Current Methods for Recurrent Events Data with Dependent Termination:

A Bayesian Perspective

Debajyoti Sinha,
Department of Statistics, Florida State University

Tapabrata Maiti,
Department of Statistics, Iowa State University

Joseph G. Ibrahim, and
Department of Biostatistics, University of North Carolina

Bichun Ouyang
Department of Biostatistics, Medical University of South Carolina

Summary

There has been a recent surge of interest in modeling and methods for analyzing recurrent events data with risk of termination dependent on the history of the recurrent events. To aid the future users in understanding the implications of modeling assumptions and modeling properties, we review the state of the art statistical methods and present novel theoretical properties, identifiability results and practical consequences of key modeling assumptions of several fully specified stochastic models. After introducing stochastic models with noninformative termination process, we focus on a class of models which allows both negative and positive association between the risk of termination and the rate of recurrent events via a *frailty* variable. We also discuss the relationship as well as the major differences between these models in terms of their motivations and physical interpretations. We discuss associated Bayesian methods based on Markov chain Monte Carlo tools, and novel model diagnostic tools to perform inference based on fully specified models. We demonstrate the usefulness of current methodology through an analysis of a data set from a clinical trial. In conclusion, we explore possible future extensions and limitations of the methodology.

Keywords

Gibbs sampling; Frailty; Mixed Poisson process; M-site model; Semiparametric Bayes

1 INTRODUCTION

Modeling and analysis of recurrent event times data in many applications range from clinical trials, engineering product testing, reliability analysis of repairable systems and so forth. References include Kuo and Yang (1995,1996) in software reliability and reliability analysis of repairable systems, Oakes (1992),Wei and Glidden (1997),Scheike (2000),Lin et al. (2000),Wang et al. (2001) and Miloslvsky et al. (2002, 2004) in clinical trials and medical applications. In this article, we investigate the theoretical properties and implications of key modeling assumptions of recently developed fully specified stochastic models of recurrent events data with dependent termination. We will demonstrate how some of the important classes of models, particularly with associated Bayesian analysis tools, can be used analysis of such data including model selection and diagnostics for adequacy of modeling assumptions.

To illustrate with a typical data example, we consider a study involving heart-transplant patients who are at risk of recurrent events of transplant rejection (treated effectively by drug therapy) as well as the non-recurrent terminal event of death. The # of recurrent rejection events to a patient by time t is $N(t) = \Sigma_j \, 1_{[Y_j \leq t]}$, where $Y_1 < Y_2 < \cdots$ are the ordered rejection times. In this situation, the non-recurrent event of death terminates/censores the observation of the recurrent event process $\{N(t)\}$, and the termination time T is not determined via the design of the study. We make a practical distinction between the terms "termination" and "censoring". The censoring due to loss of follow-up, end of study, etc., when they occur before death, is treated as a case of typical *noninformative censoring* of the termination/death time T. In general, $N(C)$ is informative of $N(T)$ even when a termination time T is non-informatively censored at C.

A simplifying assumption for modeling recurrent events data subject to termination (e.g. Sinha, 1993, Hougaard, 2000) is that given the history of $X(t) = \{x(s) : s < t\}$ of right-predictable covariates $x(s)$ and the history $\mathcal{N}(t-) = \{N(s) : s < t\}$ of recurrent events of the subject, the risk of termination $h_T(t|\mathcal{N}(t-), X(t))$ at any time point t is independent of the history $\mathcal{N}(t-) = \{N(s) : s < t\}$, that is,

$$h_T(t|\mathcal{N}(t-), X(t)) = h_T(t|X(t)). \tag{1.1}$$

In practice, one usually uses ordinary survival analysis methods for the termination time T using $\{N_*(t-), x(t)\}$ as time-dependent covariates, where $N_*(u)$ is a prespecified vector of function of $\mathcal{N}(u)$. If there are no significant effects of $N_*(t-)$ found, then the data on recurrent events can be analyzed under the assumption of (1.1). However, assumption (1.1) may not be valid in many practical examples. For example, in a organ transplant trial, the intensity of recurrent graft rejection episodes is considered an indicator of the unobservable health-state of the patient, whereas the risk of termination due to death, an important outcome by itself, is also very much related to the unobservable health-state. For an organ-transplant trial, both types of events are negative outcomes. For a product testing study of repairable systems, each system may experience recurrent failures leading to detection of faults (recurrent negative outcomes signifying unobservable quality of the product) until it is terminated (positive outcome related to unobservable quality) with the final release of the product.

In the frequentist literature, most of the statistical methods for recurrent events data, with noninformative termination, use partially specified models for robust estimation of the regression parameters associated with either the marginal hazard of Y_j (the j-th recurrent event of the subject) or $E[dN(t)]$, where $dN(t) = N[t, t + dt)$. These models are defined conditional only on a subset $\mathcal{H}^*(t)$ of the history $\{\mathcal{N}(t-), X(t)\}$ (e.g. Sun and Wei, 2000, Wang, Qin and Chang, 2001). The interpretations of the regression parameters are strictly restricted to the subset $\mathcal{H}^*(t)$ of the history the model is conditionally defined on. For example, $E[dN(t)|\mathcal{N}(t-), X(t)]$, $P\left[Y_j \in [t, t+dt)|\mathcal{N}(t-) = j - 1; X(t)\right]$ and $E[dN(t)|X(t)]$ can be very different from each other. It is not possible to automatically derive the regression parameters of one partially specified model from those of another partially specified model. We refer interested readers to the detailed and excellent reviews by Miloslavsky et al. (2002, 2004) regarding frquentist inference for recurrent events data under noninformative termination using different partially specified models. The same problems arise when partially specified models for gap times $U_j = Y_j - Y_{J-1}$ are used conditional on $\mathcal{H}^*_{j-1}$, a subset of the history up to $(j - 1)$-th recurrent event times Y_j. We also loose a sense of the progression of the process over calendar time from origin when we use models based on gap times. For some studies, it is desirable to have a fully specified unifying model such that conclusions about different components conditional on different subsets of observed history are compatible with each other. This helps

to fully comprehend the data generation process of the experiment at hand. A fully specified model can be also used for prediction, model diagnostics, simultaneous inference on multiple parameters, etc.

In Section 2, we begin with different fully specified stochastic models for recurrent events data with noninformative termination. We discuss the physical interpretations, new stochastic formulations and the relationships among the following models of recurrent events data: the frailty model (Oakes (1992), Sinha 1993 and Hougaard (2000)), the M-site model (Gail et al., 1980) and the time-varying correlated frailty model (Herdersen and Shimakura, 2003). Despite the recent interest and research activity in Bayesian survival analysis (e.g., Sinha and Dey (1998), Walker, et al. (1999), Dey, Mueller and Sinha (1998) and Ibrahim, Chen and Sinha (2001)), the current Bayesian literature on recurrent events data is very limited. One goal of this paper is to review the current state of Bayesian methodologies for recurrent events data and demonstrate the use and advantages of the Bayesian paradigm for such problems.

In later Sections, we examine extensions of these recurrent events models to recurrent events with informative termination. The extensive simulation study of Miloslavsky et al. (2002) demonstrates the issue of bias in the regression estimates of the marginal rate and marginal intensity functions when informative termination is ignored during estimation. We focus on fully specified stochastic models where the recurrent events and as well as the termination are considered as important responses. In presence of two different types of event processes of interest, to comprehend the effects of covariates and the data generation process, we need to evaluate the effects of covariates separately on two processes as well as the effect of the joint process. For example, in a transplant study, we are interested in $h_T(t|\mathcal{N}(t-);\mathcal{X}(t))$ as well as $E[dN(t)|\mathcal{H}^*(t-);\mathcal{X}(t)]$ when $\mathcal{H}^*(t)$ can be equal to either $(\mathcal{N}(t-), T \geq t)$ or $\mathcal{N}(t-)$. A careful examination of the regressions functions in all three functions can only reveal the true nature of the effects of covariates on two types of outcomes.

Methods of estimation in the presence of informative termination are presented and reviewed by Miloslavsky et al. (2002,2004) using the CAR (Coarsening At Random) assumption, (e.g., Heitjan and Rubin, 1991; Gill et al., 1997; Robbins, 1993), which implies that the risk of termination h_T has the following property:

$$h_T(t|\mathcal{H}(\tau)) = h_T(t|\mathcal{H}(t-)),$$

(1.2)

where $\mathcal{H}(t-) = \{\mathcal{N}(t-), \mathcal{X}(t)\}$, and $(0, \tau]$ for $\tau > t$ is either the study-interval or the time-interval of interest. When termination is, say, the event of death for the transplant example, it is not possible to either define or conceptualize a hazard $h_T(t)$ of death at $t < \tau$ given the history of recurrent events and covariate at an interval $(t+, \tau]$ beyond time t. One can only observe a history of recurrent rejections in $(t, \tau]$ only when the patient is alive beyond time t. The CAR assumption is difficult to understand and verify unless a bigger class of joint stochastic models of $\{N(\cdot), T\}$ is defined and then the validity of CAR assumption is verified. The CAR assumption is applicable when termination event is not of interest and it is precisely treated as a non-informative censoring. In such a case, the focus is on understanding $dN(t)$ given $(\mathcal{N}(t-);\mathcal{X}(t))$ and we only need to take care of inferential problems raised by the observed data under informative censoring. Even after having the CAR assumption, one needs to correctly specify $h_T(t|\mathcal{N}(t-);\mathcal{X}(t))$ and $E[dN^+(t)|\mathcal{N}^+(t-);\mathcal{X}(t)]$, where $N^+(t) = N(Min(t, T))$. To obtain the regression parameters of $E[dN(t)|\mathcal{H}^+(t-);\mathcal{X}(t)]$ from $E[dN^+(t)|\mathcal{N}^+(t-);\mathcal{X}(t)]$, we need additional set of assumptions (beyond CAR) whose validity cannot be evaluated solely from observed data. We will investigate what key assumptions of fully-specified stochastic models of the joint process can give rise to CAR and related assumptions.

Recent work by Lancaster and Intrator (1998), Sinha and Maiti (2004), and Liu, Wolfe and Huang (2004) offer several new fully specified models for recurrent events data with dependent termination describing $\{dN(t), d1_{[T \leq t]}\}$ sequentially unfolding over time. However, no theoretical investigations of relationships among these models and their properties have been examined in current literature. The relationships (or the lack there of) these models with CAR assumption based models have not been investigated yet. In general, a fully specified stochastic model may not even satisfy the CAR assumption. In this work, we consider novel theoretical explorations of three major aspects of modeling: (1) theoretical properties and practical consequences of key modeling assumptions on the observable quantities; (2) relationships among different models; and (3) identifiability of model parameters and distributional assumptions from observable data. For each model with dependent termination, we explore in detail different consequences resulting from modeling assumptions because these important consequences may not be reasonable in various biological and engineering applications. For example, we show that for some of the existing models, when termination times are subject to noninformative right-censoring, the induced censoring of the number of recurrent events until termination is bound to be noninformative (contrary to what we expect to see in most real-life data examples). We investigate the relationships and comparisons of related models existing in the recurrent events literature to help the user to make informed selection of the models most appropriate for the study/experiment in hand.

In Section 3, we discuss a comparison of a full Bayesian analysis (with the help of sampling based algorithms) versus a semiparametric likelihood-based analysis. A full Bayesian approach is one of the natural choices of inference based on fully specified stochastic models, and we argue that it is the most appropriate one in this case. In Section 4, we present details regarding the Bayesian computational tools used in this problem. Then, a typical Bayesian analysis of recurrent events data with informative termination is illustrated with the analysis of recurrent events data from a heart transplant study.

One important component of any model-based methodology is the development of appropriate tools for evaluating model adequacy, outlier detection and influential observations. In Section 5, we demonstrate the Bayesian case-influence diagnostics and residual analysis appropriate for the proposed models. Recurrent events data with a terminating event is also related to various other applications such as cumulative medical costs data (e.g., Lin, et al., 1997; Etzioni et al., 1999), cumulative repair/warranty costs data (e.g., Nelson, 1995), and quality-adjusted survival data (e.g., Zhao and Tsiatis, 1997). In Section 6, we explore the relationship between recurrent events data and cumulative cost process data. We conclude our article with some closing remarks and exploring further extensions of our models and methods.

2 MODELS

2.1 MPP Models Under Non-informative Termination

We first describe models for $\{N_i(t)\}$ under the non-informative termination assumption of (1.1) (e.g., Oakes, 1992;Sinha, 1993). The Mixed Poisson Process model (MPP model) for recurrent events data in Oakes (1992) and Sinha (1993) assumes that for subject $i = 1, \cdots, n$ with $T_i > t$, the probability of a new recurrent event in $[t, t + dt)$ is given by

$$P[dN_i(t) = 1 | w_i, \mathcal{N}_i(t-); \mathcal{X}_i(t)] \simeq w_i e^{\beta_1 x_i(t)} d\mu_0(t),$$

(2.1)

and $P[dN_i(t) > 1 | w_i; \mathcal{X}_i(t)] \simeq o(dt)$, where $\mu_0(t)$ is the baseline cumulative intensity and β_1 is the regression parameter of the covariate effect on the conditional intensity. The time-constant frailty random effects, $w_1, \cdots, w_n$, are assumed to be i.i.d. with common cdf $F(w|\eta)$ with $E_F[W_i] = 1$ (to assure identifiability of the model parameters). Hougaard (2000) calls this the

shared frailty model. The use of a finite mean frailty distribution is warranted here to ensure the expected number of events within any finite time interval to be finite. For simplicity of notation, we will present our models and the results assuming fixed time-constant covariate x $(t) = x$ throughout. Extending the following discussion and results to accommodate time-dependent $x(t)$ is quite straightforward. The results which are not valid for time-dependent predictable covariate $x(t)$ will be mentioned separately in this paper.

In practical applications, we only observe the process $\{N_+(t) = N(t)1_{[T \geq t]}\}$, and the conditional intensity of $\{N_+(t)\}$ given $(w_i, \mathcal{N}_i(t-), T_i \geq t; \mathcal{X}_i(t))$ should be defined, however, with the non-informative termination assumption of (1.1), we can use the simplified conditional intensity of (2.1). In Section 6, we discuss the extension of the MPP model using a non-proportional intensity for $N(t)$. One main consequence of the mixed Poisson process assumption in the MPP model is $P[dN_i(t) > 1|x_i] = o(dt)$, indicating that the MPP model is not suitable when the recurrent events within a subject tend to cluster together within a small time-interval. The number $N(t)$ carries all the predictive information of the history $\mathcal{N}(t)$ for the conditional risk of a new event (Oakes, 1992), because

$$P[dN(t)=1|\mathcal{N}(t-);x] = P[dN(t)=1|N(t-);x] \simeq \lambda_N(t|N(t-);x)\,dt, \tag{2.2}$$

where x is the vector of fixed covariates. The expression in (2.2) shows that the MPP model is a subclass of the dynamic covariate model discussed by Aalen et al. (2004) and Fosen et al. (2006), where the model is specified via $P[dN(t)=1|\mathcal{N}^*(t-);x]$, where $\mathcal{N}^*(t)$ is a pre-specified subset of $\mathcal{N}(t)$. There are many choices for a dynamic covariate model, and only a careful investigation of the underlying structure (performed in this paper) can give the user some understanding of the implications of the assumptions of a particular dynamic covariate model.

Even when $\mu_0(t)$ is left unspecified, the frailty distribution $F(w|\eta)$ is identifiable from the observed recurrent events data (Oakes, 1992). This implies that the observed recurrent events data can be used via appropriate goodness-of-fit plots for identifying a suitable $F(w|\eta)$ among various choices of frailty distributions. In susequent Sections, while discussing the recurrent events model under dependent termination, we will address whether the frailty density is identifiable from the observable data. Use of a Gamma frailty distribution with $F(w|\eta)$ assumed to be a $Ga(\eta, \eta)$ cdf with variance η^{-1}, is fairly commonplace due to its convenience and theoretical properties. However, choices other than a Gamma frailty can be considered. The frailty random effect accommodates heterogeneity among subjects due to their variable proneness to the recurrent non-fatal events via the unknown variance η^{-1} of the Gamma frailty .

2.2 M-site Models

The popular M-site model of carcinogenesis (e.g., Whittemore and Keller (1978) and Gail et al. (1980)) is based on the premise that recurrent detectable tumors (or cancers) in a finite interval of interest $(0, \tau)$ arise from proliferation of clones derived from each of M progenitor cells or tumor-sites within a patient with i.i.d. promotion times $Z_1, \cdots, Z_M$ (times to recurrent detectable tumors in cancer context) with common density $g(t)$. When the unknown and unobservable M_i given the frailty w_i and time-constant x_i has a Poisson distribution with mean $E[M_i|w_i, T_i; x_i] = \mu_i = w_i \exp(\beta_1 x_i)\gamma 0$, the resulting process $\{N_i(t)\}$ in $(0, \tau)$ is same as the MPP model of (2.1) with $d\mu_0(t) = \gamma_0 g(t)dt$ (Oakes, 1992). This result gives a strong biological justification for the use of the MPP model in cancer. In product testing, M_i can represent the unknown number of undetected faults/bugs in product i at the onset and i.i.d. $Z_{i1}, \cdots, Z_{iM_i}$ are the detection times of these faults.

A new modification of the M-site model assumes that $[M_i|T_i]$ is Poisson with mean $\mu_i = \gamma_0 \exp(\beta_1 x_i)$ (free of the patient-specific random effects) and the patient specific piecewise

exponential promotion time density $g_i(t) = U_{ij} e^{-U_{ij}}$ for $t \in (a_{j-1}, a_j)$ with a pre-specified grid $0 = a_0 < a_1 < \cdots < a_J = \tau$. When the vector of patient-specific rates $\Lambda_i = (U_{i1}, \cdots, U_{iJ})$ follows a multivariate Gamma density of Griffiths (1984), we get the serially correlated Gamma frailty model of Henderson and Shimakura (2003). We call this model the HS M-site model. This model can be used when we believe that the effect of among subjects heterogeneity on the rates $(U_{i1}, \cdots, U_{iJ})$ of latent activation times is more than its effect on number of latent tumors/faults M_i. The original work of Henderson and Shimakura uses a composite likelihood method to analyze only discrete panel count data (instead of continuous time recurrent events data) under the HS M-site model. Extensions of the composite likelihood method to analyze data based on actual recurrent event times are not trivial. To analyze continuously monitored recurrent events data with this kind of time-varying frailty model, we recommend using a multivariate log-normal density for Λ_i, instead of the multivariate gamma density of the HS M-site model. Our major concern is non-identifiability of the density $g(t|\Lambda_i)$ and the multivariate density of Λ_i from recurrent events data when both densities are left unspecified.

2.3 L-I Model for Informative Termination

In many data examples and studies, the non-informative termination assumption of (1.1) will not be valid. Lancaster and Intrator (1998), in what we call here the L-I model, formulated the association between $\mathcal{N}_i(t)$ and the risk termination at time t via the shared frailty w_i and replaced the assumption of (1.1) with the assumption

$$h_T(t|\mathcal{N}_i(t-), w_i; x_i) \propto w_i \exp(\beta_2 x_i) \mu_0'(t),$$

(2.3)

where β_2 is the regression parameter of the covariate effect on the conditional risk of termination, and the baseline intensity $\mu_0'(t)$ is the derivative of $\mu_0(t)$. The parametric L-I model allows only a certain type of association between $\mathcal{N}(t)$ and risk of termination T, and has some important consequences of its modeling assumptions (discussed below).

We now present a way of viewing the L-I model as a process that unfolds over time. This alternative description and a series of theoretical results will facilitate a better understanding of the consequences of the key assumptions. Let us assume that all events are generated via a latent nonhomogeneous mixed Poisson process (Ross, 1983) $N^*(\cdot)$ with conditional (given frailty w) intensity $w\mu^*(t; x)$, where $\mu^*(t;x) \propto \left[\exp(\beta_1 x) + \gamma \exp(\beta_2 x)\right] \mu_0'(t)$. When $dN^*(\cdot) = 1$, there are two possibilities for the new event-either a non-fatal event (Type I) with probability $P[dN(t) = 1, T > t + dt|dN^*(t) = 1, T > t, w; x] = \phi(x)$ or a termination event (Type II) with probability $P[T < t + dt, dN(t) = 0|dN^*(t) = 1, T > t, w; x] = 1 - \phi(x)$, where $0 \leq \phi(x) = \exp(\beta_1 x)\backslash[\exp(\beta_1 x) + \gamma \exp(\beta_2 x)] \leq 1$ for some $\gamma \geq 0$. Because the recurrent events and termination are essentially coming out of a single event process N^*, the L-I model implies that any subject with a history of higher than usual rate of $N(t)$ will be at a higher than usual risk of termination. As an important consequence of being an extension of MPP model, the L-I model has properties similar to (2.2). More precisely, following calculations similar to Oakes (1992), we can show that

$$h_T(t|\mathcal{N}(t-);x) = h_T(t|N(t-);x) = (1 - \phi) \frac{E_W\left[\exp\left\{-W_{\mu^*}(t;x)\right\} W^{N(t-)+1}\right]}{E_W\left[\exp\left\{-W_{\mu^*}(t;x)\right\} W^{N(t-)}\right]} \mu^*(t;x).$$

(2.4)

The expectations denoted by E_W in (2.4) are taken with respect to the frailty distribution $F(w|\eta)$. We can also show that $\lambda_N(t|\mathcal{N}(t-), T \geq t; x) = \lambda_N(t|N(t-), T \geq t; x)$, and

$$\lambda_N (t|N(t-),T \geq t;x) = \phi \frac{E_W \left[\exp \{-W\mu^*(t;x)\} W^{N(t-)+1} \right]}{E_W \left[\exp \{-W\mu^*(t;x)\} W^{N(t-)} \right]} \mu^*(t;x).$$

(2.5)

These results imply that $\lambda_N (t|N(t-),T \geq t;x) \propto h_T (t|N(t-);x)$, i.e., given the history $\{N(t-),T \geq t;X(t)\}$, the risk of a new recurrent event is proportional to the risk of termination and both risks depend on the observed history only via number of past events $N(t-)$. When $F(w|\eta)$ is $Ga(\eta, \eta)$, then $h_T (t|N(t-);x) = (1 - \phi(x))(N(t-)+\eta)^{-1}d\mu^*(t;x)$ and $\lambda_N (t|N(t-),T \geq t;x) = \phi(x)(N(t-)+\eta)^{-1}d\mu^*(t;x)$. The L-I model can be also expressed as an extension of the M-site model with the additional assumption that $N^*(t)$ within interval $(0, \tau]$ is generated via the extended M-site model of the promotion of a random number (Poisson distributed) of sites. The following consequences of the L-I model are less obvious. However, they have a major practical significance.

Theorem 2.1—Under the L-I model, the number of recurrent events before termination N(T) is geometric, that is, P[N(T) = k; x] = {ϕ(x)}^{k}{1 - ϕ(x)} for k = 0, 1, · · ·.

The proof of Theorem (2.1) is given in Appendix I. It is important to note that the marginal distribution of $N(T)$ is free of $\mu^*(t)$. In practice, it is often a reasonable assumption to allow the termination time T to be subjected to the possibility of noninformative right censoring at C, that is, , $P[T > t + u|T > t, C > t; x] = P[T > t + u|T > t; x]$ for all $t, u > 0$ (Cox and Oakes, 1984). In the transplant study, the the termination time of death is subjected to on-informative censoring due to drop-out. Theorem (2.1) implies that the L-I model is a very special class of models for recurrent events where the induced censoring of $N(T)$ by $N(C)$ is also non-informative, since $P[N(T) > u + v |N(C) > u, T > C] = P[N(T) > u + v |N(T) > u] = (1 - \phi)^v$. The proof of this result follows directly from Theorem (2.1). This restrictive property of the L-I model limits its application in most real data setting. As it is argued by Lin et al. (1997), the induced censoring of $N(T)$ by $N(C)$ should not be noninformative in general. We mention that these results are valid irrespective of the form of $F(w|\eta)$. An extension of the L-I model suggested by a reviewer assumes time-varying $\phi_t(x) = \exp(\beta_1 x)/[\exp(\beta_1 x) + \gamma(t) \exp(\beta_2 x)]$, where $\gamma(t)$ is a function of time t. In this case, the assumption of (2.3) and the result of Theorem (2.1) are not valid. The resulting model is a special case of another competing model introduced in a subsequent section.

One important question is whether the unknown and unspecified frailty distribution $F(w|\eta)$ is identifiable from the observed data under the L-I model. In absence of the covariate x, we can show that

$$\frac{h_T (t|N(t-)=1)}{h_T (t|N(t-)=0)} = \frac{\psi''(\mu^*(t)) \psi(\mu^*(t))}{\{\psi'(\mu^*(t))\}^2},$$

(2.6)

where $\psi(u) = E[\exp(-uW)]$ is the Laplace transform of $F(w|\eta)$ with first and second derivatives ψ' and ψ''. Writing this ratio as a function $\zeta(v)$ of $v = P[T > t, N(t) = 0] = \psi(\mu^*(t))$, we get

$$\zeta(v) = - \frac{v\chi''(v)}{\chi'(v)},$$

(2.7)

where χ is the inverse function of the Laplace transform $\psi(u)$. The solution of the above differential equation is $\{\chi'(v)\} = \exp[-cA(v)]$, where c is a constant of integration and

$A(v) = \exp\left(\int \frac{\zeta(v)}{v} dv\right)$. By using the boundary condition based on the assumption that $E(W) = \psi'(0) = 1$, we can show that the function $\chi(v)$ is uniquely determined by $\zeta(v)$ even when ϕ is unknown. Note that knowing $\chi(v)$ is the same as knowing the unique Laplace transform $\psi(u) = E[\exp(-uW)]$ of $F(w|\eta)$. This proves that when the observed data contain some subjects at risk of termination after their first recurrent events (i.e., $h_T(t|N(t-) = 1)$ and $h_T(t|N(t-) = 0)$ are estimable), the frailty distribution $F(w|\eta)$ is identifiable. Using similar calculations, we can also demonstrate that we can obtain the same differential equation of $\chi(v)$ using $\lambda_N(t|N(t) = 1, T \geq t)$ and $\lambda_N(t|N(t) = 0, T \geq t)$. Without giving any rigorous arguments, we claim that the L-I model is somewhat under parameterized when we expect to have a substantial amount of information on the risk of T given $N(t) \geq 2$.

2.4 GMPP Model

The generalized MPP model (GMPP model), recently introduced by Sinha and Maiti (2004) and Liu et al. (2004) (only for analysis of grouped panel count data), assumes that the hazard of T_i at time t depends on $\mathcal{N}_i(t-)$ via the subject specific unobservable frailty w_i (an assumption similar to L-I model). In the GMPP model, the assumption of (1.1) is replaced by the assumption

$$h_T(t|w, \mathcal{N}(t-); x) = w^\alpha e^{\beta_2 x} h_0(t), \tag{2.8}$$

where β_2 is the regression parameter of the covariate effect on the conditional hazard of termination. For the analysis of data example in Section 4, the form of the conditional hazard h_T in (2.8) will be further generalized by the introduction of an additional regression parameter to adjust for the direct effect of $N_i(t-)$ on $h_T(t|w, \mathcal{N}(t-); x)$. Unlike the L-I model, the conditional baseline hazard of termination $h_0(t)$ is, in general, not related to $\mu_0(t)$. However, similar to the property of the L-I model, the conditional risk of termination $h_T(t|\mathcal{N}_i(t-); x_i)$ and the conditional risk of a new recurrent event $\lambda_N(t|\mathcal{N}_i(t-), T_i \geq t; x_i)$, depend on the recurrent events history $\mathcal{N}_i(t-)$ only through the current cumulative count $N_i(t-)$. The proof of this result is very similar to the proof of (2.2) given in Oakes (1992). This property indicates that, when the actual gap-times between past events in history $\mathcal{N}(t-)$ are very informative about the risk of T in $[t, t+dt)$, the GMPP model will not be very applicable. In this case, it is not prudent to use a model where $h_T(t|\mathcal{N}(t-); x)$ depends on $\mathcal{N}(t-)$ only through $N(t-)$.

The parameter $-\infty < \alpha < \infty$ quantifies the nature of the dependence between $\mathcal{N}_i(t-)$ and the hazard of T_i. When $\alpha = 0$, we have non-informative termination, that is, $h_T(t|\mathcal{N}_i(t-); x_i)$ is free of $\mathcal{N}_i(t-)$. When $\alpha = 1$, we obtain the extension of the L-I model with time-varying $\phi_t(x)$. The theorems presented below explicitly describe the role of α, and they investigate some useful and desirable properties of the GMPP model.

Theorem 2.2—When the frailty distribution is $Ga(\eta, \eta)$, the hazard $h_T(t|N_i(t-); x_i)$ can be expressed as $h_T(t|\mathcal{N}_i(t-); x_i) = h_0(t) e^{\beta_2 x} \dfrac{E[V^\alpha \exp\{-V^\alpha H(t; x_i)\}]}{E[\exp\{-V^\alpha H(t; x_i)\}]}$ where $H(t; x) = H_0(t) \exp(\beta_2 x)$; $\mu(t; x) = \mu_0(t) \exp(\beta_1 x)$; and $V \sim Ga(N_i(t) + \eta; \mu(t; x_i) + \eta)$. For the special case $\alpha = 1$, we obtain a further simplified expression

$$h_T(t|N_i(t-); x_i) = \frac{\{\eta + N_i(t-)\} h_0(t) e^{\beta_2 x_i}}{[H(t; x_i) + \mu(t; x_i) + \eta]} = \left\{1 + \frac{N_i(t-)}{\eta}\right\} h_T(t|N(t-) = 0; x_i). \tag{2.9}$$

The proof of Theorem 2.2 is given in Appendix I. Equation (2.9) shows the role of α in describing the predictive effect of $N(t)$ on the risk of termination. We see that $h_T(t|N_i(t); x_i)$ is

an increasing function of $N_i(t-)$ for $\alpha > 0$; and $h_T(t|\mathcal{N}_i(t-);x_i)$ is a decreasing function of $N_i(t-)$ for $\alpha < 0$. We can get a crude approximation of the relationship in (2.9) as

$$h_T(t|N(t-);x) \simeq \left\{1+\eta^{-1}N(t-)\right\}^\alpha h_T(t|N(t-)=0;x).$$

(2.10)

Further, for the $Ga(\eta, \eta)$ frailty, $\lambda_N(t|N_i(t-), T_i \geq t; x_i)$ is expressed via

$$\lambda_N(t|N_i(t-),T_i \geq t;x_i)\,dt = d\mu(t;x_i)\,\frac{\{N_i(t-)+\eta\}\,E\left[\exp\left\{-V_{N+1}^\alpha H(t;x_i)\right\}\right]}{\{\mu(t;x_i)+\eta\}\,E\left[\exp\left\{-V_N^\alpha H(t;x_i)\right\}\right]},$$

(2.11)

where $H(t; x) = H_0(t)\exp(\beta_2 x)$; $\mu(t; x) = \mu_0(t)\exp(\beta_1 x)$; $V_{N+1} \sim Ga(1 + N_i(t-)+\eta; \mu(t; x_i) + \eta)$ and $V_N \sim Ga(N_i(t-)+\eta; \mu(t; x_i)+ \eta)$. For the special case $\alpha = 1$, we have

$$\lambda_N(t|N_i(t-),T_i \geq t;x_i) = \frac{\{\eta+N_i(t-)\}d\mu(t;x_i)}{\left[H(t;x_i)+\mu(t;x_i)+\eta\right]}$$
$$= \left\{1+\eta^{-1}N_i(t-)\right\}\lambda_N(t|N(t-)=0,T \geq t;x_i).$$

(2.12)

Equation (2.11) shows that the value of plays a very moderate role on the effect of $N(t-)$ in the risk of a new recurrent event, compared to the corresponding role of α for the effect of $N(t-)$ on the conditional risk for termination. Because V_{N+1} and V_N are very similar in distribution when $N(t-)$ is sufficiently large, we can get a crude approximation of the relationship in (2.11) as

$$\lambda_N(t|N_i(t-),T_i \geq t;x_i) \simeq \left\{1+\frac{N_i(t-)}{\eta}\right\}\lambda_N(t|N(t-)=0,T \geq t;x_i) \times \frac{E\left[\exp\left(-V_0 H(t;x_i)\right)\right]}{E\left[\exp\left(-V_1 H(t;x_i)\right)\right]},$$

(2.13)

where $V_0 \sim Ga(1 + \eta; \mu(t; x) + \eta)$ and $V_1 \sim Ga(\eta; \mu(t; x)+ \eta)$. Equations (2.10) and (2.13) describe the essence of the role of α in the marginal hazard of T and the risk of new recurrent events given the history $(\mathcal{N}_i(t-),T_i \geq t)$. The proof of (2.11) is similar to the proof of Theorem (2.2), and is omitted here for brevity. The properties in (2.9) and (2.12) demonstrate that the GMPP model is very different from the modulated renewal process of Oakes and Cui (1994).

The following theorem shows how α affects the relationship between the distribution of number of past recurrent events $N(t)$ and the information on whether a subject has been terminated at t. The proof is given in Appendix I.

Theorem 2.3—For the GMPP model with Ga(η,η) frailty,

$$\frac{P[N(t)=m|T=t;x]}{P[N(t)=m|T>t;x]} = \frac{E\left[\exp\left\{-W^\alpha H(t;x)\right\}\right]}{E\left[W^\alpha \exp\left\{-W^\alpha\right\}H(t;x)\right]} \times \frac{E\left[V^\alpha \exp\left\{-V^\alpha H(t;x)\right\}\right]}{E\left[\exp\left\{-V^\alpha H(t;x)\right\}\right]},$$

(2.14)

where V is distributed as Ga(m + η; μ(t; x) + η). When $\alpha > 0$, equation (2.14) is an increasing function of m and when $\alpha < 0$ it is a decreasing function of m.

One consequence of (2.14) is that $E[N(t)|T > t] < E[N(t)|T = t]$ when $\alpha > 0$ and $E[N(t)|T > t] > E[N(t)|T = t]$ when $\alpha < 0$.

The GMPP model can be also motivated via an extension of the M-site model by introducing the additional hazard (2.8) for termination as a competing risk with promotion times $Z_1, \cdots,$ Z_M of M sites. If we know that for a subject at risk of termination at time $t < \tau$, the number of past recurrent events is m (that is, m of M sites have already been promoted to recurrent events), then we expect these past event times to be i.i.d. with density $g^*(u; t) = g(u)/G(t)$ where $G(t) = P(Z < t)$. The following theorem shows this, and the common distribution of the past event times does not depend on whether the subject has been terminated at t.

Theorem 2.4—The joint conditional distribution of the recurrent event times given $\{N(t) = m, T > t, x\}$ is same as the joint conditional distribution of the recurrent event times given $\{N(t) = m, T = t, x\}$. The common joint density is given by

$$f(t_1, \cdots, t_m | N(t) = m, T > t; x) \propto \frac{d\mu(t_1; x) \cdots d\mu(t_m; x)}{\{\mu(t; x)\}^m} = \prod_{j=1}^{m} \frac{d\mu_0(t_j)}{\mu_0(t)},$$

(2.15)

where $\mu(t; x) = \exp(\beta_1 x)\mu_0(t)$.

Theorem 2.4 is another consequence of the mixed Poisson process assumptions of the GMPP model, and the proof is given in Appendix I. Theorem 2.4 implies that, if we compare a subject terminated at time y with another terminated after y, there will be no difference in their times of occurrences of past recurrent events after adjusting for their respective numbers of past events. If one believes that the termination is invariably followed by a marked change (e.g., big clustering) in the occurrences recurrent events, then this property will not hold for a given study. For example, for a study involving heart patients with recurrent acute angina, the number of past recurrences of angina as well as the exact timings and gaps between these anginas are important predictors of death. For such a study we need to explore alternative modeling strategies via directly formulating $h_T(t | \mathcal{N}(t-); x)$, were at least the time of most recent angina $Y_{N(t-)}$ has to be included as a predictor.

For the GMPP model, the distribution of the variable $N(T)$, which is the total number of recurrent events until termination, does not have a simple closed form expression. However, in some studies including the transplant study, $E[N(T)|x]$ summarizes the expected total quality of life experience over a lifetime. For a GMPP model, $E[N(T); x]$ can be computed using the marginal expectation of $N(T)$ given by

$$E[N(T)|x] = \exp(\beta_1 x + \beta_2 x) \int_0^\infty \mu_0(t) h_0(t) E_F\left[w^{\alpha+1} S_0(t)^{w^\alpha \exp(\beta_2 x)}\right] dt,$$

(2.16)

where E_F means the expectation is taken with respect to the frailty distribution $F(w|\eta)$. The proof of (2.16) follows from the fact that $E[N(T)|x] = E_G[E\{N(T)|W, x\}]$ and the property that $N(t)$ given $\{T = t, W, x\}$ is Poisson distributed.

2.5 CAR Model

Models based on CAR assumption in (1.2) are popular, since with this assumption, it is possible to make inference using proportional risks models for $h_T(t | \mathcal{N}^+(t-); x)$ and $\lambda_{N^+}(t | \mathcal{N}^+(t-), \mathcal{H}_T(t-); x)$, where $N^+(t) = N(Min\{t, T\})$. For examples, we can consider a model with $h_T(t | \mathcal{N}^+(t-); x) = h_0(t) q_2(\mathcal{N}^+(t-)|\gamma_2) \exp[\beta_2 x]$ and $\lambda_{N^+}(t | \mathcal{N}^+(t-), T \geq t; x) = \lambda_0(t) q_1(\mathcal{N}^+(t-)|\gamma_1) \exp[\beta_1 x]$, where q_k is known function of $\mathcal{N}^+(t-)$ with unknown parameter γ_k for $k = 1, 2$. These models are special cases of the multiplicative risk Andersen-Gill class of models for recurrent events (Andersen and Gill,

1982). Motivated by the expressions of the risk of new recurrent event and of the risk of termination for GMPP model obtained in previous section, we can use a CAR model with a conditional density

$$\lambda_{N_+}\left(t|\mathcal{N}^+\left(t-\right),T \geq t;x\right) = \lambda_0\left(t\right)\left(1+\eta^{-1}N\left(t-\right)\right)\exp\left[\beta_1 x\right], \tag{2.17}$$

similar to conditional density in (2.10), and a conditional hazard of termination

$$h_T\left(t|\mathcal{N}^+\left(t-\right),T \geq t;x\right) = h_0\left(t\right)\left(1+\eta^{-1}N\left(t-\right)\right)^{\alpha}\exp\left[\beta_2 x\right] \tag{2.18}$$

similar to the hazard in (2.12).

The main motivation behind using such proportional intensity and proportional hazards models of (2.17) and (2.18), is to take advantage of partial likelihood based inference (if it is justified) since we do not need to mention the baseline functions $h_0(t)$ and $\lambda_0(t)$ to estimate the parameters β_1, β_2, α and η. However, a straight-forward partial likelihood based inference for a CAR model with these proportional intensity and proportional risk models cannot be used, since the risk-set $R_{(j)}(t) = \left\{i:N_i^+\left(t-\right) = j-1, T_i \geq t\right\}$ is a biased sample of $R_{(j)}^*(t) = \left\{i':N_{i'}^+\left(t-\right) = j-1\right\}$. Works by Miloslavsky et al. (2002,2004) propose an inverse probability of censoring weighted estimator to deal with this problem. Under the assumption of CAR, the regression parameters of $\lambda_N\left(t|\mathcal{N}\left(t-\right);x\right)$ can be obtained via estimating the Anderson-Gill like intensity model $\lambda_{N_+}\left(t|\mathcal{N}^+\left(t-\right),T \geq t;x\right) = \lambda_0\left(t\right)\left(1+\eta^{-1}N\left(t-\right)\right)\exp\left[\beta_1 x\right]$. However, if one uses a proportional rate model for $E[dN^+(t)|x]$, then the corresponding regression estimates for $E[dN(t)|x]$ under CAR assumption will require further assumptions about termination risk given history. In subsequent sections, we will present a Bayesian (based on full likelihood) analysis of a transplant study under a CAR model with assumptions in (2.17) and (2.18).

There are some important differences between the CAR models and the GMPP models. First of all, a GMPP model does not satisfy the CAR assumption of (1.2) when $\alpha \neq 0$ in (2.8). We omit the proof of this result. We can further show that if we define any model (either GMPP or an extension of GMPP) based on the conditional intensity $\lambda_N\left(t|w,\mathcal{N}\left(t-\right),T \geq t;x\right)$ and conditional hazard $h_T\left(t|w,\mathcal{N}\left(t-\right);x\right)$, then we can satisfy the CAR assumption of (1.2) only when $h_T\left(t|w,\mathcal{N}\left(t-\right);x\right)$ is free of the subject-specific random effect w. Hence, the assumption of CAR implies that there is no common unobservable random-effects affecting the two responses $\{N(t)\}$ and T given x. This suggests that we should not use CAR assumption when we think that there is a latent unobservable patient-specific factor (random effect) which affects the risk of new recurrent events as well as the risk of termination. For example, in a transplant study, a CAR assumption may not be very appropriate since a patient's unobservable disease status can affect both the risk of rejection and risk of death.

3 LIKELIHOODS AND PRIORS FOR GMPP MODELS

The observable data $\mathbf{Y}$ can be split into three component vectors, $\mathbf{Y} = (D_1, D_2, X)$, where $D_1 = \{N_{io}: i = 1, \cdots, n\}$ where $\mathcal{N}_{io} = \{t_{i1} < \cdots < t_{im_i}\}$ is the data from the observed recurrent events times, $D_2 = \{t_i, \delta_i : i = 1, \cdots, n\}$ is the data from the observed termination times subject to possible censoring (δ_i is the indicator for censoring for subject i), and $X = \{x_i : i = 1, \cdots, n\}$ is the data on observed covariates.

Using the conditional intensity of $\mathcal{N}_i(t)$ in (2.1), the likelihood contribution from the sampling distribution of D_1 is

$$L_1\left(\beta_1,\mu_0,\mathbf{w}|D_1,X\right) \propto \prod_{i=1}^{n}\left[\left\{e^{\beta_1 x_i}w_i\right\}^{N_i(t_i)}\exp\left\{-e^{\beta_1 x_i}w_i\mu_0\left(t_i\right)\right\}\prod_{t\in\mathcal{N}_{io}}d\mu_0\left(t\right)\right],$$

(3.1)

where $t \in \mathcal{N}_{io}$ means that the product is taken over the set of m_i observed recurrent event times of subject i. Similarly, based on the sampling distributions of D_2 in (2.8), we get

$$L_2\left(\beta_2,h_0,\mathbf{w}|D_2,X\right) \propto \prod_{i=1}^{n}\exp\left[-w_i^{\alpha}H_0\left(t_i\right)e^{\beta_2 x_i}\right]\left\{w_i^{\alpha}h_0\left(t_i\right)e^{\beta_2 x_i}\right\}^{\delta_i}.$$

(3.2)

The observed data likelihood $L_o(\beta_1, \beta_2, \mu_0, h_0, \alpha, \eta|\mathbf{Y})$ of the GMPP model is obtained via multiple integration,

$$L_o\left(\theta|\mathbf{Y}\right) \propto \int L_1\left(\beta_1,\mu_0,\mathbf{w}|D_1 X\right) \times L_2\left(\beta_2,h_0,\alpha,\mathbf{w}|D_2,X\right)g\left(\mathbf{w}|\eta\right)\prod_{i=1}^{n}dw_i,$$

(3.3)

where $\theta = (\beta_1, \beta_2, \alpha, \mu_0(\cdot), h_0(\cdot), \eta)$, and $g(\mathbf{w}|\eta)$ is the joint density of $\mathbf{w} = (w_1, \cdots, w_n)$ given the frailty parameter η. The semiparametric maximum likelihood estimator of θ can be obtained via maximizing $L_o(\theta|\mathbf{Y})$ via a Monte Carlo EM (see Liu et al., 2004) algorithm. The E-step of each iteration requires Metropolis-Hasting simulations from the conditional density $f\left(w_i|\widehat{\theta},\mathbf{Y}\right)$, making the computation too intensive. The algorithm is also heavily dependent on the starting values, since it needs to estimate (using Breslowtype non-parametric MLEs) two nonparametric functions $\widehat{h_0}\left(t\right)$ and $\widehat{\mu_0}\left(t\right)$ at the M-step of each iteration. We note that the asymptotic justifications and convergence issues for the EM algorithm in semiparametric maximum likelihood estimation of this model are not available.

In recent years, there has been remarkable progress and growing interest in Bayesian methods for event history analysis and survival analysis (e.g., Ibrahim, Chen and Sinha (2001) and the references therein). We advocate using a semiparametric Bayes approach with an independent increments gamma process prior (Kalbfleisch, 1978; Walker, Damien, Laud and Smith, 1999; Ibrahim, Chen and Sinha, 2001) for specifying the prior distribution on the unknown function $\mu_0(\cdot)$ and the rest of the parameters. In practice, for many of the studies, when the available covariates are important well-studied variables such as treatment, race and gender, one can use informative priors on β_1 and β_2 to balance both the skeptical and the enthusiastic views about the benefits/effects of the covariates. It is also a common practice in Bayesian survival analysis to assume that priors on parameters associated with baseline functions μ_0 and h_0 be independent of the rest of the parameters. The mutual independence between the set of parameters associated with nonfatal events and the set of parameters associated with termination is also justifiable, as we can expect to have independent expert opinions about these two different types of events. Similarly, the prior belief for α, which quantifies the association between the two types of events, can be obtained independently from the others. Thus the joint prior of the parameter vector θ is often assumed to be

$$\pi\left(\theta\right) = \pi\left(\beta_1\right) \times \pi\left(\beta_2\right) \times \pi\left(\mu_0\right) \times \pi\left(h_0\right) \times \pi\left(\alpha\right) \times \pi\left(\eta\right).$$

(3.4)

The joint posterior for the GMPP model is given by

$$p(\theta, \mathbf{w}|\mathbf{Y}) \propto L_1(\beta_1, \mu, \mathbf{w}|D_1, X) \times L_2(\beta_2, \phi, \mathbf{w}|D_2, X) \times g(\mathbf{w}|\eta) \times \pi(\theta), \tag{3.5}$$

where $g(\mathbf{w}|\eta) = \Pi_i \left[\eta^\eta w_i^{\eta-1} \exp(-w_i, \eta) / \Gamma(\eta) \right]$ and $\pi(\theta)$ is given in (3.4). It is obvious that we need to use MCMC tools to sample from (3.5), and then use those MCMC samples to carry out the Monte Carlo integrations (Chen, Shao and Ibrahim (2000) and Robert and Casella (1999)). At each iterative stage of the MCMC algorithm, we only need to sample the full conditional distributions of the parameters and the frailty random effects.

For a semiparametric Bayesian analysis of GMPP model, we need to express the prior opinions about $\lambda_0(t)$ and $h_0(t)$ (baseline functions) using prior processes. The gamma prior process GP $(\gamma(t), c)$ for $\mu_0(t)$ is determined by the known prior mean $E_\pi[\mu_0(t)] = \gamma(t)$ and the known prior precision c of $\mu_0(t)$ (with prior variance $Var_\pi[\mu_0(t)] = \gamma(t)/c$). It is possible to put another level of hierarchy in this model using an unknown c and a functional form of $\gamma(t|\gamma_0)$ with some unknown γ_0. For a parametric model of $h_0(t)$ in (3.2), we suggest a Weibull hazard $h_0(t) = h_0(t|\lambda, v) = \lambda v t^{v-1}$ with a prior distribution (possibly noninformative) on $\psi = (\lambda, v)$ using a product of two independent Gamma densities (with known small mean and large variance). The Weibull assumption allows for a very wide class of choices for $h_0(t)$. However, one can also use a nonparametric $h_0(t)$ with, say, a nonparametric Gamma process on $H_0(t)$.

We need to choose the forms of the priors $\pi(\beta_1)$, $\pi(\beta_2)$, $\pi(\alpha)$ and $\pi(\eta)$, that are flexible, yet they facilitate easy to sample conditional posteriors. In particular, we use $N(\beta_{10}, \sigma_1)$ for $\pi(\beta_1)$ and $N(\beta_{20}, \sigma_2)$ for $\pi(\beta_2)$. The $\pi(\alpha)$ is assumed to be $N(0, \sigma_3)$, when we are not sure whether the association between the two classes of events should be either positive or negative. The hyperparameters of the priors are specified via a careful prior elicitation procedure. We would like to emphasize that the sampling based MCMC computational technique can handle more complicated prior structures including a $\pi(\theta)$ which induces prior dependencies among the parameters.

For MCMC sampling, we need to derive the exact full conditional distributions of each of the model parameters and random effects. Using the expression of $L_1(\beta_1, \mu_0, \mathbf{w}|D_1, X)$ in (3.1) and the property of the Gamma process prior of $\mu_0(t)$, we obtain the full conditional of $\mu_0(t)$ as

$$p(d\mu_0(t)|\text{rest}) \sim Ga(cd\gamma(t) + dN_+(t), c + R_+(t; \beta_1, \mathbf{w})), \tag{3.6}$$

where $N_+(t) = \Sigma_{\{i:t_i>t\}} N_i(t)$, $R_+(t; \beta_1, \mathbf{w}) = \Sigma_{\{i:t_i>t\}} w_i \exp(\beta_1 x_i)$ The full conditional in (3.6) can be derived following the arguments similar to previous work on recurrent events data using the Gamma process prior (e.g., Sinha (1993), Ibrahim, Chen and Sinha (2001)). We note that for computational purposes, we only need to sample from the increments of $\mu_0(t)$ in between consecutive observed recurrent event times and termination times when $N_+(t)$ and $R_+(t)$ (two step functions) change their values.

The expressions of the full conditionals for all parameters used in the MCMC chain are given in Appendix II. Except for $\mu_0(t)$, all other full conditional distributions are not in standard form. The full conditional for β_1, β_2, h_{0j}, α and η are log-concave. Thus, the adaptive rejection sampling algorithm of Gilks and Wild (1992) can be used to sample from these log-concave densities. We use the Metropolis-Hastings algorithm to sample w_i.

4 Applications

To illustrate the use of Bayesian methodology for the analysis of recurrent events with dependent termination, we use a subset of the cardiac transplant data from the Medical

University of South Carolina. Cardiac transplant recipients may experience non-fatal recurrent graft rejections after the transplants. By modeling the intensity of non-acute graft rejections and risk of death, we can understand the roles of the risk factors for death and graft rejection, and predict a patient's risk of death/rejection given history of rejections, thus it can help clinicians to make decision on patients' schedule of biopsy (a very invasive monitoring plan with potentially serious side-effects) given the history of graft rejections. The main aims of the actual study also included investigation of the effects of the time-independent covariates race and gender on the risk of recurrent graft rejections and risk of death. For this article, we are going to only analyze the data from 21 female patients from certain ages and years of entry (the details are withheld because the results of the analysis of the entire study have not been published yet). The numbers of recurrent events among these women vary between 0 to 7. The time-independent covariate available for each patient is a binary variable of race (1 for non-Caucasian women and 0 for Caucasian women). It is often argued in the transplant literature that non-Caucasian women are somewhat at higher risk of graft rejections compared to Caucasian women, and marginally, non-Caucasian women are also at higher risk of death. However, the effect of race on death for such transplant patients may be confounded by the effect of history of graft rejections on death. The complex relationship among transplant rejections, death and race are not very well-understood. A graphical representation of the data is given in Figure 1. From this figure, it is difficult to immediately comprehend the relationship of race with the observed data of these two types of responses.

A Cox partial likelihood analysis based on the response variable termination time shows that the effects of both race and the time-dependent covariate consisting of the number of previous rejections, are statistically highly significant (indicating increased risk of death for non-Caucasian women). This confirms that we have informative termination in this example. Analyzing the recurrent transplant rejections using either the methods of Sun and Wei (2000) or Wang et al. (2001) shows that the race effect is statistically significant. However, these analyses fail to explain the complex relationship of gender, rejections and death. They do not explain whether the effect of gender on death is only via its effect on rejection followed by rejection's effect on death. We also emphasize that the CAR assumption of (1.2) is clearly not physically interpretable in this situation.

For this data example, we consider following two models. (1) Model-0: a GMPP model for the $\{N(t)\}$ process with conditional intensity and hazard (respectively for recurrent rejections and death) given in (2.1) and (2.8), with a time-independent covariate x_i (gender) and patient-specific frailty w_i; (2) Model 1: an extension of the GMPP model with the conditional hazard of the termination T_i given $(x_i, \mathcal{N}_i(t-), w_i)$ in (2.8) replaced by

$$h_T(t|\mathcal{N}_i(t-), w_i; x_i)\, dt \propto w_i^\alpha \exp(\beta_2 x_i + \beta_3 N_i(t-))\, h_0(t), \tag{4.1}$$

where β_3 is an additional regression parameter measuring the direct effect of number of previous rejections $N_i(t\text{-})$ on the risk of termination. The baseline hazard is assumed to be monotone $h_0(t) = \lambda v t^{v\text{-}1}$. A credible interval of α in Model-1 can be used to evaluate the justification for a CAR assumption, since Model-1 gets reduced to a CAR model for $\alpha = 0$. It is of particular interest to know whether $v < 0$ and the actual value of v to decide whether the baseline risk of death after adjusting for effects of race and history of rejections decrease over time, and how much the risk can decrease over a certain interval, say, on interval of 3 years. The frailty density is assumed to be $Ga(\eta, \eta)$ with η^{-1} as the variability of the frailty.

For Gibbs sampling, we use the publicly available popular software WinBUGS. We follow the recommendations of Gelman and Rubin (1992) to monitor convergence. In particular we used 10 Gibbs chains with 50,000 iterations in each chain. We checked that the estimated scale

reduction factor $\hat{R}$ of the Gelman and Rubin statistic is less than 1.2 for each of the parameters. We also checked several other convergence criteria using CODA (Best et. al., 1995).

We used zero means and large variances for the hyperparameters of the prior distributions of the regression parameters β_1, β_2 and β_3 to reflect prior opinions of no direct regression effects on the rejection intensity as well as on the risk of termination. For the gamma priors for λ, ν and η, we used small scales and shapes so that the prior distributions are diffuse. We observed that our analysis is not sensitive to these choices, as long as the priors remain sufficiently diffuse. In this analysis, another crucial elicitation step is to choose the hyperparameters $\tau = 1/c$ and $\gamma_0(t)$ associated with the gamma process of $\mu_0(t)$ in (3.6). We take $\gamma_0(t) = \lambda_0 t^{\alpha 0}$ with $\lambda_0 = .01$, $\alpha_0 = .5$ and $\tau = 5.0$. The rationale behind choosing α_0 is that a plot of the prior baseline intensity over the time intervals reduces approximately according to a square root scale. We also use a very diffuse gamma process for $\mu_0(t)$ to represent low confidence in our prior guess $\gamma_0(t)$ of $\mu_0(t)$.

The results of our Bayesian analysis using Model-0 suggest that there is good evidence for a direct race effect on the intensity of recurrent graft events with 95% credible interval (0.5133, 1.55) for β_1. The 95% credible interval (0.828, 2.99) of α lies above zero indicating strong evidence that patients who are more prone to recurrent graft rejections are at higher risk of death because a higher unobservable patient-specific frailty increases the risk of recurrent rejections as well as death. The evidence for the direct race-effect on the increased risk of death is also significant, since the 95% credible interval (0.6, 2.93) for β_2 is also above zero, and the posterior median is 1.74. This suggests that the race effect on risk of death may not be explained adequately by the effect of race on death via the intermediate variable of the history of past rejections. The variance parameter η of the frailty density has a credible interval of (0.6,2.9) indicating evidence of high level of patient-heterogeneity for frequency of graft rejections.

For the extension of the GMPP model (Model-1), the corresponding credible intervals of β_1 and β_2 are (0.29,1.6) and (0.8,3.9). They are slightly wider than the credible intervals obtained from Model-0. The credible interval of β_3 (direct of number of past rejections $N(t-)$ on risk of death) is (0.16,1.17) supporting the anticipated effect of past number of rejections for increasing the risk of death. The posterior analysis of this model does not support the CAR assumption for this study, since the credible-interval of α for this model turns out to be (1.23,503), which is way above zero. In a subsequent section, we discuss about how to decide about the most appropriate model (based on observed data) among these two competing models. One of the major advantages of a Bayesian analysis is that one can compute the predictive distributions of the future event process for different races. We have omitted the predictive distributions here for brevity.

5 MODEL DIAGNOSTICS

Model validation diagnostics are very important components of Bayesian survival analysis (Sinha and Dey (1997); Ibrahim, Chen and Sinha (2001)). We present a method for Bayesian model validation for recurrent events data via the conditional predictive ordinate (CPO) (Gelfand, Dey and Chang, 1992). For subject i, the cross-validated posterior predictive probability evaluated at the observed data point is given by

$$CPO_{(i)} = P\left[\mathbf{N}_i = \mathbf{N}_{i,obs}, T_i = t_i, \delta_i | \mathbf{Y}_{(-i)}\right] = \int P\left[\mathcal{N}_i(t_i) = \mathbf{N}_{i,obs}, T_i = t_i, \delta_i | \theta\right] p\left(\theta | \mathbf{Y}_{(-i)}\right) d\theta, \tag{5.1}$$

where $\mathbf{Y}_{(-i)}$ is the observed data with subject i removed, θ is the vector of model parameters, and $\mathbf{N}_{i,obs}$ is the observed recurrent events history data for subject i. Using a result of Gelfand, Dey and Chang (1992), we have $CPO_{(i)} = \{E[P(\mathbf{N}_i = \mathbf{N}_{i,obs}, T_i = t_i | \theta)^{-1} | \mathbf{Y}]\}^{-1}$, which can be

computed using MCMC samples from the full posterior $p(\theta|\mathbf{Y})$. The method for calculating the $CPO_{(i)}$'s is particularly simple within WinBUGS. In WinBUGS, we need to specify n additional 'nodes' (which are treated as posterior quantities to be computed)

$\{P[\mathcal{N}_i(t_i)=N_{i,obs},T_i=t_i,\delta_i|\theta,w_i;x_i]\}^{-1}$ for $i=1,\cdots,n$. After computing the posterior means of these n nodes, the $CPO_{(i)}$'s are obtained as inverses of these posterior means. Typically, the $CPO_{(i)}$'s (treated like Bayesian residuals) are plotted against either x_i or t_i (or even $\mathbf{N}_{i,obs}$) to check whether the $CPO_{(i)}$'s values have any relationship with the covariates and observed data values. From this plot, we can identify possible outliers and possible evidence of a covariate dependence pattern. A plot of $-\log(CPO_{(i)}$ vs x_i can evaluate whether these cross-validated residuals depend on certain ranges of the covariate values. For the transplant data under Model-0, we found that non-Caucasian women tend to have lower $CPO_{(i)}$'s values compared to African American Women, indicating a poorer fit of Model-0 to the data from African American as compared to Caucasian women. This suggests that, in future, race should be treated as stratification variable instead of a binary covariate for the analysis of such data. A plot of $CPO_{(i)}$ vs t_i can evaluate whether the posterior analysis has any unusually small prediction capabilities for certain values of the termination times. For the transplant data under Model-0, we found no association between $CPO_{(i)}$ and t_i in Figure 2.

Ratio of $CPO_{(i)}$ under two competing models (e.g., ratio of Model-1 and Model-0 for transplant data) evaluates the relative support of observation i to Model-1 compared to Model-0. Figure 3 presents a plot of these ratios of CPO from two competing models (CPO of Model-1 divided by CPO of Model-0) versus the times of death. The ratio of greater than 1 indicates the observation from the patient prefers Model-1 against Model-0 (and vice-versa). Most observations have the ratios below 1 indicating a strong data support for Model-0 (GMPP model) over Model-1. The plot uses different symbols for races to help us evaluate whether the degree of data-support to a particular model depends on the covariate value and observed death-time. The figure suggests that support to Model-0 compared to Model-1 from observations from African American patients is somewhat stronger than the support form most of the observations from Caucasian patients. All the observations favoring Model-1 over Model-0 come from Caucasian women indicating an association between model preference and race.

6 COMMENTS AND FUTURE RESEARCH

We have proposed a comprehensive Bayesian model for recurrent events data with dependent termination whose parameters are fully interpretable. We have provided theoretical justifications for each part of the model and established relationships with other existing models. The novelty of this research is to view the recent developments of recurrent events with dependent termination data models in an unified and theoretically valid Bayesian model. Bayesian model diagnostics are also provided. The advantage of the Bayesian methodology is the ability of predicting future observations. Below, we mention some of the immediate extensions and future developments.

6.1 Cumulative Medical Care Costs Data

The inference problem of recurrent events data described here is very related to other important statistical problems such as the estimation of the mean of quality adjusted lifetime (Zhao and Tsiatis, 1997) and lifetime medical costs data. For simplicity, we focus here on the latter to identify its relationship with our models and methods. Studies of medical costs often focus on estimating the cumulative costs of care over a specified time interval, say, time of diagnosis to the time of a terminating event such as death or cure (e.g., Etzioni et al., 1996; Lin et al., 1997; Etzioni et al., 1999). For such experiments, for patient i, we can let $N_i(t)$ denote the continuous variable of cumulative cost at time t and T_i is the time to the terminating event. Our

main interest is in the distribution of $N_i(T_i)$, the cumulative medical cost at the termination time T_i. The extension of the GMPP model to cumulative medical cost data is achieved via the assumption that $\{N(t) : t \leq T\}$ given the frailty random effect w, is a realization of an independent-increment nonnegative process with conditional characteristic function

$$E\left[\exp\{iy\left(N\left(t+dt\right)-N\left(t\right)\right)|w,\mathcal{N}\left(t\right),T>t,x\right\}\right] = \psi(y)^{we^{\beta_1 x}\{\mu_0(t+dt)-\mu_0(t)\}}, \tag{6.1}$$

where $\mu_0(t)$ is a non decreasing function and $\psi(y)$ is the characteristic function of an infinitely divisible distribution function with unit mean. One example of such an $N(t)$ is the Gamma process where $\psi(y) = \{\rho/(\rho - iy)\}^\rho$ and the resulting distributions of the disjoint increments in $N(t)$ given w are independent and $N(u+\Delta)-N(u) \sim Ga(\rho(\mu_0(u+\Delta)-\mu_0(u)), \rho)$. We obtain discrete Poisson distributed increments of recurrent events data when we use $\psi(y) = \exp\{1 - \exp(iy)\}$.

For the medical costs data problem, two important quantities of much practical interest are the expected lifetime cost $C_T = E[N(T)|x]$ and the rate of cost accumulation over the lifetime $R_T = E[N(T)/T^* |x]$. Here, we define T^* as the minimum of the last inspection time a_J and the termination time T. The posterior means and credible intervals of C_T and R_T given x are useful posterior summaries of a Bayesian analysis for such data. For our model, the general simplified expression of $E[N(T)|x]$ in (2.16) is valid even for this case. The posterior mean of $E[N(T)|x]$ can be computed using MCMC samples from the posterior distribution $p(w,\theta|\mathbf{Y})$.

6.2 Extensions of the Model

One key assumption of the GMPP model is the proportional intensity of $N(t)$ given $\{w, x, T > t\}$. A possible criticism of this assumption can be addressed by extending the model to a nonproportional intensity for $N(t)$ given $\{w, x, T > t\}$. We address two ways of extending the model to a nonproportional intensity—(1) nonproportional with respect to w, and (2) nonproportional with respect to x. We now discuss possible computational difficulties and challenges in extending our Bayesian methodology to these two cases.

The accelerated time intensity model is given by replacing the assumption in (2.1) by

$$P\left[dN\left(t\right)=1|w,\mathcal{N}\left(t\right),T>t;x\right] \simeq \exp\left(\beta_1 x\right)d\mu_0\left(tw\right), \tag{6.2}$$

where $\mu_0(t)$ is the baseline cumulative intensity and its prior is modeled by an independent increment Gamma process similar to what was described in Section 3.

It is important to note that, now, $L_1(\beta_1, \mathbf{w}, \mu_0|D_1, X)$ in (3.1) involves $\mu_0(w_i t_i)$ and $d\mu_0(w_i t_i)$. For the GMPP model, at each iteration of MCMC chain, we need to only sample from the joint full conditional distribution of n increments $\mu_0(t_{(i)}) - \mu_0(t_{(i-1)})$ in between two consecutive events, and this set of increments are independent given $\mathbf{w}$. On the other hand, for the accelerated intensity model, in each iteration of the MCMC chain, after we sample the vector $\mathbf{w}$, we now need to sample from the joint full conditional distribution of the n^2 increments of $\mu_0(t)$ at time-points $t_i w_k$ for $i = 1, \cdots, n$ and $k = 1, \cdots, n$. For the accelerated intensity model, the increments $\{\mu_0(w_i t_{(k)}) - \mu_0(w_i t_{(k-1)}) : i = 1, \cdots, n; k = 1, \cdots, n\}$ are not independent. This is increases the dimensionality of the MCMC algorithm and results in slow convergence of the MCMC chain, because now the full conditional involves the entire sample path of the function $\mu_0(t)$ instead of its increments at fixed time points $t_1, \cdots, t_n$.

Another possible extension is to use the transformation intensity model (Lin, Wei and Ying, 2001) to replace the assumption of (2.1) as

$$P\left[dN(t)=1|w,\mathcal{N}(t),T>t;x\right] \simeq w \quad d\psi\left(\mu_0(t),x\right),$$ (6.3)

where ψ is a known link with (possibly) unknown parameters. One example is $\psi(\mu_0(t), x) = log[\mu_0(t)e^{\beta 1x} + 1]$. For this case, the MCMC based Bayesian methodology described in Section 3 can be (in principle) extended to the transformation intensity model. Details of these methods are yet to be explored.

6.3 Interval Censoring

In some real data settings, a continuous termination time T is monitored through scheduled inspections at $a_1, \cdots, a_J$, and it gives rise to interval censoring of the termination time within the inspection intervals. For example, patients within a trial can be monitored through monthly clinic visits and a patient's termination time is recorded only up to the last month's clinic visit before termination (death). Therefore, for each subject i, we only observe that $T_i \in I_{ji+1}$ for some inspection interval $I_{ji+1} = (a_{ji}, a_{ji+1}]$, and the observed counts N_{ik} for $k = 1, \cdots, j_i$. Thus, similar to Lin et al. (1997) and Lin (2000), here $N(a_{ji})$ is informatively censored for $N(T)$. In practice, if our modeling assumptions of the GMPP model of Section 2.4 are correct, we can make correct inference on the model parameters using a likelihood based on $P(T_i \in I_{ji+1}, N_{i1}, \cdots, N_{ij_i}|\theta)$. This likelihood will be similar to the likelihood in Section 3, except that now,

$P\left[T_i \in \left(a_j, a_{j+1}\right)|x_i, w_i, T_i > a_j\right] = 1 - h_{0j}^{\exp(\beta_2 x_i)w_i^\alpha}$, where $h_{0j} = \exp\left[-\int_{a_j}^{a_{j+1}} h_0(u)\,du\right]$. However, from this available interval-censored data, it is not possible to check whether our modeling assumptions are valid. To be more specific, if our assumptions fail within a very short interval around T_i, then we will not be able to detect that using the available data, because the available data has no information about the relationship between $N(T) - N(a_k)$ and $T - a_k$ where $a_k < T < a_{k+1}$. Even if we have available data on $N(T) - N(a_k)$, but no available data on $T - a_k$, then we cannot check the validity of our modeling assumptions.

Acknowledgments

APPENDIX I

Proof of Theorem (2.1)

We simplify the following equations by suppressing the covariate x from the expressions. We begin with the conditional distribution of $N(T)$ given the frailty w. Under the L-I model, we can show that

$$\begin{aligned} P\left[N(T)=k,T\in(t,t+dt)|w\right] &= P\left[N^*(t)=k,N^*(t,t+dt)=1|w\right] \times P\left[N(t)=k,N(t,t+dt)=0|N^*(t)=k,N^*(t,t+dt)=1\right] \\ &= \frac{\left[\mu^*(t)w\right]^k}{k!}\exp\left[-\mu^*(t)w\right]w\phi^k(1-\phi)\,d\mu^*(t). \end{aligned}$$ (6.4)

Also, note that

$$\int_0^\infty \frac{\left[\mu^*(t)w\right]^k}{k!}\exp\left[-\mu^*(t)w\right]w d\mu^*(t)=1.$$

Using the above two results, we get

$$P[N(T)=k|w] = \int_0^\infty P[N(T)=k, T \in (t, t+dt)|w] \, dt = \phi^k(1-\phi).$$

(6.5)

This completes the proof of the theorem.

Proof of Theorem (2.2)

To simplify the expressions, we suppress the covariate x and the subscript i and use the abbreviated notation $h(t) = h(t; x_i) = h_0(t) \exp(\beta_2 x_i)$; $\mu(t) = \mu(t; x_i) = \mu_0(t) \exp(\beta_1 x)$;
$H(t) = H(t; x) = \int_0^t h(u; x) \, du$. Using properties of a mixed Poisson process, we have

$P \quad [T \in (t, t+dt), N(t)=m]$
$\simeq \int_0^\infty [w^\alpha h(t) \, dt \exp\{-w^\alpha H(t)\}] \frac{[\exp\{-w\mu(t)\}\{w\mu(t)\}^m]}{m!} f(w|\eta) \, dw$
$= \frac{\{\mu(t)\}^m}{m!} h(t) \, dt \frac{\eta^\eta}{\Gamma(\eta)} \left[\int_0^\infty w^\alpha \exp\{-w^\alpha H(t)\} w^{m+\eta-1} \exp[-w\{\mu(t)+\eta\}] \, dw \right]$
$= \frac{\{\mu(t)\}^m}{m!} h(t) \, dt \frac{\eta^\eta}{\Gamma(\eta)} \frac{\Gamma(\eta+m)}{\{\mu(t)+m\}^{\eta+m}} E\left[-V^\alpha \exp\{-V^\alpha H(t)\}\right],$

where $f(w|\eta)$ is the density of the univariate frailty random variable w and V has a $Ga(m + \eta;$ $\mu(t) + \eta)$ distribution. Using the same argument, we can show that

$P \quad [T>t, N(t)=m]$
$= \frac{\{\mu(t)\}^m}{m!} \int_0^\infty [\exp\{-w\mu(t) - w^\alpha H(t)\} w^m f(w|\eta) \, dw]$
$= \frac{\{\mu(t)\}^m}{m!} \frac{\eta^\eta}{\Gamma(\eta)} \left[\int_0^\infty \exp\{-w^\alpha H(t)\} w^{m+\eta-1} \exp[-w\{\mu(t)+\eta\}] \, dw \right]$
$= \frac{\{\mu(t)\}^m}{m!} \frac{\eta^\eta}{\Gamma(\eta)} \frac{\Gamma(\eta+m)}{\{\mu(t)+m\}^{\eta+m}} E\left[\exp\{-V^\alpha H(t)\}\right].$

Dividing $P[T \in (t, t+dt), N(t)=m]$ with $P[T>t, N(t)=m]$, we can obtain the limiting expression for $h_T(t/N(t-); x)$. For the case $\alpha = 1$, using the formula for the Gamma integral, we can further show that

$$h_T(t|N(t)) = h(t) \left[\frac{E\{V\exp(-V \quad H(t))\}}{E\{\exp(-V \quad H(t))\}} \right] = h(t) \left[\frac{\eta + N(t)}{H(t) + \mu(t) + \eta} \right],$$

(6.6)

which completes the proof.

Proof of Theorem (2.3)

Using the modeling assumptions in (2.1) and (2.8), we can show that

$$P[N(t)=m|T>t] = \frac{\mu(t)^m}{m!} \frac{E\left[W^m \exp\left(-W_\mu(t) - W^\alpha H(t)\right)\right]}{E\left[\exp\left(-W^\alpha H(t)\right)\right]},$$

(6.7)

and

$$P[N(t)=m|T=t] = \frac{\mu(t)^m}{m!} \frac{E\left[W^{m+\alpha} \exp\left(-W_\mu(t) - W^\alpha H(t)\right)\right]}{E\left[W^\alpha \exp\left(-W^\alpha H(t)\right)\right]}$$

(6.8)

where the expectations are taken with respect to the $Ga(\eta, \eta)$ distribution of the frailty W.

Proof of Theorem (2.4)

Using the property of a non-homogeneous Poisson process (Ross 1983), from (2.5) and (2.6) we get

$$f(t_1, \cdots, t_m, N(t) = m, T > t | x, w) = e^{-w^\alpha H(t|x) - w\mu(t|x)} \frac{w^m}{m!} \prod_{j=1}^{m} d\mu\left(t_j | x\right)$$

(6.9)

and

$$P[N(t) = m, T > t | x, w] = e^{-w^\alpha H(t|x)} \frac{w^m}{m!} \{\mu(t|x)\}^m,$$

(6.10)

where $H(t|x) = H_0(t)e^{\beta_2 x}$. From the above two equations, we can show that

$$f(t_1, \cdots, t_m | N(t) = m, T > t, x, w) = f(t_1, \cdots, t_m | N(t) = m, T > t; x) = \frac{\prod_{j=1}^{m} d\mu\left(t_j | x\right)}{\{\mu(t|x)\}^m}.$$

(6.11)

Using the nonhomogeneous Poisson process property of $N(t)|T > t, w; x$ and conditional independence of $N(t)$ and $T \in (t, t + dt)$ given w, we get

$$P[t_1, \cdots, t_m, N(t) = m, T \in (t, t+dt) | w; x] \simeq e^{-w\mu(t;x)} \frac{w^m}{m!} \prod_{j=1}^{m} d\mu\left(t_j; x\right) \times e^{-w^\alpha H(t;x)} w^\alpha h(t; x) \, dt$$

(6.12)

and

$$P[T \in (t, t+dt), N(t) = m | w; x] \simeq e^{-w\mu(t;x)} \frac{\{w\mu(t;x)\}^m}{m!} e^{-w^\alpha H(t;x)} w^\alpha h(t|x) \, dt.$$

(6.13)

Dividing these two quantities, we can show that

$$f(t_1, \cdots, t_m | N(t) = m, T = t, w; x) = \frac{\prod_{j=1}^{m} d\mu\left(t_j; x\right)}{\{\mu(t; x)\}^m}.$$

(6.14)

This result completes the proof.

APPENDIX II

The conditional posterior distributions for the Gibbs sampling steps for discrete termination time (panel count) data are given as follows. Define

$$V_{ij} = \begin{cases} 1 & \text{if subject} \quad i \quad \text{examined at time} \quad a_j \\ 0 & \text{otherwise} \end{cases}$$

We denote the full conditional distribution of θ given all other parameters by $[\theta|rest]$.

I.
$$\left[\mu_j|\text{rest} \right] \sim Ga\left(\tau\Lambda_j + \Sigma_i N_{ij} V_{ij}, \tau + \Sigma_i w_i \exp\left(\beta_1 x_i\right) V_{ij} \right)$$

II.
$$\left[h_{0j}|\text{rest} \right] \propto \Pi_i \left[\left\{ 1 - h_{0j}^{w_i^\alpha \exp(\beta_2 x_i)} \right\}^{\delta_i V_{ij}} \left\{ h_{0j} \right\}^{(1-\delta_i)V_{ij} w_i^\alpha \exp(\beta_2 x_i)} \right] \times \pi\left(h_{0j}|e_j, v_j \right)$$

III.
$$[\beta_1|\text{rest}] \propto \exp\left\{ \Sigma_i \beta_1 x_i \Sigma_j N_{ij} V_{ij} - \Sigma_i w_i \exp\left(\beta_1 x_i\right) \Sigma_j \mu_j V_{ij} \right\} \times \pi(\beta_1)$$

IV.
$$[\beta_2|\text{rest}] \propto \exp\left\{ \Sigma_i \beta_2 x_i \delta_i \right\} \left\{ S_0(t_i) \right\}^{w_i^\alpha \exp(\beta_2 x_i)} \times \pi(\beta_2)$$

V.
$$[\alpha|\text{rest}] \propto \Pi_i w_i^{\delta_i \alpha} \{ S_0(t_i) \}^{w_i^\alpha \exp(\beta^2 x_i)} \times \pi(\alpha)$$

VI.
$$[\eta|\text{rest}] \propto \Pi_i \frac{\eta^\eta}{\Gamma(\eta)} w_i^{\eta-1} \exp\left(-\eta w_i\right) \times \pi(\eta)$$

VII.
$$[w_i|\text{rest}] \propto w_i^{\Sigma_j N_{ij} V_{ij} + \delta_i \alpha + \eta - 1} \exp\left[-w_i \left\{ \exp\left(\beta_1 x_i\right) \Sigma_j \mu_j V_{ij} + \eta \right\} \right] \{ S_0(t_i) \}^{w_i^\alpha \exp(\beta_2 x_i)}$$

REFERENCES

1. Aalen OO, Fosen J, Weedon-Fekjr H, Borgan, Husebye E. Dynamic analysis of multivariate failure time data. Biometrics 2004;60:764–773. [PubMed: 15339300]

2. Best, NG.; Cowles, MK.; Vines, SK. CODA Manual version 0.30. MRC Biostatistics Unit; Cambridge, UK: 1995.

3. Chen, M-H.; Shao, Q-M.; Ibrahim, JG. Monte-Carlo Methods in Bayesian Computation. Springer-Verlag; NY: 2000.

4. Cox DR. Regression models and life tables. Journal of the Royal Statistical Society, B 1972;34:187–220.

5. Cox, DR.; Oakes, DO. Analysis of survival data. Chapman & Hall Ltd; London; New York: 1984.

6. Dey, DK.; Mueller, P.; Sinha, D., editors. Practical Nonparametric and Semiparametric Bayesian Analysis. Springer-Verlag; New York: 1998.

7. Etzioni R, Feuer EJ, Sullivan SD, Lin D-Y, Hu C-C, Ramsey SD. On the use of survival analysis techniques to estimate medical care costs. Journal of Health Economics 1999;18:365–380. [PubMed: 10537900]

8. Fosen J, Borgan, Weedon-Fekjr H, Aalen OO. Dynamic analysis of recurrent event data using the additive hazard model. Biometrical Journal 2006;48:381–398. [PubMed: 16845903]

9. Gail MH, Santner TJ, Brown CC. An analysis of comparative carcinogenesis experiments based on multiple times to tumor. Biometrics 1980;36:255–266. [PubMed: 7407314]

10. Gelfand, AE.; Dey, DK.; Chang, H. Model determination using predictive distributions with implementation via sampling-based methods. In: Bernardo, JM.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. Bayesian Statistics. 4, Ed.. Oxford University Press; 1992. p. 147-167.

11. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science 1992;7:457–472.

12. Gilks WR, Wild P. Adaptive rejection sampling for Gibbs sampling. Applied Statistics 1992;41:337–348.

13. Griffiths RC. Characterization of infinitely divisible multivariate gamma distributions. Journal of Multivariate Analysis 1984;15:13–20.

14. Heitjan DF, Rubin DB. Ignorability and coarse data. The Annals of Statistics 1991;19:2244–2253.

15. Henederson R, Shimakura S. A serially correlated gamma frailty model for longitudinal count data. Biometrics 2003;90:355–366.

16. Hougaard, P. Analysis of Multivariate Survival Data. Springer; NY: 2000.

17. Ibrahim, JG.; Chen, M-H.; Sinha, D. Bayesian Survival Analysis. Springer-Verlag; 2001.

18. Ibrahim JG, Chen M-H, Sinha D. Criterion based methods for Bayesian model assessment. Statistica Sinica 2001b;11:419–443.

19. Kalbfleisch JD. Nonparametric Bayesian analysis of survival time data. Journal of Royal Statistical Society, B 1978;40:214–221.

20. Kuo L, Yang TY. Bayesian computation of software reliability. Journal of Computational and Graphical Statistics 1995;4:65–82.

21. Kuo L, Yang TY. Bayesian computation for nonhomogeneous Poisson processes in software reliability. Journal of the American Statistical Association 1996;91:763–773.

22. Lancaster T, Intrator O. Panel data with survival: hospitalization of HIV-positive patients. Journal of the American Statistical Association 1998;93:46–51.

23. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics 1997;53:419–434. [PubMed: 9192444]

24. Lin DY. Proportional means regression for censored medical costs. Biometrics 2000;56:775–778. [PubMed: 10985215]

25. Lin DY, Wei LJ, Ying Z. Semiparametric transformation models for point processes. Journal of the American Statistical Association 2001;96:620–628.

26. Little, RJA.; Rubin, DB. Statistical Analysis With Missing Data. Wiley; New York: 1987.

27. Liu L, Wolfe RA, Huang X. Shared frailty model for recurrent events and terminating event. Biometrics 2004;60:747–756. [PubMed: 15339298]

28. McCulloch RE. Local model influence. Journal of the American Statistical Association 1989;84:473–478.

29. Miloslavsky, M.; Keles, S.; Van der Laan, MJ.; Butler, S. U.C. Berkeley Division of Biostatistics Working Paper Series. 2002. Recurrent Events Analysis in the Presence of Time Dependent Covariates and Dependent Censoring. Working Paper 123

30. Miloslavsky M, Keles S, Laan, Mark J, Butler S. Recurrent events analysis in the presence of time-dependent covariates and dependent censoring. Journal of the Royal Statistical Society 2004;66:239–257.Series BSeries B

31. Nelson W. Confidence limits for recurrence data - Applied to cost or number of product repairs. Technometrics 1995;37:147–157.

32. Oakes D. A model for association in bivariate survival data. Journal of the Royal Statistical Society, B 1982;44:414–422.

33. Oakes, D. Frailty models for multiple event times. In: Klein, JP.; Goel, PK., editors. Survival Analysis: State of the Art. Kluwer Academic Publisher; Netherlands: 1992. p. 371-379.

34. Oakes D, Cui L. On Semiparametric Inference for Modulated Renewal Processes. Biometrika 1994;81:83–90.

35. Robert, CP.; Casella, G. Monte Carlo Statistical Methods. Wiley; New York: 1999.

36. Robins JM. Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. ASA Proceedings of the Biopharmaceutical Section 1993:24–33.

37. Ross, SM. Stochastic Process. Wiley; New York: 1983.

38. Scheike TH. Comparison of non-parametric regression functions through their cumulatives. Statistics & Probability Letters 2000;46:21–32.

39. Sinha D. Semiparametric Bayesian analysis of multiple event time data. Journal of the American Statistical Association 1993;88:979–983.

40. Sinha D, Dey DK. Semiparametric Bayesian analysis of survival data. Journal of the American Statistical Association 1997;92:1195–1212.

41. Sinha D, Maiti T. Analyzing panel-count data with dependent termination: a Bayesian approach. Biometrics 2003;60:34–40. [PubMed: 15032771]

42. Sun J, Wei LJ. Regression analysis of panel count data with covariatedependent observation and censoring times. Journal of the Royal Statistical Society, B 2000;62:293–302.

43. Tanner, MA. Tools for Statistical Inference. Third Edition. Springer-Verlag; New York: 1996.

44. Walker, Stephen G, Damien P, Laud PW, Smith AFM. Bayesian Nonparametric Inference for Random Distributions and Related Functions (with discussion). Journal of the Royal Statistical Society, B 1999;61:485–509.

45. Wang M-C, Qin J, Chang C-T. Analyzing recurrent event data with informative censoring. Journal of the American Statistical Association 2001;96:1057–1065.

46. Wei LJ, Glidden DV. An overview of statistical methods for multiple failure time data in clinical trials (with discussion). Statistics in Medicine 1997;16:833–839. [PubMed: 9160483]

47. Whittemore A, Keller JB. Quantitative theories of carcinogenesis. Society for Industrial and Applied Mathematics Review 1978;20:1–30.

48. Zhao H, Tsiatis AA. A consistent estimator of the distribution of quality adjusted survival time. Biometrika 1997;84:339–348.

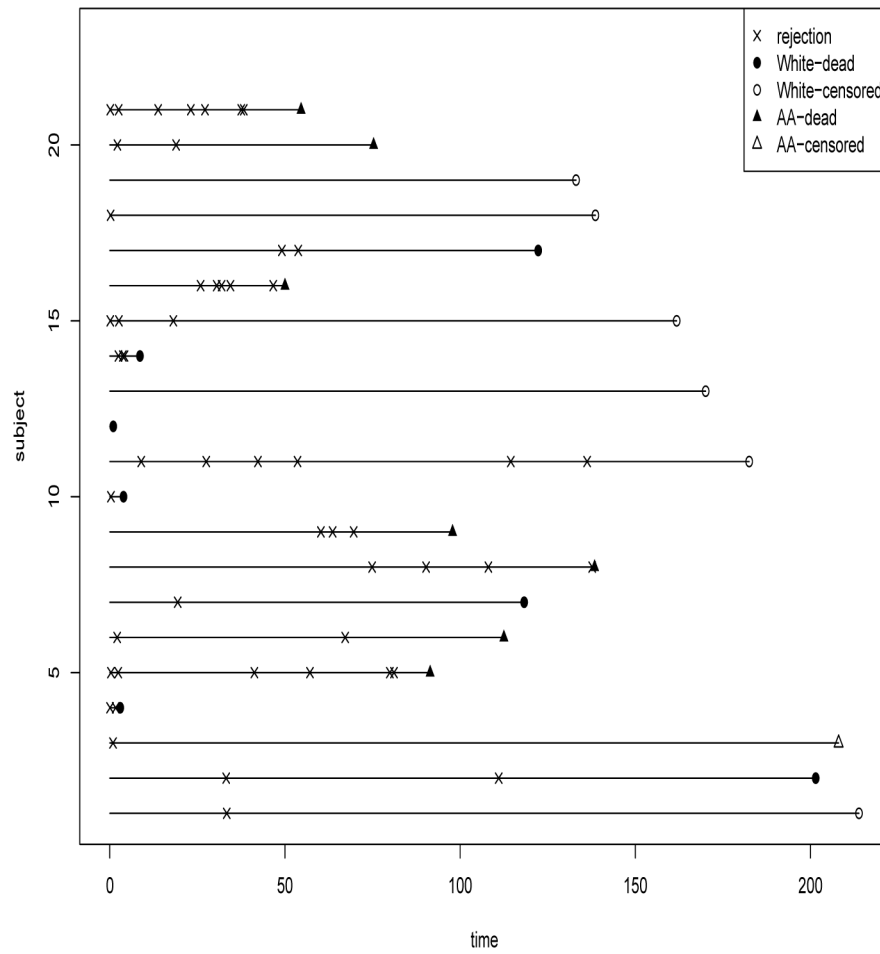**Figure 1.**
Plot of the rejection times and death for African American (AA) and Caucasian (White) females
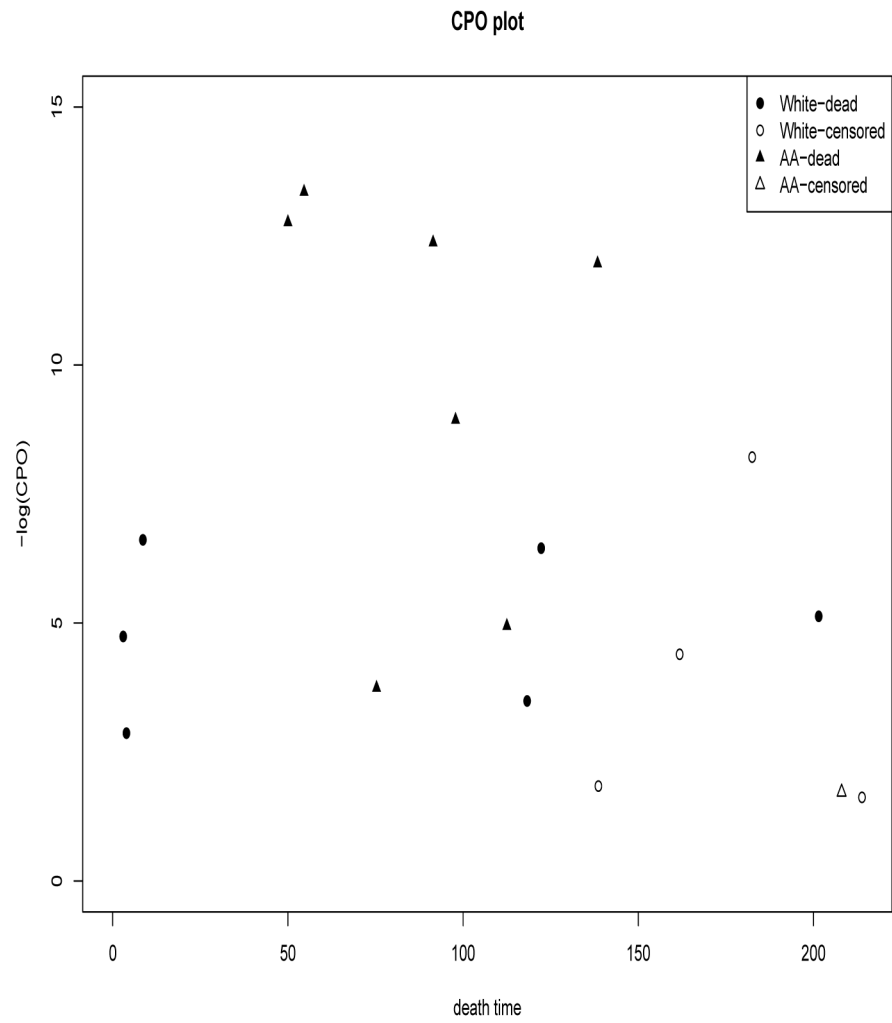from transplant study

Figure 2.
Plot of the -log(CPO) under Model-0 versus the observed time of death y_i for each patient
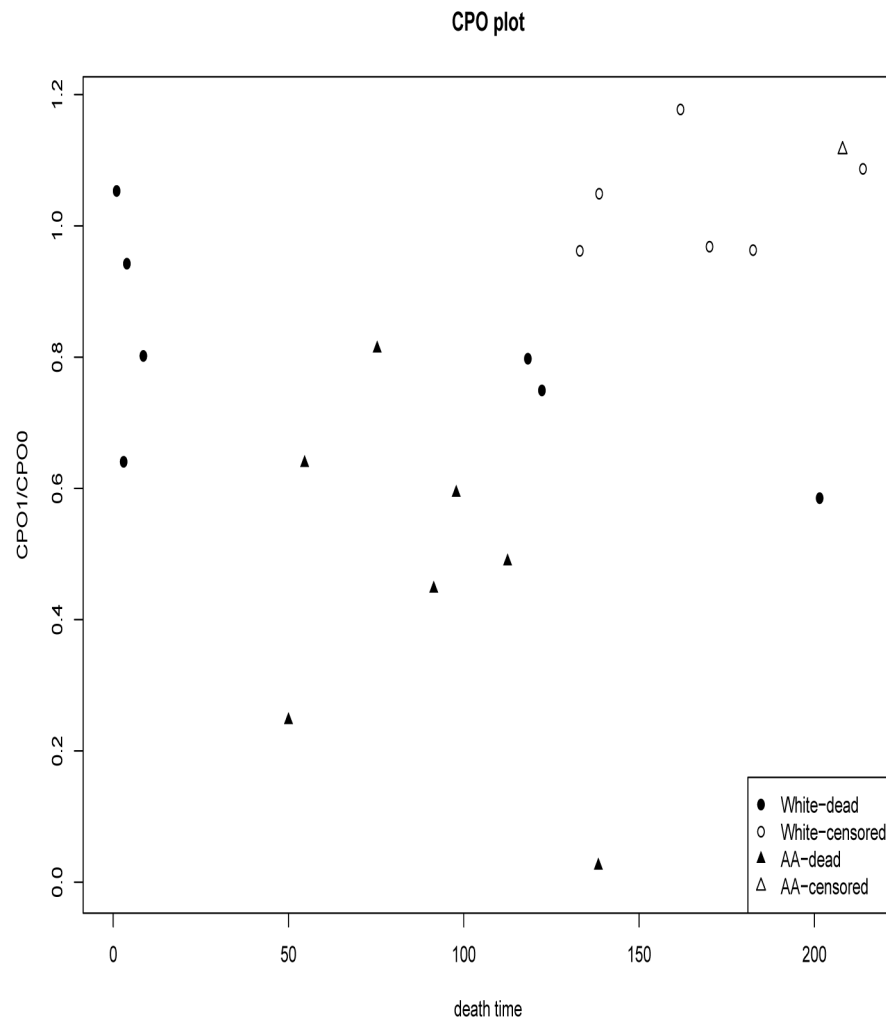
CPO plot

Figure 3.
Plot of the ratio of Model-1 CPO and Model-0 CPO versus the observed time of death y_i for each patient