \theta = 1$ | | | |
|---|---|---|---|---|---|---|---|---|
| | MR Model | | H&W Model | | MR Model | | H&W Model | |
| Parameter | Bias | ESE | Bias | ESE | Bias | ESE | Bias | ESE |
| $\beta = 0.5$ | -0.003 | 0.152 | 0.0052 | 0.201 | -0.005 | 0.213 | 0.0037 | 0.249 |
| $\alpha = 0.5$ | 0.011 | 0.232 | 0.014 | 0.262 | 0.002 | 0.278 | 0.006 | 0.293 |
| $\theta$ | -0.022 | 0.087 | N/A | N/A | -0.031 | 0.147 | N/A | N/A |

| | Setting $I_e$: $\gamma \sim$ log-Normal | | | | Setting $I_f$: $\gamma \sim$ Poisson | | | |
|---|---|---|---|---|---|---|---|---|
| | MR Model | | H&W Model | | MR Model | | H&W Model | |
| Parameter | Bias | ESE | Bias | ESE | Bias | ESE | Bias | ESE |
| $\beta = 0.5$ | -0.010 | 0.145 | 0.003 | 0.248 | 0.005 | 0.103 | 0.003 | 0.173 |
| $\alpha = 0.5$ | -0.002 | 0.213 | 0.008 | 0.275 | 0.002 | 0.189 | -0.000 | 0.225 |

Bias: Empirical Bias; ESE: Empirical Standard Error

**Table 3: Simulation Results for the MR Model under Settings $I_a$ to $I_f$**

**Based on 1000 Simulated Sample**

*Setting I (n=200)*

| | Setting $I_a$ | | | | Setting $I_b$ | | | |
|---|---|---|---|---|---|---|---|---|
| MR model | Bias | CSE | ESE | 95% C.P. | Bias | CSE | ESE | 95% C.P. |
| $\hat{\beta}$ | -0.004 | 0.154 | 0.149 | 0.956 | 0.004 | 0.209 | 0.220 | 0.937 |
| $\hat{\alpha}$ | 0.002 | 0.227 | 0.229 | 0.943 | 0.012 | 0.275 | 0.275 | 0.948 |
| $\hat{\theta}$ | -0.01 | 0.085 | 0.091 | 0.916 | -0.022 | 0.145 | 0.154 | 0.925 |
| | Setting $I_c$ | | | | Setting $I_d$ | | | |
| MR model | Bias | CSE | ESE | 95% C.P. | Bias | CSE | ESE | 95% C.P. |
| $\hat{\beta}$ | -0.004 | 0.152 | 0.160 | 0.935 | -0.013 | 0.204 | 0.206 | 0.946 |
| $\hat{\alpha}$ | -0.011 | 0.231 | 0.242 | 0.94 | -0.007 | 0.274 | 0.277 | 0.95 |
| $\hat{\theta}$ | -0.012 | 0.088 | 0.091 | 0.934 | -0.022 | 0.148 | 0.154 | 0.919 |
| | Setting $I_e$ | | | | Setting $I_f$ | | | |
| MR model | Bias | ESE | CSE | 95% CI | Bias | ESE | CSE | 95%CI |
| $\beta = 0.5$ | -0.010 | 0.145 | 0.147 | 0.954 | 0.005 | 0.103 | 0.098 | 0.936 |
| $\alpha = 0.5$ | -0.002 | 0.213 | 0.222 | 0.959 | 0.002 | 0.189 | 0.185 | 0.943 |

CSE: mean of calculated standard error; ESE: empirical standard error; 95% C.P.: 95% confidence interval coverage probability

**Table 4: Simulation Results for the MR Model under Settings $II_a$ to $II_d$ Based on 1000 Simulated Sample**

*Setting II (n=100)*

| | Setting $II_a$ | | | | Setting $II_b$ | | | |
|---|---|---|---|---|---|---|---|---|
| MR model | Bias | CSE | ESE | 95% C.P. | Bias | CSE | ESE | 95% C.P. |
| $\hat{\beta}$ | -0.006 | 0.214 | 0.224 | 0.946 | 0.019 | 0.291 | 0.304 | 0.94 |
| $\hat{\alpha}$ | 0.015 | 0.321 | 0.333 | 0.936 | 0.034 | 0.390 | 0.400 | 0.958 |
| $\hat{\theta}$ | -0.026 | 0.117 | 0.121 | 0.890 | -0.046 | 0.202 | 0.207 | 0.902 |
| | Setting $II_c$ | | | | Setting $II_d$ | | | |
| MR model | Bias | CSE | ESE | 95% C.P. | Bias | CSE | ESE | 95% C.P. |
| $\hat{\beta}$ | -0.011 | 0.212 | 0.220 | 0.939 | -0.015 | 0.281 | 0.275 | 0.953 |
| $\hat{\alpha}$ | -0.007 | 0.327 | 0.353 | 0.935 | -0.005 | 0.387 | 0.385 | 0.955 |
| $\hat{\theta}$ | -0.020 | 0.123 | 0.127 | 0.903 | -0.06 | 0.204 | 0.216 | 0.891 |

CSE: mean of calculated standard error; ESE: Empirical standard error; 95% C.P.: 95% confidence interval coverage probability

**Table 5: Analysis of CANUSA Study**

| Covariate | Peritoneal Dialysis Failure $\hat{\alpha}$ | $\widehat{SE}$ | $P$ | Hospitalization $\hat{\beta}$ | $\widehat{SE}$ | $P$ |
|---|---|---|---|---|---|---|
| Country | | | | | | |
| *USA* | 0.316 | 0.233 | 0.175 | 0.33 | 0.168 | 0.0499 |
| *Canada* | 0 | . | . | 0 | . | . |
| Gender | | | | | | |
| *Female* | -0.454 | 0.173 | 0.009 | -0.097 | 0.13 | 0.456 |
| *Male* | 0 | . | . | 0 | . | . |
| Race | | | | | | |
| *Non-Caucasian* | -0.365 | 0.225 | 0.104 | -0.261 | 0.153 | 0.089 |
| *Caucasian* | 0 | . | . | 0 | . | . |
| Causes of Renal Failure | | | | | | |
| *Polycystic kidney disease* | 0.330 | 0.352 | 0.349 | 0.199 | 0.270 | 0.463 |
| *Diabetes* | 0.160 | 0.248 | 0.518 | 0.818 | 0.198 | $< 0.001$ |
| *Vascular* | 0.388 | 0.358 | 0.278 | 0.803 | 0.283 | 0.005 |
| *Other* | 0.226 | 0.242 | 0.350 | 0.543 | 0.193 | 0.005 |
| *Glomerulonephritis* | 0 | . | . | 0 | . | . |
| Age (per year) | -0.006 | 0.006 | 0.365 | -0.007 | 0.005 | 0.166 |
| Non-Protein catabolic rate | 0.096 | 0.367 | 0.794 | 0.382 | 0.269 | 0.155 |
| Percent lean body mass | -0.026 | 0.008 | 0.001 | -0.014 | 0.005 | 0.010 |
| Subjective global assessment | -0.058 | 0.056 | 0.297 | -0.066 | 0.041 | 0.107 |
| Cardiovascular disease | 0.141 | 0.169 | 0.402 | 0.162 | 0.130 | 0.213 |
| Karofsky score | -0.081 | 0.069 | 0.242 | -0.114 | 0.053 | 0.031 |
| Baseline renal clearance measure (per 10 units) | 0.122 | 1.081 | 0.910 | -0.143 | 0.861 | 0.868 |
| Serum albumin (per 10 gram per liter) | -0.311 | 0.174 | 0.073 | -0.238 | 0.122 | 0.051 |

**Table 6: Analysis of CANUSA Study**

**(Naive Method)**

| Covariate | Peritoneal Dialysis Failure $\hat{\alpha}$ | $\widehat{SE}$ | $P$ | Hospitalization $\hat{\beta}$ | $\widehat{SE}$ | $P$ |
|---|---|---|---|---|---|---|
| Country | | | | | | |
| *USA* | 0.225 | 0.158 | 0.160 | 0.254 | 0.127 | 0.045 |
| *Canada* | 0 | . | . | 0 | . | . |
| Gender | | | | | | |
| *Female* | -0.331 | 0.123 | 0.007 | 0.010 | 0.103 | 0.920 |
| *Male* | 0 | . | . | 0 | . | . |
| Race | | | | | | |
| *Non-Caucasian* | -0.258 | 0.159 | 0.110 | -0.164 | 0.117 | 0.162 |
| *Caucasian* | 0 | . | . | 0 | . | . |
| Causes of Renal Failure | | | | | | |
| *Polycystic kidney disease* | 0.242 | 0.251 | 0.340 | 0.123 | 0.215 | 0.568 |
| *Diabetes* | 0.125 | 0.170 | 0.46 | 0.789 | 0.163 | $< 0.001$ |
| *Vascular* | 0.285 | 0.241 | 0.24 | 0.789 | 0.218 | 0.001 |
| *Other* | 0.174 | 0.168 | 0.300 | 0.499 | 0.160 | 0.002 |
| *Glomerulonephritis* | 0 | . | . | 0 | . | . |
| Age (per year) | -0.004 | 0.005 | 0.400 | -0.006 | 0.004 | 0.192 |
| Non-Protein catabolic rate | 0.085 | 0.265 | 0.750 | 0.375 | 0.224 | 0.094 |
| Percent lean body mass | -0.019 | 0.006 | $< 0.001$ | -0.007 | 0.005 | 0.100 |
| Subjective global assessment | -0.042 | 0.041 | 0.300 | -0.052 | 0.032 | 0.111 |
| Cardiovascular disease | 0.092 | 0.121 | 0.450 | 0.116 | 0.101 | 0.251 |
| Karofsky score | -0.056 | 0.047 | 0.230 | -0.091 | 0.040 | 0.023 |
| Baseline renal clearance measure (per 10 units) | 0.090 | 0.877 | 0.920 | -0.168 | 0.782 | 0.830 |
| Serum albumin (per 10 gram per liter) | -0.215 | 0.119 | 0.071 | -0.155 | 0.096 | 0.107 |