



Bayesian regression model for recurrent event data with event-varying covariate effects and event effect

Li-An Lin , Sheng Luo  and Barry R. Davis

Department of Biostatistics, The University of Texas School of Public Health, Houston, TX, USA

ABSTRACT

In the course of hypertension, cardiovascular disease events (e.g. stroke, heart failure) occur frequently and recurrently. The scientific interest in such study may lie in the estimation of treatment effect while accounting for the correlation among event times. The correlation among recurrent event times comes from two sources: subject-specific heterogeneity (e.g. varied lifestyles, genetic variations, and other unmeasurable effects) and event dependence (i.e. event incidences may change the risk of future recurrent events). Moreover, event incidences may change the disease progression so that there may exist event-varying covariate effects (the covariate effects may change after each event) and event effect (the effect of prior events on the future events). In this article, we propose a Bayesian regression model that not only accommodates correlation among recurrent events from both sources, but also explicitly characterizes the event-varying covariate effects and event effect. This model is especially useful in quantifying how the incidences of events change the effects of covariates and risk of future events. We compare the proposed model with several commonly used recurrent event models and apply our model to the motivating lipid-lowering trial (LLT) component of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) (ALLHAT-LLT).

ARTICLE HISTORY

Received 27 January 2017
Accepted 13 July 2017


KEYWORDS

Event dependence; frailty model; heterogeneity; Markov chain Monte Carlo; survival model

1. Introduction

Cardiovascular disease (CVD) is a class of diseases that involves the heart or blood vessels [35]. The incidences of CVD events (e.g. myocardial infarction, congenital heart disease, and stroke) can be either fatal or non-fatal. The non-fatal CVD events may occur multiple times for some high-risk individuals, generating a statistical issue of recurrent events. These events may impair an individual's ability to live a normal daily life, resulting in enormous healthcare costs to society. In the United States alone, CVD is responsible for around 17% of national health expenditure [21]. Globally, CVD is the leading cause of death, accounting for the greatest burden of morbidity and mortality in both developed and developing countries [35]. Considering rising healthcare costs and enormous impact on human

CONTACT Sheng Luo  sheng.t.luo@uth.tmc.edu  Department of Biostatistics, The University of Texas School of Public Health, 1200 Pressler St, Houston 77030, TX, USA

 Supplemental data for this article can be accessed here. <https://doi.org/10.1080/02664763.2017.1367368>

life, multiple epidemiological and clinical trial studies have been devoted to identify and manage risk factors at both the population and subject levels through primordial, primary, and secondary prevention [43].

In the statistical analysis of single time-to-event outcomes, the standard survival model can be used [8]. When the scientific interest lies in recurrent events (e.g. recurrent CVD events), the assumption of the Cox proportional hazards model is violated due to the correlation among event times from the same individual. To adjust for this correlation, frailty models and variance-corrected models have been widely used. Frailty models are first used for cluster data where frailty is an independent and identically distributed random variable that describes the within-cluster common risks. In the context of recurrent event data, the study population is heterogeneous in lifestyle, genetic makeup, and environmental exposure, in addition to other unmeasured variables. Each study subject acts as a cluster whose frailty term introduces subject-specific heterogeneity and produces within-subject correlation in the occurrence and timing of recurrent events [1,23]. However, frailty models do not account for event dependence, that is, the occurrence of a past event may change the risk of a later event [24]. In the context of CVD, when a CVD event occurs, it can cause damage to the cardiovascular system. The damage may accumulate as the number of events increases, changing the disease progression and increasing the risk of future CVD events. Therefore, in the analysis of recurrent CVD event data, it is essential to incorporate and quantify event dependence. To this end, variance-corrected models adjust the correlated event times using robust standard error or an event-specific baseline hazard (a stratified baseline hazard that is different for each event). Specifically, there are five prominent variance-corrected models for the recurrent event data: AG model by Andersen and Gill [3]; gap time (PWP-GT) and total time (PWP-CP) models by Prentice, Williams, and Peterson [36]; LWA model by Lee, Wei, and Amato, and Leurgans [32]; and WLW model by Wei, Lin, and Weissfeld [42]. A systematic review and comparison study of these models is provided by Kelly and Lim [30]. Among these five models, PWP-GT and WLW models use stratification strategy by event-specific baseline hazard to incorporate event dependence. This type of stratified models is useful in handling event dependence because the incidence of an event may change the hazard function.

Both frailty models and stratified variance-corrected models account for either subject-specific heterogeneity or event dependence, but not both. Simulation studies have shown that the statistical inference is biased and inefficient if any of these two types of dependence are misspecified or unaccounted [4,30,41]. There is only a limited literature accounting for both subject-specific heterogeneity and event dependence in the analysis of recurrent event data. Box-Steffensmeier and De Boef [4] proposed a conditional frailty model with a frailty term and the event-stratified baseline hazard function. Similar to the stratified variance-corrected models, they utilized event-specific stratification to incorporate event dependence where different parameters for different event occurrences were adopted in the model. However, a critical issue with such stratification strategy is that the event-stratified statistical inference is inefficient and disconnects the relation between consecutive events because it essentially fits a separate model for each event [30].

Moreover, the effects of treatment and other covariates may vary as disease progresses. For example, a patient with more CVD events may have more severe disease. Some therapies may work for the first event but not for second and further events. Consequently, it is desirable to estimate the event-specific covariate effects for different event occurrences,

which may change. This would give more informative results than the classical frailty models which use common hazard function and can only estimate the ‘average effect’. Event-specific covariate effects also have the advantage of observing how effects change with each subsequent event [26,30].

In this article, we develop a novel Bayesian framework for the analysis of recurrent event data. Specifically, the framework uses: (1) a frailty distribution to characterize the subject-specific heterogeneity and within-subject dependence; (2) event-varying covariate effects to study how effects change with each subsequent event; (3) an event effect to explore and assess the event dependence among consecutive event times. The remainder of this article is organized as follows. In Section 2, we describe the lipid-lowering trial (LLT) component of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) that motivates this methodological development and propose a Bayesian regression model for recurrent event data to address the correlation among event times. Section 3 presents extensive simulation studies to evaluate the performance of the proposed model, in comparison to multiple commonly used recurrent event models. In Section 4, we apply the proposed models to the ALLHAT-LLT study and present analysis results. The concluding remarks and discussion are given in Section 5.

2. Data and models

2.1. A motivating clinical trial

This methodological research is motivated by the LLT component of the ALLHAT study [9], which was a randomized, non-blinded, large pragmatic trial conducted from February 1994 through March 2002 at 513 clinical centers in the United States, Puerto Rico, US Virgin Islands, and Canada [2]. The study was conducted in participants ages 55 and older with stage 1 and 2 hypertension and at least one additional cardiovascular risk factor. The principal objective of the ALLHAT-LLT study was to evaluate the impact of large sustained cholesterol reduction on all-cause mortality in a hypertensive cohort. Other fatal/non-fatal outcomes were prespecified as secondary endpoints, including CVD events, including myocardial infarction, stroke, and heart failure. A total of 10,355 subjects were randomly assigned to receive either pravastatin (5170 subjects) or usual care (5185 subjects). The detailed results of the ALLHAT-LLT study are previously described [2].

The follow-up time was 4.8 years in median duration, with a maximum of 7.8 years. We exclude the subjects with death confirmation pending and those with missing data in covariates and obtained the final dataset with a sample size of 9901. The recurrent events under consideration are the CVD events including non-fatal myocardial infarction, stroke, and heart failure. In the pravastatin group, 546 subjects had one CVD event and 183 subjects had at least two CVD events, while in the usual care group, 561 subjects had one CVD event and 184 subjects had at least two CVD events. Because the recurrent CVD events are of primary interest in this article, we treat as censoring the subjects who died, or lost to follow-up, or withdrew from the study. To visualize the data structure, Figure 1 displays the timing and number of CVD events for three selected subjects. The follow-up times are shown on the x -axis and the subjects’ IDs are on the y -axis. Figure 1 suggests that the frequency and event times of recurrent events vary across subjects and some subjects may experience multiple occurrences of CVD events. For example, Subject 1 had one CVD

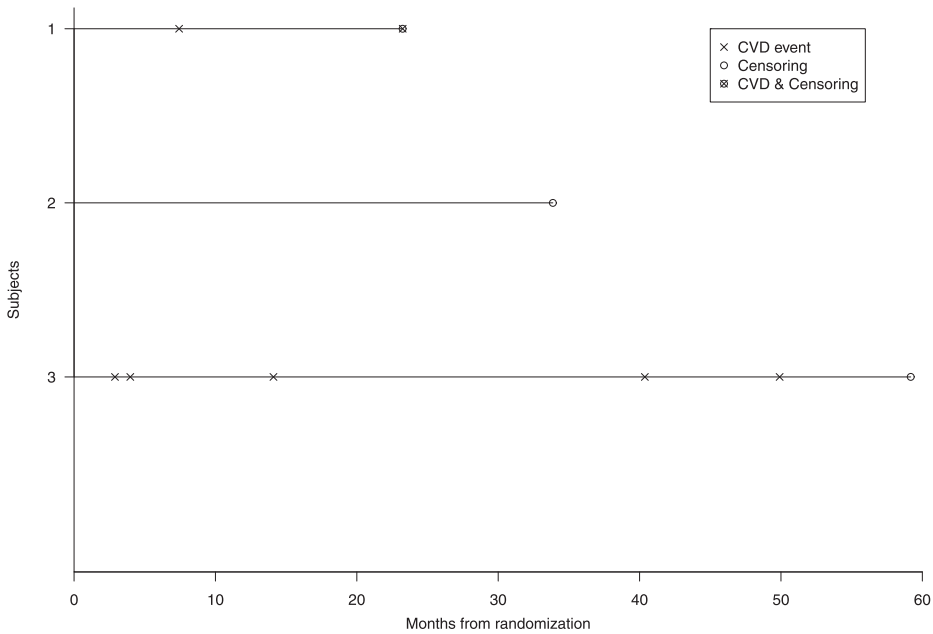


Figure 1. Three selected subjects' time to CVD events (including myocardial infarction, stroke, and heart failure).

event at 7.4 month and another CVD event at 23.2 month and died immediately after, while Subject 2 did not have any CVD event before censoring at 33.9 month, and Subject 3 had 5 CVD events at various time points. The occurrence of CVD events may increase the risk of future CVD events. Previous studies have showed that prior CVD event occurrence is a major risk factor of CVD events [14,25]. To visualize this strong correlation, Figure 2 displays cumulative hazard functions of the first CVD event after study onset for subjects with and without prior CVD events before entering the study. This figure suggests that subjects with prior CVD events had much higher risk of having another CVD event during the study than the subjects without. This phenomenon manifests the strong event dependence among the sequences of recurrent CVD events.

2.2. Models

Assuming a subject is at risk of J occurrences of recurrent events prior to censoring, we denote gap time j (time between events $j-1$ and j) by T_j and total time j (time from study onset to event j) as R_j ($R_j = \sum_{k=1}^j T_k$). Let \mathbf{Z}_j be the covariate vector related to gap time j . For each subject, the frailty model assumes that all gap times T_j , where $j = 1, \dots, J$, share a common unobserved frailty term W . Therefore, gap times for each subject are independent conditional on the frailty W and the covariate vector \mathbf{Z}_j . The hazard function for gap time j in the form of Cox proportional hazards model can be specified as

$$\lambda_j(t_j | \mathbf{Z}_j, W, T_k, k < j) = \lambda_0(t_j) \exp(\mathbf{Z}_j \boldsymbol{\beta}' + W), \quad (1)$$

where $\lambda_0(t)$ is the baseline hazard function for recurrent events, and $\boldsymbol{\beta}$ is the covariate effects of \mathbf{Z}_j . We assume frailty term W has a normal distribution with $E(W) = 0$ and

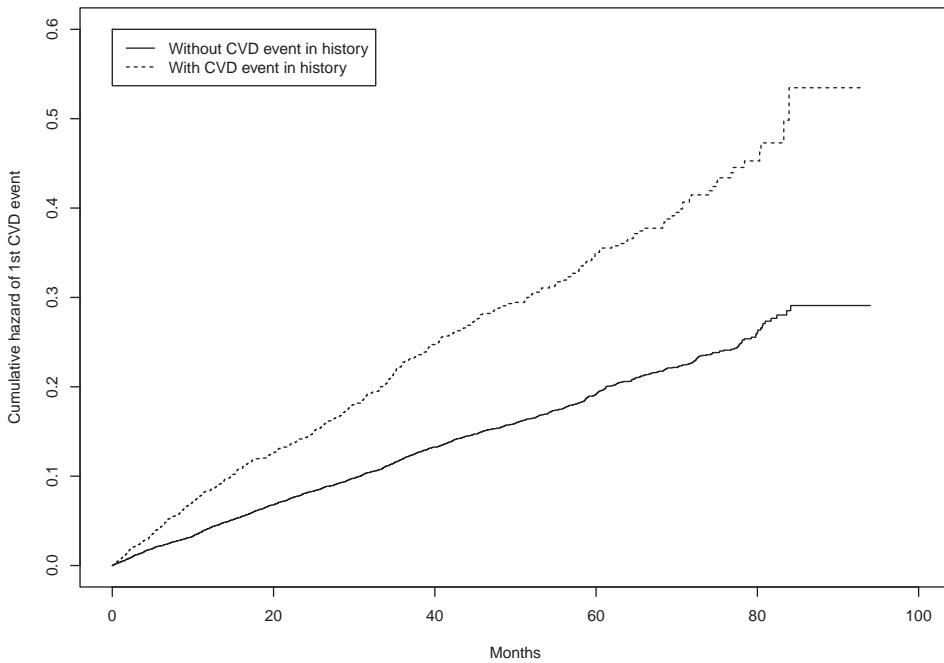


Figure 2. Cumulative hazard functions of the first CVD event after study onset for subjects with and without prior CVD events before entering the study.

$\text{Var}(W) = \tau$. The variance τ specifies the within-subject correlation between gap times and a large value of τ indicates a strong correlation among within-subject events and a large between-subject variation. We refer to this model as frailty model.

Alternatively, a subset of variance-corrected models accounts for event dependence through event stratification via event-specific baseline hazard, which is a stratified baseline hazard function that is different for each event. For example, in Prentice, Williams and Peterson gap time model [36], gap time t_j has hazard function of the form

$$\lambda_j(t_j | \mathbf{Z}_j, T_k, k < j) = \lambda_{0j}(t_j) \exp(\mathbf{Z}_j \boldsymbol{\beta}'_j), \tag{2}$$

where $\lambda_{0j}(t)$ is baseline hazard function for event j , and $\boldsymbol{\beta}_j$ is covariate effects of \mathbf{Z}_j for event j . In contrast to model (1), both $\lambda_{0j}(t)$ and $\boldsymbol{\beta}_j$ are event-specific parameters. The advantage of model (2) is that it accounts for event dependence by allowing both the covariate effects and the baseline hazard function to be affected by the earlier events. We refer to this model as PWP-GT model.

However, neither the frailty model (1) nor the stratified variance-corrected model (2) presents a general modeling strategy for the recurrent event data characterized by both subject-specific heterogeneity and event dependence [4]. A model that combines a frailty term to incorporate unobserved subject-specific heterogeneity with event-based stratification to incorporate event dependence has been studied [4,26]. The hazard function of event j is

$$\lambda_j(t_j | \mathbf{Z}_j, W, T_k, k < j) = \lambda_{0j}(t_j) \exp(\mathbf{Z}_j \boldsymbol{\beta}'_j + W). \tag{3}$$

Simulation studies showed that both models (1) and (2) give biased estimation for the fixed effects when both subject-specific heterogeneity and event dependence exist [4]. In contrast, model (3) provides robust estimation when the data exhibit either subject-specific heterogeneity or event dependence, or both. Note that both models (2) and (3) utilize event-stratification to estimate the event-specific parameters, e.g. β_j . We refer to Model (3) as conditional frailty model [4].

Although models (2) and (3) have accounted for the event dependence via event stratification, some studies have showed that the event stratification is inefficient because it essentially fits a separate model for each event and fails to make use of all the available information contained in the ordering of events [30]. Furthermore, recurrent events occur in a time-sequential fashion, for example, the second event can only occur after the first event occurs. Due to censoring, fewer subjects can have higher number of events. Models (2) and (3) become unreliable as the sample size of subjects with later events become very small and cannot provide accurate estimation of parameters for each gap [26,30]. A convenient ad hoc solution is to set a fixed number l and assume $\lambda_{0j} = \lambda_{0l}$ and $\beta_j = \beta_l$ for $j \geq l$.

Another issue with models (2) and (3) is that the event dependence is not modeled explicitly. In the context of hypertension studies, physicians and patients are often interested in knowing how event occurrences change the risk of future events. For example, survivors of a stroke may have increased risk of another stroke, which is a major source of increased mortality and morbidity [39]. To model the event effect, the hazard function of event j can be written as

$$\lambda_j(t_j | \mathbf{Z}_j, W, T_k, k < j) = \lambda_0(t_j) \exp \left(\mathbf{Z}_j \boldsymbol{\beta}'_j + \sum_{k=0}^{j-1} \alpha_k + W \right), \quad (4)$$

where $\boldsymbol{\beta}_j$ is event-stratified effect of covariate vector \mathbf{Z}_j , α_k is the additional risk introduced by the k th event (with $j = 1, \dots, J$ and $\alpha_0 = 0$), and the parameter vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{J-1})$ is defined as event effect. The quantity of α_k can be interpreted as log relative risk of next CVD event for a subject with k past events relative to a subject with $k-1$ past events. Alternatively, $\sum_{k=0}^{j-1} \alpha_k$ is log relative risk of next CVD event for a subject with $j-1$ past events relative to subject without any past event. To this end, model (4) explicitly accounts for event dependence because $\boldsymbol{\beta}_j$ and α_k specify how the covariate effects and hazard function, respectively, change as the number of past events increases. Unlike the event-stratification used for Models (2) and (3), we model the covariate effects $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_J)$ and the event effect $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{J-1})$ as regression functions varying by the number of past events. This model strategy can borrow information from adjacent events (detailed in Section 2.3), thereby improving estimation efficiency and avoiding over-fitting.

The proposed model is more comprehensive and can reduce to other models with additional restriction being imposed. If $\boldsymbol{\beta}_1 = \dots = \boldsymbol{\beta}_J$, model (4) reduces to event-invariant covariate effects model. If $\alpha_1 = \dots = \alpha_{J-1}$, model (4) reduces to linear event effect model where each event occurrence is assumed to increase the risk of next event by the same amount [18,19]. Moreover, model (1) is a special case of model (4) with $\boldsymbol{\beta}_1 = \dots = \boldsymbol{\beta}_J$ and

$\alpha_1 = \dots = \alpha_{j-1} = 0$, assuming neither event-variant covariate effects nor event dependence effects. Furthermore, our model is more informative than models (2) and (3) because event dependence is modeled explicitly using event effect parameter vector α .

2.3. Likelihood function and Bayesian inference

Because the recurrent events are subject to independent censoring, not all the event times can be observed. For subject i ($i = 1, \dots, n$), let C_i be the censoring time, T_{ij} be gap time j (time between the events $j-1$ and j), $R_{ij} = \sum_{k=1}^j T_{ik}$ be total time j (time from study onset to the event j), and $C_{ij} = C_i - R_{i,j-1}$ be gap censoring time j , where $j = 1, \dots, m_i$ with m_i being the number of recurrent events before censoring. We let the observed gap time be $X_{ij} = \min(T_{ij}, C_{ij})$ and the censoring indicator be $\delta_{ij} = I(T_{ij} \leq C_{ij})$, where $I(\cdot)$ is the indicator function. We assume that the distribution of $\{W_i, \mathbf{Z}_{ij}, C_i, X_{ij}, j = 1, \dots, m_i\}$ is independent and identically distributed for all subjects. In the article, we use lowercase letters to denote realization of the corresponding random variables, that is, $\{w_i, \mathbf{z}_{ij}, c_i, x_{ij}, j = 1, \dots, m_i\}$. Based on the hazard function (4), the likelihood function conditional on frailty $\mathbf{w} = (w_1, \dots, w_n)$ is

$$L(\cdot | \mathbf{w}) = \prod_{i=1}^n \prod_{j=1}^{m_i} \left\{ [\lambda_0(x_{ij}) \exp(\mathbf{z}_{ij} \boldsymbol{\beta}_j' + \sum_{k=0}^{j-1} \alpha_k + w_i)]^{\delta_{ij}} \times \exp \left[- \int_0^{x_{ij}} \lambda_0(t) \exp(\mathbf{z}_{ij} \boldsymbol{\beta}_j' + \sum_{k=0}^{j-1} \alpha_k + w_i) dt \right] \right\}. \quad (5)$$

In likelihood function (5), the baseline hazard function $\lambda_0(t)$ is unspecified. To increase modeling flexibility, we model $\lambda_0(t)$ as a smoothing function without specifying a specific parametric functional form. Specifically, we adopt a penalized spline with a set of sufficiently flexible spline basis. Rupper *et al.* [38] suggested that quadratic or cubic regression splines can fit the data better than linear splines with small number of knots. For cubic splines, the columns of model matrix tend to be highly correlated because each column is a transformed version of t , which can introduce considerable collinearity [29]. The collinearity may result in a nearly singular model matrix and imprecision in the spline fit [38]. The cubic B-spline basis re-scales the cubic spline basis, and is more numerically stable than truncated power bases in computation. The baseline hazard function using the cubic B-spline basis is $\lambda_0(t) = \sum_{k=1}^{K+3} u_k B_k(t)$, where $B(t) = (B_1(t), \dots, B_{K+3}(t))$ are re-scaled spline basis of $t, t^2, t^3, (t - \kappa_1)_+^3, \dots, (t - \kappa_K)_+^3$. The knot locations $(\kappa_1, \dots, \kappa_K)$ are equally spaced as suggested by Eilers and Max [12]. Gray [20] suggested to use the cubic B-spline with fairly small number of basis functions (10–20 knots) and that the number of knots has little effect on the fixed effects parameter estimation. Alternatively, other strategy (e.g. model selection criteria) can be applied to choose the optimal number and location of knots [38]. To avoid over-fitting by splines, we impose a penalty based on finite differences of the coefficients of adjacent B-splines [12]. This type of penalty is a discrete approximation to the traditional integrated square of the finite derivation and it reduces the computation complexity. The penalty term with the first order differences can be written as $\zeta \sum_{k=2}^{K+3} (u_k - u_{k-1})^2$, where ζ is the smoothing parameter [12]. In Bayesian framework,

Downloaded by [66.159.100.248] at 12:35 26 August 2017

Lang and Brezger [31] replaced the penalty term by the random walk prior distribution, that is, $u_1 \propto \text{constant}$ and $u_k \sim N(u_{k-1}, \tau_u), k = 2, \dots, K + 3$, where the variance parameter τ_u controls the amount of smoothness of spline functions, and it corresponds to the smoothing parameter ζ .

The covariate effects $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_J)$ and the event effect $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{J-1})$ are event-specific parameters which may vary by the number of past events. Since the biological mechanism of covariate and event effect on the risk function may be similar across different event incidences, it is more realistic and efficient to borrow strengths from adjacent events. To do this, we use a random walk prior process [13] that assumes a smoothing marginal dependence structure on $\boldsymbol{\beta}_j$ (or α_j). Formally, the random walk prior process is specified as $\boldsymbol{\beta}_j \sim N(\boldsymbol{\beta}_{j-1}, \tau_\beta)$, for $j = 2, \dots, J - 1$, with diffuse priors $\boldsymbol{\beta}_1 \propto \text{constant}$. A similar random walk prior with τ_α as the variance of normal prior can be used for the event effects α_j . By random walk prior specification, the estimation of $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ can borrow strength from the adjacent events and improve efficiency in the estimation process, especially when the sample sizes are small for the later events.

Because there is no prior information about the model parameters, we specify diffuse prior distributions for all other parameters. Specifically, we assign the $N(0, 100)$ prior distribution for event-invariant covariate effects, and assign the inverse gamma distribution $\text{Inv-Gamma}(a, c)$ (with very small a and c , e.g. $a = c = 0.001$) for the variance parameters τ , τ_u , τ_α , and τ_β . We have investigated other selections of diffuse prior distributions with various hyper-parameters and have obtained very similar results in both the simulation studies and the application to the ALLHAT-LLT study.

To obtain the estimates of the parameter vector, we generate samples using Hamiltonian Monte Carlo (HMC) [11] and No-U-Turn Sampler (NUTS) [22], which are MCMC algorithms that avoid random walk behavior by using the gradient of the log-posterior. Both HMC and NUTS algorithms are implemented in `Stan`, a probabilistic programming language implementing statistical inference. The model fitting is performed in `Stan` (version 2.14.0) (Stan Development Team, 2016) by specifying the full likelihood function and the prior distributions of all unknown parameters. For large datasets, `Stan` may be more efficient than the BUGS language in achieving faster convergence and requiring smaller number of samples [22]. The convergence of the MCMC chains is assessed via monitoring MCMC chain histories, autocorrelation plots, and density plots. In addition, we run multiple chains with different initial values and compute the Gelman–Rubin scale reduction statistics (\widehat{R}) to ensure \widehat{R} of all parameters smaller than 1.1 [16]. From the MCMC samples, we estimate the posterior means, posterior standard deviations, and 95% credible intervals of parameters. To facilitate easy reading and implementation of the proposed model, a `Stan` code has been posted in the Web Supplement.

2.4. Model selection criteria

We use the conditional predictive ordinate (CPO) [15] statistic for our model selection, derived from the posterior predictive distribution (PPD). Let the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \tau, \tau_u, \tau_\alpha, \tau_\beta, \mathbf{u})$ with $\mathbf{u} = (u_1, \dots, u_{K+3})$ being the vector of spline basis coefficients. For subject i , the CPO statistic is defined as $\text{CPO}_i = \int f(\mathbf{x}_i | \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | \mathcal{D}^{(-i)}) d\boldsymbol{\theta} = \{\int (\pi(\boldsymbol{\theta} | \mathcal{D}) / f(\mathbf{x}_i | \boldsymbol{\theta})) d\boldsymbol{\theta}\}^{-1}$, where $\mathbf{x}_i = (x_{i1}, \dots, x_{im_i})$ is the vector of the observed gap times, $f(\mathbf{x}_i | \boldsymbol{\theta})$ is the density function obtained from model (5), $\pi(\boldsymbol{\theta} | \mathcal{D})$ is the posterior

distribution, \mathcal{D} is the full data, and $\mathcal{D}^{(-i)}$ is the data with subject i deleted. There is no analytic solution for the above integration. In the absence of a closed form, a Monte Carlo approximation of CPO_i can be obtained using the harmonic-mean approximation [10] as $\widehat{\text{CPO}}_i = \{(1/M) \sum_{m=1}^M 1/f(\mathbf{x}_i | \boldsymbol{\theta}^{(m)})\}^{-1}$, where $\boldsymbol{\theta}^{(m)}$ is the parameter samples at the m th iteration of a total of M posterior samples after burn-in. A large CPO value indicates a better fit. The summary statistic of the CPO is the log pseudo-marginal likelihood (LPML), defined as $\text{LPML} = \sum_{i=1}^n \log(\widehat{\text{CPO}}_i)$. We can compare different models using LPML, where larger value of LPML indicates better fit.

Another widely used model selection approach is deviance information criterion (DIC) [40] which can be computed from the posterior samples. DIC provides an assessment of model fitting and a penalty for model complexity. Because of the mixture framework in our model, we use the DIC_3 measurement [6]. The DIC_3 is defined as $\text{DIC}_3 = \overline{D(\boldsymbol{\theta})} + \tau_D$, where $\overline{D(\boldsymbol{\theta})} = -2E_{\boldsymbol{\theta} | \mathcal{D}}\{\log[\prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\theta})]\}$ is the posterior mean deviance, $\tau_D = \overline{D(\boldsymbol{\theta})} + 2 \log\{E_{\boldsymbol{\theta} | \mathcal{D}}[\prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\theta})]\}$ is a measure of the effective number of parameters in the model, and $E_{\boldsymbol{\theta} | \mathcal{D}}(\cdot)$ is the expectation with respect to the posterior $\pi(\boldsymbol{\theta} | \mathcal{D})$. Thus, we have $\text{DIC}_3 = -4E_{\boldsymbol{\theta} | \mathcal{D}}\{\log[\prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\theta})]\} + 2 \log\{E_{\boldsymbol{\theta} | \mathcal{D}}[\prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\theta})]\}$. Applying Monte Carlo approximation, $\widehat{\text{DIC}}_3 = -(4/M) \sum_{m=1}^M \sum_{i=1}^n \log\{f(\mathbf{x}_i | \boldsymbol{\theta}^{(m)})\} + 2 \log\{(1/M) \sum_{m=1}^M \prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\theta}^{(m)})\}$. A smaller value of DIC_3 indicates a better-fitting model.

Moreover, we use the expected Akaike information criterion (EAIC) and the expected Bayesian information criterion (EBIC) as the additional model selection tools [5]. The EAIC and EBIC can be estimated as $\widehat{\text{EAIC}} = \overline{D(\boldsymbol{\theta})} + 2\nu$ and $\widehat{\text{EBIC}} = \overline{D(\boldsymbol{\theta})} + \nu \log n$, where ν is the number of parameters in the model. Smaller values of EAIC and EBIC indicate a better-fitting model.

3. Simulation study

We conduct simulation studies to compare the performance of four candidate models: frailty model (1), PWP-GT model (2), conditional frailty model (3), and the proposed model (4). In each simulation study, we generate 600 datasets with 400 subjects within each dataset. The event times are generated from the hazard function $\lambda_{ij}(t) = \lambda_0(t) \exp(X_{i1}\beta_j + X_{i2}\eta + w_i + \alpha_{j-1})$, where $\lambda_0(t)$ is the baseline hazard function, X_{i1} is a binary variable taking values 0 or 1 with probability 0.5, β_j is event-specific parameter, and X_{i2} is a continuous variable generated from the normal distribution with mean 0 and standard deviation 0.5. The coefficient for X_{i2} is $\eta = -0.5$. We consider frailty w_i following a normal distribution with mean 0 and variance $\tau = 0.5$. The baseline hazard function follows Weibull distribution with $\lambda_0(t) = 0.04t^{0.25}$. The event effect α_k is assumed to decrease as event increasing where $\alpha_j = 0.4j^{-0.6}$. The censoring time c_i is generated from a uniform distribution with the range of 12 and 15.

We use the inference procedure in Section 2.3 to estimate the unknown parameters and set the number of knots for baseline hazard function to be 10. We have explored different number of knots and obtained very similar results. In each simulation study, we run two parallel MCMC chains with over-dispersed initial values. Each chain is run for 10,000 iterations. The first 5000 iterations are discarded as burn-in, and the remaining 5000 samples

Table 1. Simulation results of the simulation study i.

	Frailty model				PWP-GT model			
	Bias	SEM	SE	CP(%)	Bias	SEM	SE	CP(%)
$\beta = 0.5$	0.172	0.137	0.132	76.7	-0.067	0.082	0.088	84.8
$\eta = -0.5$	-0.185	0.137	0.145	70.7	0.075	0.078	0.084	81.6
$\tau = 0.5$	0.475	0.086	0.076	0				
	Conditional frailty model				Proposed model			
	Bias	SEM	SE	CP(%)	Bias	SEM	SE	CP(%)
$\beta = 0.5$	0.028	0.111	0.105	94.2	0.008	0.109	0.100	96.7
$\eta = -0.5$	-0.034	0.110	0.109	93.3	-0.010	0.109	0.102	94.4
$\tau = 0.5$	0.071	0.115	0.107	89.4	0.023	0.118	0.116	95.6
$\alpha_1 = 0.4$					-0.026	0.089	0.085	93.3
$\alpha_2 = 0.264$					0.003	0.070	0.060	96.7
$\alpha_3 = 0.207$					-0.010	0.049	0.053	96.6
$\alpha_4 = 0.174$					-0.029	0.053	0.057	95.6

are used to obtain the posterior distributions of the parameters. We compute bias (the average of posterior means minus the true values), standard error (SE, the standard deviation of the posterior means), mean standard error (SEM, the sample mean of the square root of posterior variances), and coverage probability (CP) of 95% equal-tail credible intervals.

In the simulation study I, we assume that the event-invariant covariate effects $\beta_j \equiv 0.5$ for all j . The data are fitted by all four candidate models. The simulated data exhibit both subject-specific heterogeneity and event dependence. Therefore, we expect to see biased estimates from the models which do not consider both of these data features. Table 1 displays the results from all four models. The estimation from the frailty model is biased because it does not consider the event dependence. Moreover, the random effects variance (τ) is overestimated because under the assumption of no event dependence, the frailty model treats part of event dependence as subject-specific heterogeneity. Large bias is observed in the results from the PWP-GT model. Similar simulation results have been reported in the literature [4,30,41]. By accounting for both subject-specific heterogeneity and event dependence, the conditional frailty model provides the parameter estimates close to the true values, with SEM being close to SE and CP being close to the nominal value. In general, the conditional frailty model outperforms the frailty model and PWP-GT model, in the presence of both subject-specific heterogeneity and event dependence. However, it does not model the event dependence explicitly and cannot explicitly inform how the past events impact the risk of a next event. Compared to the other models, the proposed model provides additional information related to event dependence which may be critically important for researchers, because the hazard ratio of a subject with any number of past events having the next event as compared to a subject without event history (α) can be obtained.

In the simulation study II, we allow both the event effect and the treatment effect to be event-varying ($\alpha_j = 0.4j^{-0.6}$ and $\beta_j = 1.5 - 0.06j^{1.5}$) and compare the proposed model with the conditional frailty model. Table 2 suggests that the proposed model generally provides estimates with negligible bias, SE being close to SD, and CP being reasonably close to 95%, indicating that the proposed model can successfully recover the true parameters

Table 2. Simulation results from the proposed model and conditional frailty model in the simulation study II.

	Proposed model				Conditional frailty model			
	Bias	SEM	SE	CP(%)	Bias	SEM	SE	CP(%)
$\eta = -0.5$	0.016	0.088	0.093	94.2	0.051	0.082	0.081	90.1
$\tau = 0.5$	-0.008	0.084	0.081	95.7	-0.108	0.081	0.078	78.2
$\beta_1 = 1.440$	-0.023	0.096	0.103	93.9	-0.059	0.138	0.138	93.9
$\beta_2 = 1.330$	-0.011	0.101	0.102	96.0	-0.088	0.173	0.171	93.2
$\beta_3 = 1.188$	0.012	0.121	0.114	94.5	-0.080	0.231	0.210	94.2
$\beta_4 = 1.020$	0.034	0.141	0.134	93.3	-0.092	0.315	0.200	93.9
$\beta_5 = 0.829$	0.011	0.160	0.146	92.9	-0.296	0.320	0.370	77.5
$\alpha_1 = 0.4$	-0.038	0.093	0.085	94.2				
$\alpha_2 = 0.264$	0.012	0.065	0.064	92.3				
$\alpha_3 = 0.207$	0.010	0.057	0.059	95.7				
$\alpha_4 = 0.174$	0.012	0.052	0.041	97.5				

in the presence of both event-varying treatment effect and event dependence. In comparison, the conditional frailty model gives large bias and higher SEM for the event-varying covariate effects (β) because it essentially fits a separate model for each event. Note that the bias for β_5 from the conditional frailty model is particularly large because we assume $\lambda_{0j} = \lambda_{05}$ and $\beta_j = \beta_5$ for $j \geq 5$, to ensure sufficient sample size in the last strata (subjects with at least five recurrent events). Consequently, the estimated β_5 is the ‘average effect’ for the fifth event and the events thereafter. However, the proposed model does not have such restriction where each event effect and covariate effect are modeled explicitly.

The simulation studies I and II assume that both subject-specific heterogeneity and event dependence exist. In the simulation study III, we follow similar setting as simulation study I and assume no subject-specific heterogeneity ($\tau = 0$) and no event dependence ($\alpha_j \equiv 0$). Thus, all four models (frailty model, PWP-GT model, conditional frailty model, and the proposed model) are overparameterized. And, the simulation results (Web Table 1) suggest that the proposed model still performs reasonably well under model overparameterization.

In conclusion, the simulation results suggest that the proposed model can account for both subject-specific heterogeneity and event dependence, can explore the event-varying effects of covariates and past events, and can provide less biased parameter estimates than other candidate models. Compared to the conditional frailty model, the proposed model, by introducing event effect parameter α_k , explicitly estimates the event dependence and explores the function form of the event effect for the risk of future events. Further, the proposed model provides more accurate parameter estimates than the conditional frailty model because it can borrow information from adjacent events.

4. Application to ALLHAT data

In this section, we apply the proposed model to the motivating ALLHAT-LLT study. In the analysis of recurrent CVD events (including myocardial infarction, stroke, and heart failure), we use two parallel chains with over-dispersed initial values and run each chain for 10,000 iterations. The first 5000 iterations are discarded as burn-in and the parameter estimates are based on the remaining 5000 iterations from each chain. For the baseline hazard function, we use Bayesian B-splines with 10 knots. We have investigated other

selections of knots and obtained very similar results. For the ease of illustration, we apply the Bayesian random walker prior as detailed in Section 2.3 to estimate the event effect and the event-varying effect of one covariate (treatment), while the effects of more event-varying covariates can be easily modeled in a similar fashion. Good mixing properties of the MCMC chains are observed in the trace plots which are posted in Web Figure 1.

Previous studies suggest a number of risk factors for CVDs, including older age, male gender, hypertension, diabetes, tobacco use, excessive alcohol consumption, excessive sugar consumption, prior history of CVDs, obesity, lack of physical activity, air pollution, etc. [14,25]. Based on the data availability, we include the following covariates in the model for CVD events: lipid-lowering group ($llt = 0$ for usual care, 1 for pravastatin), gender ($gender = 0$ for female, 1 for male), race ($race = 0$ for white and others, 1 for black), diabetes ($diabetes = 0$ for no, 1 for yes), cigarette smoking ($smoker = 0$ for never smoker, 1 for current smoker, 2 for past smoker), aspirin use ($aspirin = 0$ for no, 1 for yes), antihypertensive treatment before study onset ($blmeds = 0$ for no, 1 for yes), antihypertensive randomization group ($att = 0$ for doxazosin, 1 for chlorthalidone, 2 for amlodipine, 3 for lisinopril), occurrence of CHD before study onset ($lchd = 0$ for no, 1 for yes), age, body mass index (bmi), systolic blood pressure (sbp), diastolic blood pressure (dbp), high-density lipoprotein cholesterol (hdl), and low-density lipoprotein cholesterol (ldl).

It is expected that the gap times between consecutive CVD events are positively correlated. To accommodate this correlation, we fit the proposed model with the event-varying treatment effect (llt_j) and event effect (α_j) while adjusting for all the CVD risk factors. Table 3 presents the estimated the event-varying treatment effect llt_j and the event effect α_j , $j = 1, \dots, 10$. Table 3 suggests that the overall treatment effect is not statistically significant in reducing the risk of CVD events because the 95% credible intervals of llt_j all cover 0. Furthermore, the treatment effect is rather invariant to the number of past events because the estimates of llt_j are very similar. Table 3 also suggests that the effect of each event varies. For example, the log relative risk of next CVD event for a subject with one CVD event relative to a subject without any CVD event is $\hat{\alpha}_1 = 0.654$ (95% CI [0.457, 0.834]). After the first CVD event, comparing subjects with j past events and ones with $j-1$ past events, $j > 1$, the log relative risk of having a CVD event again (α_j) is not significantly different and the event effect (α_j , $j > 1$) are smaller than α_1 . This pattern is not unexpected. The first incidence of CVD event may damage the cardiovascular system and may significantly increase the risk of CVD event recurrence. The CVD event incidences after the first one also increase the risk of future CVD events, but the magnitude of risk increasing may not be as high as the first CVD event. Based on the findings in Table 3, we reduce our proposed model to a more parsimonious one with event-invariant treatment effect ($llt_j = llt$, for $j = 1, \dots, 10$) and linear event effect function with intercept ($\alpha_j = \mu_1 + \mu_2(j-1)$, $j = 1, \dots, 9$). We refer to this model as reduced proposed model.

To assess the performance of the reduced proposed model and the conditional frailty model, we use the model selection criteria in Section 2.4. Table 4 presents DIC_3 , EAIC, and EBIC, LPML values for these two models. The results suggest that the proposed model outperforms the conditional frailty model in all criteria with smaller DIC_3 , EAIC, EBIC, and larger LPML, and hence it is selected as the final model. Table 5 displays the posterior mean, standard deviation (SD), and 95% equal-tail credible intervals from the

Table 3. Parameter estimations for event-varying treatment effects λ_{jt} and event effect α_j .

	Proposed model			
	Mean	SD	95% CI	
λ_{t_1}	-0.034	0.049	-0.131	0.061
λ_{t_2}	0.035	0.089	-0.131	0.223
λ_{t_3}	-0.022	0.133	-0.287	0.246
λ_{t_4}	-0.024	0.149	-0.481	0.166
λ_{t_5}	-0.106	0.201	-0.493	0.181
λ_{t_6}	-0.086	0.238	-0.514	0.303
λ_{t_7}	0.045	0.251	-0.556	0.331
λ_{t_8}	0.035	0.263	-0.673	0.435
λ_{t_9}	0.067	0.273	-0.865	0.417
$\lambda_{t_{10}}$	0.082	0.317	-0.895	0.532
α_1	0.654	0.097	0.457	0.834
α_2	0.138	0.069	-0.009	0.241
α_3	0.161	0.117	-0.010	0.252
α_4	0.142	0.161	-0.013	0.279
α_5	0.115	0.209	-0.046	0.445
α_6	0.072	0.247	-0.073	0.506
α_7	0.179	0.322	-0.109	0.577
α_8	0.114	0.360	-0.165	0.661
α_9	0.135	0.401	-0.254	0.597

Table 4. Model comparison statistics for the ALLHAT-LLT study.

	Conditional frailty model	Reduced proposed model
DIC ₃	50,492.77	50,270.66
EAIC	50,053.93	49,459.58
EBIC	50,241.38	49,618.19
LPML	-25,894.58	-25,387.04

Notes: DIC₃, deviance information criterion; EAIC, expected Akaike information criterion; EBIC, expected Bayesian information criterion; LPML, log pseudo-marginal likelihood.

reduced proposed model and the conditional frailty model. We find that there is no significant difference between usual care and pravastatin treatment in the risk of CVD events ($RR = \exp(-0.027) = 0.973$; 95% CI [0.894, 1.061]). The results are consistent with the final ALLHAT-LLT article [2]. Table 5 identifies some significant risk factors of CVD events (e.g. gender, race, age, diabetes, etc). For example, having diabetes significantly increases the risk of CVD events ($RR = \exp(0.370) = 1.448$; 95% CI: [1.319, 1.592]) which is also consistent with previous studies [33]. Moreover, the results suggest considerable subject-specific heterogeneity (random effects variance $\hat{\tau} = 0.734$, 95% CI: [0.594, 0.867]) and strong event dependence (i.e. $\hat{\mu}_1$ and $\hat{\mu}_2$ are significantly larger than 0). The log relative risk of next CVD event for a subject with one past events relative to a subject without any past event is 2.018 ($\exp(0.702)$, 95% CI: [1.728, 2.342]). The log relative risk of next CVD event for one additional CVD event after the first one is 1.112 ($\exp(0.106)$, 95% CI: [1.015, 1.214]). Compared with the conditional frailty model, our reduced proposed model can explicitly model the effect of past CVD events on the risk of future events.

5. Discussion

In this article, we propose a Bayesian recurrent event model to account for both the subject-specific heterogeneity and event dependence. The proposed method provides

Downloaded by [66.159.100.248] at 12:35 26 August 2017

Table 5. Results of fitting the proposed model and conditional frailty model.

	Reduced proposed model				Conditional frailty model			
	Mean	SD	95% CI		Mean	SD	95% CI	
llt (pravastatin)	-0.027	0.044	-0.112	0.059	-0.025	0.041	-0.105	0.057
att (chlorthalidone)	-0.056	0.059	-0.173	0.065	-0.049	0.056	-0.161	0.061
att (amlodipine)	-0.016	0.066	-0.147	0.113	-0.008	0.062	-0.129	0.112
att (lisinopril)	-0.012	0.067	-0.145	0.117	-0.012	0.062	-0.134	0.109
Gender (male)	0.187	0.051	0.087	0.288	0.171	0.049	0.077	0.269
Race (black)	0.131	0.048	0.035	0.226	0.126	0.044	0.039	0.211
Age (years)	0.044	0.003	0.038	0.051	0.042	0.003	0.035	0.048
bmi (kg/m ²)	0.004	0.004	-0.004	0.013	0.004	0.004	-0.004	0.011
sbp (mmHg)	0.008	0.002	0.004	0.011	0.007	0.002	0.004	0.011
dbp (mmHg)	-0.011	0.003	-0.016	-0.006	-0.010	0.002	-0.015	-0.005
Diabetes (yes)	0.370	0.048	0.277	0.465	0.343	0.045	0.257	0.435
hdl (mg/dl)	-0.010	0.002	-0.014	-0.006	-0.010	0.002	-0.013	-0.006
ldl (mg/dl)	0.003	0.001	0.001	0.005	0.003	0.001	0.001	0.005
Smoker (current)	0.436	0.064	0.310	0.560	0.408	0.062	0.286	0.532
Smoker (past)	0.104	0.054	-0.002	0.209	0.096	0.051	-0.002	0.198
Aspirin (yes)	0.238	0.047	0.146	0.330	0.218	0.045	0.131	0.309
blmeds (yes)	0.300	0.087	0.134	0.473	0.289	0.082	0.132	0.451
lchd (yes)	0.457	0.061	0.339	0.579	0.412	0.061	0.297	0.534
τ	0.734	0.070	0.594	0.867	0.441	0.168	0.194	0.799
μ_1	0.702	0.079	0.547	0.851				
μ_2	0.106	0.046	0.015	0.194				

robust estimation when the data exhibit both the subject-specific heterogeneity and event dependence, while frailty models and variance-corrected models ignore one of the data features and produce biased estimation. To model the event-varying covariate effects and event effect, we utilize the Bayesian random walk prior. This model strategy can borrow strength from adjacent events and perform better than the event-stratification strategy used in the conditional frailty model. Moreover, our model can inform how the covariate effects and the risk of future disease events change as the number of past events accumulates. The simulation studies suggest that our model is flexible and it can estimate different functional forms of event-varying covariate effects and event effect. In the analysis of the ALLHAT-LLT study, the proposed model has a better fit than the conditional frailty model. We have identified event-invariant and insignificant pravastatin effect, positive event effect on the risk of CVD events, some significant risk factors of recurrent CVD events, in addition to sizeable subject-specific heterogeneity.

There are some limitations in our proposed model that we will address in a future study. First, the recurrent event process may be stopped by some terminal events such as dropout or death. The terminal events are often correlated with the recurrent process (e.g. the subjects with more CVD occurrences are more likely to die), leading to the issue of dependent termination. The dependent termination has a non-neglectable impact on the occurrences of the recurrent events [7,17,27]. In this scenario, the shared frailty model is commonly used to introduce the correlation between the processes of recurrent events and termination [26]. However, the recurrent events may change the disease progression and may consequently change the risk of death. Thus, the event dependence between recurrent events and terminal event also needs to be accounted for and quantified. How to address the mutual event dependence in the joint analysis of recurrent events and dependent termination is an important direction of future research. Moreover, we have chosen

a normal distribution for the subject-specific heterogeneity. In generalized linear mixed models, misspecification of random effects distribution has little impact on the parameters that are not associated with the random effects [28,34,37]. The impact of random effects misspecification in our model framework warrants further investigation.

Acknowledgments

The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high-performing computing resources that have contributed to the research results reported within this article. URL: <http://www.tacc.utexas.edu>.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Sheng Luo's research was supported in part by the National Institute of Neurological Disorders and Stroke under Award Numbers R01NS091307 and 5U01NS043127. Barry Davis's research was supported in part by Health and Human Services contracts N01-HC-35130 and HHSN268201100036C from the National Heart, Lung, and Blood Institute, National Institutes of Health, US Department of Health and Human Services, Bethesda, MD.

ORCID

Li-An Lin  <http://orcid.org/0000-0003-2731-1346>

Sheng Luo  <http://orcid.org/0000-0003-4214-5809>

References

- [1] O.O. Aalen and E. Husebye, *Statistical analysis of repeated events forming renewal processes*, Stat. Med. 10 (1991), pp. 1227–1240.
- [2] ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial, *Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to pravastatin vs usual care: The antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT-LLT)*, J. Am. Med. Assoc. 288 (2002), pp. 2998–3007.
- [3] P.K. Andersen and R.D. Gill, *Cox's regression model for counting processes: A large sample study*, Ann. Stat. 10 (1982), pp. 1100–1120.
- [4] J.M. Box-Steffensmeier and S. De Boef, *Repeated events survival models: The conditional frailty model*, Stat. Med. 25 (2006), pp. 3518–3533.
- [5] B. Carlin and T. Louis, *Bayesian Methods for Data Analysis*, Chapman & Hall/CRC, Boca Raton, 2009.
- [6] G. Celeux, F. Forbes, C.P. Robert, and D.M. Titterton, *Deviance information criteria for missing data models*, Bayesian Anal. 1 (2006), pp. 651–673.
- [7] R.J. Cook and J.F. Lawless, *The Statistical Analysis of Recurrent Events*, Springer, Berlin, 2007.
- [8] D.R. Cox, *Partial likelihood*, Biometrika 62 (1975), pp. 269–276.
- [9] B.R. Davis, J.A. Cutler, D.J. Gordon, C.D. Furberg, J.T. Wright Jr, W.C. Cushman, R.H. Grimm, J. LaRosa, P.K. Whelton, H.M. Perry, M.H. Alderman, C.E. Ford, S. Oparil, C. Francis, M. Proschan, S. Pressel, H.R. Black, and C.M. Hawkins, *Rationale and design for the antihypertensive and lipid lowering treatment to prevent heart attack trial (ALLHAT)*. ALLHAT research group, Am. J. Hypertens 9 (1996), pp. 342–360.

- [10] D.K. Dey, M.H. Chen, and H. Chang, *Bayesian approach for nonlinear random effects models*, *Biometrics* 53 (1997), pp. 1239–1252.
- [11] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth, *Hybrid Monte Carlo*, *Phys. Lett. B* 195 (1987), pp. 216–222.
- [12] P.H.C. Eilers and B.D. Marx, *Flexible smoothing with B-splines and penalties*, *Stat. Sci.* 11 (1996), pp. 89–121.
- [13] L. Fahrmeir and S. Lang, *Bayesian inference for generalized additive mixed models based on Markov random field priors*, *J. R. Stat. Soc.: Ser. C (Appl. Stat.)* 50 (2001), pp. 201–220.
- [14] V. Fuster and B.B. Kelly, *Promoting Cardiovascular Health in the Developing World: a Critical Challenge to Achieve Global Health*, National Academies Press, Washington, DC, 2010.
- [15] S. Geisser and W.F. Eddy, *A predictive approach to model selection*, *J. Am. Stat. Assoc.* 74 (1979), pp. 153–160.
- [16] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin, *Bayesian Data Analysis*, CRC Press, Boca Raton, 2013.
- [17] D. Ghosh and D. Lin, *Marginal regression models for recurrent and terminal events*, *Stat. Sin.* 12 (2002), pp. 663–688.
- [18] H.K. Gjessing, K. Røysland, E.A. Pena, and O.O. Aalen, *Recurrent events and the exploding Cox model*, *Lifetime Data Anal.* 16 (2010), pp. 525–546.
- [19] J.R. González and E.H. Slate, *Modelling intervention effects after cancer relapses*, *Stat. Med.* 24 (2005), pp. 3959–3975.
- [20] R.J. Gray, *Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis*, *J. Am. Stat. Assoc.* 87 (1992), pp. 942–951.
- [21] P.A. Heidenreich, J.G. Trogon, O.A. Khavjou, J. Butler, K. Dracup, M.D. Ezekowitz, E.A. Finkelstein, Y. Hong, S.C. Johnston, A. Khera, D.M. Lloyd-Jones, S.A. Nelson, G. Nichol, D. Orenstein, P.W.F. Wilson, and Y.J. Woo, *Forecasting the future of cardiovascular disease in the United States: A policy statement from the American Heart Association*, *Circulation* 123 (2011), pp. 933–944.
- [22] M.D. Homan and A. Gelman, *The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo*, *J. Mach. Learn. Res.* 15 (2014), pp. 1593–1623.
- [23] P. Hougaard, *Frailty models for survival data*, *Lifetime Data Anal.* 1 (1995), pp. 255–273.
- [24] P. Hougaard, *Analysis of Multivariate Survival Data*, Springer, Berlin, 2000.
- [25] B.V. Howard and J. Wylie-Rosett, *Sugar and cardiovascular disease: A statement for healthcare professionals from the committee on nutrition of the council on nutrition, physical activity, and metabolism of the American Heart Association*, *Circulation* 106 (2002), pp. 523–527.
- [26] X. Huang and L. Liu, *A joint frailty model for survival and gap times between recurrent events*, *Biometrics* 63 (2007), pp. 389–397.
- [27] X. Huang and R.A. Wolfe, *A frailty model for informative censoring*, *Biometrics* 58 (2002), pp. 510–520.
- [28] H. Jacqmin-Gadda, S. Sibillot, C. Proust, J.M. Molina, and R. Thiébaud, *Robustness of the linear mixed model to misspecified error distribution*, *Comput. Stat. Data Anal.* 51 (2007), pp. 5142–5154.
- [29] L.J. Keele, *Semiparametric Regression for the Social Sciences*, John Wiley & Sons, Hoboken, NJ, 2008.
- [30] P.J. Kelly and L.L.Y. Lim, *Survival analysis for recurrent event data: An application to childhood infectious diseases*, *Stat. Med.* 19 (2000), pp. 13–33.
- [31] S. Lang and A. Brezger, *Bayesian P-splines*, *J. Comput. Graph. Stat.* 13 (2004), pp. 183–212.
- [32] E.W. Lee, L. Wei, D.A. Amato, and S. Leurgans, *Cox-type regression analysis for large numbers of small groups of correlated failure time observations*, in *Survival Analysis: State of the Art*, Springer, Berlin, 1992, pp. 237–247.
- [33] J.E. Manson, G.A. Colditz, M.J. Stampfer, W.C. Willett, A.S. Krolewski, B. Rosner, R.A. Arky, F.E. Speizer, and C.H. Hennekens, *A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women*, *Arch. Intern. Med.* 151 (1991), pp. 1141–1147.

- [34] C.E. McCulloch and J.M. Neuhaus, *Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter*, Stat. Sci. 26 (2011), pp. 388–402.
- [35] S. Mendis, P. Puska, B. Norrving, *Global Atlas on Cardiovascular Disease Prevention and Control*, World Health Organization, Geneva, 2011.
- [36] R.L. Prentice, B.J. Williams, and A.V. Peterson, *On the regression analysis of multivariate failure time data*, Biometrika 68 (1981), pp. 373–379.
- [37] D. Rizopoulos, G. Verbeke, and G. Molenberghs, *Shared parameter models under random effects misspecification*, Biometrika 95 (2008), pp. 63–74.
- [38] D. Ruppert, M.P. Wand, and R.J. Carroll, *Semiparametric Regression*, Cambridge University Press, Cambridge, 2003.
- [39] R.L. Sacco, P.A. Wolf, W.B. Kannel, and P.M. McNamara, *Survival and recurrence following stroke. The Framingham study*, Stroke 13 (1982), pp. 290–295.
- [40] D.J. Spiegelhalter, N.G. Best, B.P. Carlin, and A. Van der Linde, *Bayesian measures of model complexity and fit*, J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 64 (2002), pp. 583–639.
- [41] T.M. Therneau, *Modeling Survival Data: Extending the Cox Model*, Springer, Berlin, 2000.
- [42] L.J. Wei, D.Y. Lin, and L. Weissfeld, *Regression analysis of multivariate incomplete failure time data by modeling marginal distributions*, J. Am. Stat. Assoc. 84 (1989), pp. 1065–1073.
- [43] N.D. Wong, *Epidemiological studies of CHD and the evolution of preventive cardiology*, Nat. Rev. Cardiol. 11 (2014), pp. 276–289.