HHS Public Access

Author manuscript

Stat Methods Med Res. Author manuscript; available in PMC 2017 May 06.

Published in final edited form as:

Stat Methods Med Res. 2017 December ; 26(6): 2869–2884. doi:10.1177/0962280215613378.

Bayesian analysis of multi-type recurrent events and dependent termination with nonparametric covariate functions

Li-An Lin[*], **Sheng Luo**[†], **Bingshu E. Chen**[‡], and **Barry R. Davis**[§]

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

[‡]Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada

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

Abstract

Multi-type recurrent event data occur frequently in longitudinal studies. Dependent termination may occur when the terminal time is correlated to recurrent event times. In this article, we simultaneously model the multi-type recurrent events and a dependent terminal event, both with nonparametric covariate functions modeled by B-splines. We develop a Bayesian multivariate frailty model to account for the correlation among the dependent termination and various types of recurrent events. Extensive simulation results suggest that misspecifying nonparametric covariate functions may introduce bias in parameter estimation. This method development has been motivated by and applied to the lipid-lowering trial (LLT) component of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT).

Keywords

Recurrent events; joint model; multivariate frailty model; Markov Chain Monte Carlo; Hypertension

1 Introduction

Recurrent events of the same type occur frequently in longitudinal clinical trials and observational studies. Examples include repeated lung infections in people with cystic fibrosis[1], recurrent shunt failures in children with hydrocephalus[2], and recurrent strokes in older adults[3]. An important feature of recurrent events is that the event times within the same subject are correlated. The correlation among event times violates the independence assumption of the Cox proportional hazards model. As a result, the cox model produces estimates that are both biased and insufficient in the context of repeated events[4–6]. Many authors have investigated the analysis of recurrent event data, e.g., Vaida and Xu[7], Ibrahim *et al*[8], Pepe and Cai[9], and Lin *et al*[10]. Moreover, multi-type recurrent event data arise when one subject experiences two or more different types of recurrent events during the course of the study. A scientific objective is to examine covariate effects on the risks of different types

[†]Corresponding author: Sheng Luo is Associate Professor, Department of Biostatistics, The University of Texas School of Public Health, 1200 Pressler St, Houston, TX 77030, USA (sheng.t.luo@uth.tmc.edu; Phone: 713-500-9554).

of recurrent events. Different types of recurrent events may be correlated to each other, e.g., acute myocardial infarction increases the occurrence of heart failure[11]. Therefore, it may not be sufficient to perform separate analyses for each type of recurrent events, while ignoring the correlation among event types[12]. To this end, Cai and Schaubel[13] proposed a class of semiparametric marginal mean/rates models for multi-type recurrent event data with a general relative risk form and established a partial likelihood score function to estimate the covariate effects. Chen *et al* [12] used the Gibbs sampling algorithm to fit a Bayesian model for the interval-censored recurrent event data.

Another important feature of recurrent event data is that recurrent events are often subject to termination. Such termination may be either independent (the terminal time is independent of the recurrent event times, e.g., administrative censoring, dropout due to moving) or dependent (the terminal time is correlated to the recurrent event times) on the recurrent events. Dependent termination has a non-neglectable impact on the occurrences of the recurrent events. For example, the risk of recurrent strokes is positively associated with the risk of death and no further stroke events can occur once patients die. Ignoring dependent termination leads to bias in model parameter estimation[14–16]. To resolve this issue, Chen and Cook[17] developed methods for treatment comparison in the presence of multi-type recurrent events and a terminal event. Zhu *et al* [18] treated both the distributions of the dependent termination and latent variables as nuisance parameters and developed statistical methods for estimating regression parameters of recurrent events. Mazroui *et al* [19] later developed a multivariate frailty model for multi-type recurrent events with dependent termination and applied a Gauss-Hermite quadrature approximation with the penalized likelihood method for statistical inference.

In the regression analysis of multi-type recurrent events with a dependent termination, it is common that the true underlying covariate effect functions are nonlinear, rather than linear. For example, age is one of the largest risk factors for cardiovascular diseases, and its functional form for risk of cardiovascular diseases may not be linear[20]. Multiple studies have reported that the risk of cardiovascular diseases increases greatly for older adults, e.g., the risk of stroke doubles for every decade of age increase after age 55[21]. Modeling nonparametric covariate functions in the presence of single type recurrent events and a dependent termination has been discussed by Yu and Liu[22], who used the penalized partial likelihood (PPL) method[23] to estimate regression parameters. However, to the best of our knowledge, there is no research conducted to simultaneously consider multi-type recurrent events, dependent termination, and nonparametric covariate functions in a Bayesian inference framework. It is not clear how the dependence structure of these three features interact and influence the statistical inference.

The goal of this article is to investigate the effects of nonparametric covariate functions in the joint modeling framework of multi-type recurrent events with dependent termination. We use multivariate frailty distributions to model the heterogeneity among subjects, the correlation among different event types, in addition to the correlation between recurrent events and dependent termination. Gaussian quadrature method and penalized partial likelihood have been developed to fit frailty models in recurrent event data. However, Gaussian quadrature method cannot incorporate the penalty term of the penalized spline

function into the convenient SAS procedure PROC NLMIXED (SAS Institute Inc., Cary, NC, USA)[22]. And, the penalized partial likelihood method [23] underestimates the frailty variance which is an essential parameter to quantify the association between different types of recurrent events and a terminal event. Therefore, we will use Bayesian methodology for statistical inference.

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 the joint models for the multi-type recurrent events, dependent terminal event, and nonparametric functions of covariates. Section 3 discusses Bayesian inference and Bayesian model selection criteria. Section 4 presents an extensive simulation study to evaluate the performance of the proposed models. In Section 5, we apply the proposed models to the LLT component of the ALLHAT study and present analysis results. In Section 6, we provide concluding remarks and discussion.

2 Data and Joint Models

2.1 A Motivating Clinical Trial

This methodological research is motivated by the lipid-lowering trial (LLT) component of the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) study[24], referred to as the ALLHAT-LLT study, 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[25]. 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 outcomes included CHD (composite of fatal coronary heart disease and nonfatal myocardial infarction), stroke (fatal and nonfatal), and heart failure (hospitalized or fatal). A total of 10,355 subjects were randomly assigned to receive either pravastatin (5,170 subjects) or usual care (5,185 subjects). The detailed results of the ALLHAT-LLT study are described in the final ALLHAT-LLT article[25].

The follow-up time was 4.8 years in mean duration, with a maximum of 7.8 years. At the end of the trial, 84.8% of subjects were known to be alive, 12.3% were confirmed dead, 0.5% were reported dead with confirmation pending, and 2.4% were lost to follow-up or withdrawal from the study. We excluded the subjects with death confirmation pending and those with missing data in covariates, and obtained the final dataset with a sample size of 9,901. The subjects who were alive at the end of the trial, or lost to follow-up during the trial are considered as independent termination (right censored) and the time to the last follow-up visit is used as the survival time. During the follow-up of the study, subjects may experience three types of recurrent cardiovascular disease events (CHD, stroke, and heart failure) before death or censoring. To visualize the data structure, Figure 1 displays the event occurrence times (from randomization) for four selected subjects. The follow-up times are displayed on the x-axis and the subjects' IDs are on the y-axis. For example, subject 2 experienced one CHD event at 4.4 months and one stroke event at 35.7 months before he/she died at 67.6 months from randomization. Some important characteristics of the data are: (1) each subject

may experience multiple occurrences of one type of events (e.g., subject 3 had 4 recurrent heart failure events, while subject 4 had 3 recurrent CHD events); (2) subjects may experience more than one type of recurrent events prior to death or censoring (e.g., subjects 2, 3, and 4); (3) the frequencies and event times of the recurrent events vary across subjects. Table 1 lists the numbers of subjects with various types of recurrent CHD, stroke, and heart failure events. For example, in the pravastatin group, 331 subjects had one CHD occurrence and 29 subjects had at least two CHD occurrences. There were 583 and 596 all-cause mortalities in the pravastatin and usual care groups, respectively. If mortality occurred immediately after the occurrence of CHD, stroke, or heart failure, then the time to all-cause mortality and the time to the last recurrent event are identical.

Subjects with at least one type of cardiovascular disease events have higher risks of all-cause mortality and other types of cardiovascular disease events than subjects without any cardiovascular disease events[20,26,27]. To visualize these correlations, Figure 2 displays the estimates of cumulative hazard rates of death (panel a), CHD (panel b), stroke (panel c), and heart failure (panel d). Panel a suggests that subjects with occurrences of CHD, stroke, or heart failure had higher cumulative hazard of death than those without. This phenomena manifests the strong correlation between the recurrent cardiovascular disease events and death. Similarly, panels b, c, and d suggest that subjects with occurrence of one type of cardiovascular disease events had higher cumulative hazards of other types of cardiovascular disease events than those without. Moreover, among the risk factors for cardiovascular disease events, age is the most important one; and it has been suggested that the age effect on the risk of cardiovascular disease events is nonlinear[20,28]. Therefore, it is essential to develop a joint model framework for the analysis of the multi-type recurrent events (CHD, stroke, heart failure) and the dependent termination (all-cause mortality) with nonlinear covariate functions.

2.2 Joint Models of Multi-type Recurrent Events and Dependent Termination with Nonparamertric Covariate Function

Let $r_{ij}(t)$ be the hazard function of type j recurrent events for subject i at time t since study onset, where $i = 1, \ldots, I$, and $j = 1, \ldots, J$. To analyze recurrent event data, the time scale can be gap times (time between two successive events) or calendar time (time since study onset). This article focuses on calendar time while the proposed method can be applied to both time scales. Let n_{ij} be the number of type j recurrent events, and t_{ijr} be the time to the rth ($r = 1, \ldots, n_{ij}$) occurrence of type j recurrent event of subject i, since study onset. All types of recurrent event are stopped by $T_i = \min(D_i, C_i)$, where D_i is time to a terminal event (e.g., all-cause mortality), C_i is time to censoring (e.g., end of study or dropout). Let $\lambda_i(t)$ be the hazard of death for subject i at time t and the death indicator be $\Delta_i = I(D_i < C_i)$, where $I(\cdot)$ is the indicator function. The observed data structure is $(\mathbf{t}_{ij}, T_i)$ with $\mathbf{t}_{ij} = (t_{ij1}, \ldots, t_{ijn_{ij}})'$.

For simplicity of illustration, we consider a single nonparametric covariate function in the hazard functions of each type of recurrent events and the terminal event. Additional nonparametric covariate functions can be incorporated in a similar fashion. The joint model with nonparametric covariate functions can be defined as

$$r_{ij}(t|b_{ij})=r_{0j}(t)\exp\left(f_j(Z_i)+\boldsymbol{X}_i'(t)\boldsymbol{\beta}_j+b_{ij}\right),$$
$$\lambda_i(t|\boldsymbol{b}_i)=\lambda_0(t)\exp\left(f_0(Z_i)+\boldsymbol{X}_i'(t)\boldsymbol{\beta}_0+\sum_{j=1}^{J}\delta_j b_{ij}\right), \quad (1)$$

where $r_{01}(t)$, …, $r_{0J}(t)$, and $\lambda_0(t)$ are baseline hazard functions for all types of recurrent events and the terminal event, respectively, $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J$ are the regression coefficient vectors associated with the covariate vector $\boldsymbol{X}_i'(t)$. The covariate vector could be different in the hazard functions of different events and could be time-dependent. The covariate Z_i has nonlinear effects on the hazard functions of the recurrent events and terminal event. And, $f_0(\cdot), f_1(\cdot), \ldots, f_J(\cdot)$ are unspecified (nonparametric) smooth functions whose functional forms are of interest. The random effects vector $\boldsymbol{b}_i = (b_{i1}, \ldots, b_{iJ})'$ is assumed to follow a J-dimensional multivariate normal distribution with mean 0 and covariance matrix $\boldsymbol{\Sigma}$, whose off-diagonal entries govern the correlation among risks of multi-type recurrent events. Specifically, positive entry $\sigma_{jj'}$ of matrix $\boldsymbol{\Sigma}$ indicates positive correlation between the risks of type j and type j' recurrent events. The vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_j, \ldots, \delta_J)'$ incorporates the correlation between time to all-cause mortality and times to multi-type recurrent events. Indeed, a positive parameter δ_j indicates that the patients with higher risk of type j recurrent events tend to have all-cause mortality earlier.

The nonparametric covariate functions in survival analysis have been widely investigated in the literature[29–31]. In Eq (1), the nonparametric functions can be modeled by penalized splines. The crucial problem is how to select the number and locations of knots. In general, a large number of knots can ensure satisfactory fit and can capture the variability of the data, but can also dramatically increase the computational overhead. Thus, the idea of using fewer spline knots has been discussed, e.g., O'Sullivan[32], and Kelly and Rice[33]. Higher degree basis functions may need smaller number of knots than the linear spline basis functions. To this end, Ruppert and Wand[34] suggested that quadratic or cubic regression splines can fit the data better than the linear splines when the number of knots is small. However, the columns of the model matrix for cubic splines tend to be highly correlated because each column is a transformed version of z, which can induce considerable collinearity[35], resulting in a nearly singular model matrix and imprecision in the spline fit[34]. In comparison, the cubic B-spline basis rescales the cubic spline basis, and is remarkably more stable in numerical computation. Therefore, we adopt the cubic B-spline basis and denote the spline model as $f_j(z)=\sum_{k=1}^{m}u_{jk}B_k(z)$, where $\boldsymbol{B}(z) = (B_1(z), B_2(z), \ldots, B_m(z))$ is the vector of rescaled spline basis as $(z, z^2, z^3, (z-\kappa_1)_+^3, \ldots, (z-\kappa_K)_+^3)$, $m = K+3$, $(z-\kappa_k)_+ = z-\kappa_k$ if $z > \kappa_k$ and 0 otherwise. Gray[36] suggested to use 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.

Moreover, we approximate the baseline hazard functions using piecewise constant functions, which have been frequently used in survival and event history data analysis[37,38]. Lawless and Zhan[39] demonstrated that models with piecewise constant baseline hazard functions using 8 to 10 intervals often yield excellent estimation for fixed and random effects

parameters. More precisely, we divide the follow-up time into S intervals where each interval has $1/S$ quantile of the distinct terminal event times, denoted by $0=Q_0^D<Q_1^D<\ldots<Q_S^D$. Then, the piecewise constant baseline hazard function for the terminal event can be expressed as $\lambda_0(t)=\sum_{s=1}^{S}\lambda_{0s}I(Q_{s-1}^D<t\leq Q_s^D)$. Similarly, we can define the piecewise constant baseline hazard function for each type of recurrent events.

2.3 Likelihood Formulation

For notational ease, we denote the vector of B-spline coefficients for event type j as $\boldsymbol{u}_j=(u_{j1},\ldots,u_{jm})$ and denote the corresponding smoothing parameter as ζ_j, for $j=0,1,\ldots,J$, with $j=0$ denoting the terminal event. To avoid over-fitting by splines, we impose the penalty based on finite differences of the coefficients of adjacent B-splines[40]. This type of penalty is a discrete approximation to the traditional integrated square of the finite derivation and it reduces the computation complexity. Define the coefficients for the first order differences as $\Delta u_{jl}=u_{jl}-u_{jl-1}$, the penalized log-likelihood conditional on the random effects vector $\boldsymbol{b}_i$ is

$$l(\boldsymbol{\theta},\boldsymbol{u};\cdot)=\sum_{i=1}^{I}\sum_{j=1}^{J}l_{ij}^R(\boldsymbol{\theta};b_{ij},\boldsymbol{t}_{ij},T_i,Z_i,\boldsymbol{X}_i(t))+\sum_{i=1}^{I}l_i^D(\boldsymbol{\theta};\boldsymbol{b}_i,Z_i,\boldsymbol{X}_i(t),T_i,\Delta_i)-\sum_{j=0}^{J}\zeta_j\sum_{l=2}^{m}(\Delta u_{jl})^2, \quad (2)$$

where

$$\begin{aligned}
l_{ij}^R(\boldsymbol{\theta};b_{ij},\boldsymbol{t}_{ij},T_i,Z_i,\boldsymbol{X}_i(t)) &= \sum_{r=1}^{n_{ij}}\left\{\log\left[\sum_{s=1}^{S}r_{0js}I(Q_{s-1}^{(j)}<t_{ijr}\leq Q_s^{(j)})\right]+\boldsymbol{B}_j(Z_i)\boldsymbol{u}_j+\boldsymbol{X}_i'(t_{ijr})\boldsymbol{\beta}_j+b_{ij}\right\}\\
&- \sum_{s=1}^{S}\int_{Q_{s-1}^{(j)}}^{Q_s^{(j)}\wedge T_i}r_{0js}\exp\{\boldsymbol{B}_j(Z_i)\boldsymbol{u}_j+\boldsymbol{X}_i'(t)\boldsymbol{\beta}_j+b_{ij}\}\mathrm{d}t,\\
l_i^D(\boldsymbol{\theta};\boldsymbol{b}_i,Z_i,\boldsymbol{X}_i(t),T_i,\Delta_i) &= \Delta_i\left\{\log\left[\sum_{s=1}^{S}\lambda_{0s}I(Q_{s-1}^D<T_i\leq Q_s^D)\right]+\boldsymbol{B}_0(Z_i)\boldsymbol{u}_0+\boldsymbol{X}_i'(T_i)\boldsymbol{\beta}_0+\sum_{j=1}^{J}\delta_j b_{ij}\right\}\\
&- \sum_{s=1}^{S}\int_{Q_{s-1}^D}^{Q_s^D\wedge T_i}\lambda_{0s}\exp\{\boldsymbol{B}_0(Z_i)\boldsymbol{u}_0+\boldsymbol{X}_i'(t)\boldsymbol{\beta}_0+\sum_{j=1}^{J}\delta_j b_{ij}\}\mathrm{d}t,
\end{aligned}$$

where the parameter vector $\boldsymbol{\theta}=(\boldsymbol{\beta}_0,\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_J,\delta_1,\ldots,\delta_J,\boldsymbol{\Sigma})$ and $l\wedge m=\min(l,m)$.

3 Bayesian Inference

3.1 Prior Specification

In the classical approach, the selection of smoothing parameters via cross validation is difficult to implement and often fails when the number of smooth functions is large[41]. Lang and Brezger[41] derived the Bayesian B-spline which replaced the penalties by random walk prior distributions. To do this, the prior distributions for the vector of B-spline coefficients $\boldsymbol{u}_j$ can be expressed as $u_{j1}\sim\text{Uniform}(-10,10)$ and $u_{jl}|u_{jl-1}\sim N(u_{jl-1},\tau_j)$, where $j=0,\ldots,J$ and

$l = 2, \ldots, m$, and the variance parameter τ_j controls the amount of smoothness of spline functions, and it corresponds to the smoothing parameter ζ_j in model (2). We augment the parameter vector as $\boldsymbol{\theta} = (\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J, \delta_1, \ldots, \delta_J, \boldsymbol{\Sigma}, \tau_0, \tau_1, \ldots, \tau_J)$. We impose the prior distribution $\tau_j \sim$ Gamma(0.01, 0.01), which is a non-informative prior distribution with mean being 1 and variance being 100.

Because no prior information of the model parameters are available, we specify non-informative prior distributions for all unknown parameters. A commonly-used prior distribution for the covariance matrix $\boldsymbol{\Sigma}$ is a inverse-Wishart distribution because it is a conjugate prior distribution. However, the inverse-Wishart distribution uses a single parameter to control the precision of all entries in matrix $\boldsymbol{\Sigma}$. Thus, using the standard "noninformative" inverse-Wishart prior distribution is not essentially noninformative[42]. To this end, we reparameterize the covariance matrix $\boldsymbol{\Sigma}$ using the Cholesky decomposition. This approach is widely used in survival and longitudinal data analysis, and yields satisfactory results[12,37]. Specifically, we let $\boldsymbol{\Sigma} = \boldsymbol{\Psi}\boldsymbol{\Psi}'$, where $\boldsymbol{\Psi}$ is a lower triangular matrix with real and positive diagonal entries. Let $\boldsymbol{\phi}_{lm}$ be the (l,m)th entry of $\boldsymbol{\Psi}$ for $1 \le m \le l \le J$ and let the vector $\boldsymbol{a}_i = (a_{i1}, \ldots, a_{iJ})'$ follow a standard normal distribution. Then the random effects vector $\boldsymbol{b}_i = \boldsymbol{\Psi}\boldsymbol{a}_i$ has mean 0 and variance $\boldsymbol{\Sigma}$. Moreover, the entries of matrix $\boldsymbol{\Sigma}$ can be expressed as $\sigma_{lm} = \sum_{k=1}^{l \wedge m} \phi_{lk}\phi_{mk}$, where $1 \le l,m \le J$ and $l \wedge m = \min(l,m)$. We impose Uniform(0, 20) prior distribution on the diagonal entries $\boldsymbol{\phi}_{kk}$, $k = 1, \ldots, J$, to ensure non-negativity and Uniform(−10, 10) prior distribution on the off-diagonal entries of Ψ to allow for possible negative correlation. We assign the $N(0, 100)$ prior distribution for covariate effects $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_J$ and for the parameters $\delta_1, \ldots, \delta_J$. The baseline hazard functions for the recurrent and terminal events are divided into 10 pieces and the baseline hazard within each piece has a prior distribution Gamma(0.01, 0.01). We have investigated other selections of non-informative prior distributions and have obtained very similar results in both the simulation studies and the application to the ALLHAT-LLT study.

3.2 Model Fitting

To obtain the estimates of the parameter vector, we use Bayesian inference based on Markov Chain Monte Carlo (MCMC) to sample from the joint posterior distribution. The MCMC sampler is implemented using BUGS language[43] by specifying the log-likelihood function (2) and the prior distributions of all unknown parameters. 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 disperse initial values and compute the Gelman-Rubin scale reduction statistics ($\hat{R}$) to ensure $\hat{R}$ of all parameters smaller than 1.1[44]. 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 methodology, we include a sample BUGS code in the Web Supplement.

3.3 Model Selection Criteria

Among the various model selection methods available in Bayesian inference, we select the deviance information criterion (DIC), expected Akaike information criterion (EAIC), expected Bayesian information criterion (EBIC)[45], and log pseudo-marginal likelihood

(LPML). The deviance information criterion (DIC) assesses model fit based on the posterior mean of the deviance and a penalty on the model complexity[46]. Because of the mixture framework in our model, we use the DIC_3 measurement[47]. The DIC_3 is defined as $DIC_3=\overline{D(\boldsymbol{\theta})}+\tau_D$, where $\overline{D(\boldsymbol{\theta})}=-2\mathrm{E}_{\boldsymbol{\theta}|\mathscr{D}}\{\log[\prod_{i=1}^{I}f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})]\}$ is the posterior mean deviance, $\tau_D=\overline{D(\boldsymbol{\theta})}+2\log\{\mathrm{E}_{\boldsymbol{\theta}|\mathscr{D}}[\prod_{i=1}^{I}f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})]\}$ is a measure of the effective number of parameters in the model, and $\mathrm{E}_{\boldsymbol{\theta}|\mathscr{D}}(.)$ is the expectation with respect to the joint posterior distribution $\pi(\boldsymbol{\theta}|\mathscr{D})$. Thus, we have $DIC_3=-4E_{\boldsymbol{\theta}|\mathscr{D}}\{\log[\prod_{i=1}^{I}f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})]\}+2\log\{\mathrm{E}_{\boldsymbol{\theta}|\mathscr{D}}[\prod_{i=1}^{I}f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})]\}$. Applying Monte Carlo approximation,

$$\widehat{DIC}_3=-\frac{4}{M}\sum_{m=1}^{M}\sum_{i=1}^{I}\log\{f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta}^{(m)})\}+2\log\{\frac{1}{M}\sum_{m=1}^{M}\prod_{i=1}^{I}f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta}^{(m)})\}.$$

A smaller value of DIC_3 indicates a better-fitting model. The EAIC and EBIC can be estimated as $\widehat{EAIC}=\overline{D(\boldsymbol{\theta})}+2\nu$ and $\widehat{EBIC}=\overline{D(\boldsymbol{\theta})}+\nu\log I$, where ν is the number of parameters in the model, and I is the number of subjects. Smaller values of EAIC and EBIC indicate a better-fitting model.

Conditional predictive ordinate (CPO)[45] is a cross-validation predictive method that evaluates the predictive distribution of the model conditioning on the data but with one subject deleted. The CPO for subject i is defined as $CPO_i=\int f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathscr{D}^{(-i)})d\boldsymbol{\theta}=\{\int\frac{\pi(\boldsymbol{\theta}|\mathscr{D})}{f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta})}d\boldsymbol{\theta}\}^{-1}$, where $\mathscr{D}$ is the full data, $\mathscr{D}^{(-i)}$ is the data with subject i deleted. In the absence of a closed form, a Monte Carlo approximation of CPO_i can be obtained using the harmonic-mean approximation[48] as $\widehat{CPO}_i=\{\frac{1}{M}\sum_{m=1}^{M}\frac{1}{f(\boldsymbol{t}_{ij},T_i|\boldsymbol{\theta}^{(m)})}\}^{-1}$, where $\boldsymbol{\theta}^{(m)}$ is the mth sample of parameter vector $\boldsymbol{\theta}$ after burn-in. A larger CPO value indicates a better model fit. The summary statistic of the CPO is the log pseudo-marginal likelihood (LPML), defined as $\mathrm{LPML}=\sum_{i=1}^{I}\log(\widehat{CPO}_i)$. A larger value of LPML suggests better model fit.

4 Simulation Study

We conduct a simulation study with two settings to compare the performance of three models: joint model (Eq (1)), reduced model (the recurrent and terminal events are modeled independently with $\delta_j = 0$ for $j = 1, \ldots, J$ in Eq (1)), and parametric model (model the nonparametric covariate functions in Eq (1) as linear functions). In each simulation setting, we generate 400 datasets with 500 subjects in each dataset. For each subject, we simulate two types of recurrent events and a terminal event. The hazard functions of the multi-type recurrent events and terminal event are defined as follow:

$$\begin{aligned} r_{i1}(t) &= r_{01}(t)\exp(f_1(Z_i)+X_i\beta_1+b_{i1}), \\ r_{i2}(t) &= r_{02}(t)\exp(f_2(Z_i)+W_i\beta_2+b_{i2}), \\ \lambda_{ij}(t) &= \lambda_0(t)\exp(Z_i\alpha_0+X_i\alpha_1+W_i\alpha_2+\delta_1 b_{i1}+\delta_2 b_{i2}), \end{aligned}$$

where the baseline hazard functions follow Weibull distributions with $r_{01}(t) = 0.2t^{0.5}$, $r_{02}(t) = 0.1t$, and $\lambda_0(t) = 0.02t$. The nonlinear smooth functions are $f_1(Z) = (\sin(3Z) + Z^2)/5$, and $f_2(Z) = \cos(4Z)/5 + Z^2/3$. Note that variable Z is linear in the terminal event model to be consistent with our data application because the preliminary analysis of the ALLHAT-LLT dataset does not support nonparametric age function as detailed in Section 5. We generate variables X and W as binary variables taking values 0 and 1 with probability 0.5, and generate variable Z from the uniform distribution between −2 and 2. The regression coefficients are $\beta_1 = 1$, $\beta_2 = 1.5$, $\alpha_0 = 0.5$, $\alpha_1 = -1$, and $\alpha_2 = 2$. We assume the random effects vector $b_i = (b_{i1}, b_{i2})' \overset{iid}{\sim} N(0, \sum)$ where the covariance matrix $\Sigma = \{(0.64, 0.32), (0.32, 1.6)\}$. The censoring time is generated as $C_i = 15 + \text{uniform}(0, 6)$.

We use the statistical inference procedure described in Section 3 to estimate the unknown parameters in the model. We set the number of pieces to be 10 in all baseline hazard functions and set the number of knots to be 10 in all nonparametric functions. We have used different numbers of knots and pieces and have obtained very similar results. In each simulation setting, we run two parallel MCMC chains with over-dispersed initial values. Each chain is run for 10, 000 iterations. The first 5, 000 iterations are discarded as burn-in, and the remaining 5, 000 samples are used to obtain the posterior distributions of the parameters. The computation time of all three models is approximately 25 minutes. We compute bias (the average of posterior means minus the true values), the standard deviation of the posterior means (SD), the square root of the average of the posterior variance (SE), and coverage probability (CP) of 95% equal-tail credible intervals. Table 2 displays the simulation results of both settings (upper and lower table for simulation settings I and II, respectively).

In simulation setting I, there is no correlation between the multi-type recurrent events and terminal event ($\delta_1 = \delta_2 = 0$), and the percentage of censoring is around 20%. The reduced model is the true model in this setting. The simulation results suggest that both the joint and reduced models generate comparable results, with small bias, SE close to SD, and coverage probabilities close to the nominal value of 95%. Under model overparameterization, the estimates of parameters δ_1 and δ_2 from the joint model are correctly close to zero, suggesting that the joint model is still a reasonable model in this simulation setting. However, by misspecifying the nonparametric covariate functions, the parametric model gives biased estimates and low coverage probabilities for all parameters, except the regression coefficients in the survival sub-model. In Figure 3, we present the joint model estimates (dotdash lines) with 95% pointwise credible intervals (dashed lines) of the nonparametric covariate functions $f_1(Z)$ (panel a) and $f_2(Z)$ (panel b) with the true functions (solid lines), in addition to their coverage probabilities of the 95% pointwise credible intervals (panel c for $f_1(Z)$ and panel d for $f_2(Z)$). Panels a and b suggest that the estimated nonparametric functions by the joint model are reasonably close to the true functions, with

95% pointwise credible intervals always covering the true functions. In panels c and d, the empirical coverage probabilities on all values of Z are close to the nominal level of 95%.

In simulation setting II, there is a positive correlation between the recurrent events and terminal event ($\delta_1 = 0.3$ and $\delta_2 = 0.5$), so that the subjects with higher risks of recurrent events tend to have a terminal event earlier. The joint model is the true model in this setting. The percentage of censoring is also around 20%. The joint model generally provides estimates with negligible bias, SE close to SD, and CP reasonably close to 95%. These results suggest that the joint model can successfully recover the true parameters in the presence of a dependent terminal event and nonparametric covariate functions. In contrast, the reduced model gives severely biased parameter estimates and low coverage probabilities, especially for the regression coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$, because it misspecifies the model for the terminal event. The parametric model provides poor estimation on all parameters including $\boldsymbol{\alpha}$, even though there is no nonparametric covariate function in the hazard function of the terminal event. This is because the parametric model treats the nonparametric functions as linear, and it fails to capture the true between-subject variation in the risks of recurrent events. Figure 4 displays the joint model's estimated nonparametric functions and their coverage probabilities in the simulation setting II. Similar to setting I, the estimated nonparametric function performs well in recovering the true covariate functions.

To justify the use of non-parametric covariate functions, we have added a simulation study to examine whether polynomial regression models can sufficiently estimate the nonlinear covariate functions. Polynomial regression models using higher order polynomials are commonly used to fit a nonlinear relationship between the explanatory variable and response variable. In simulation settings I and II, we adopt cubic polynomial models to estimate the nonlinear smooth functions $f_1(Z)$ and $f_2(Z)$, respectively, using $f_1(Z)=z_i\beta_{11}+z_i^2\beta_{12}+z_i^3\beta_{13}$ and $f_2(Z)=z_i\beta_{21}+z_i^2\beta_{22}+z_i^3\beta_{23}$. The estimated covariate functions from the polynomial regression models are given in Web Figures 1–2. The simulation results suggest that the polynomial regression models are unable to capture all the variation of the true covariate functions and the 95% pointwise credible intervals do not always cover the true functions. Moreover, the empirical coverage probabilities on all values of Z deviate markedly from the nominal level of 95%. These undesirable properties from the polynomial regression models have been reported in the literature[49,50].

In conclusion, the simulation results suggest that in the scenario of independent termination, the joint model provides results comparable to the reduced model, and both outperform the parametric models with both linear functions and polynomials. In the presence of a dependent terminal event and nonparametric covariate functions, the joint model provides more accurate parameter estimates than the reduced model and the parametric models with both linear functions and polynomials.

5 Application to ALLHAT-LLT Data

In this section, we apply all three models to the motivating ALLHAT-LLT study. We consider three types of recurrent events (CHD, stroke, and heart failure) and a terminal event (all-cause mortality). We use two parallel chains with overdispersed initial values and run

each chain for 20,000 iterations. The first 10,000 iterations are discarded as burn-in and the parameter estimates are based on the remaining 10,000 iterations from each chain. For baseline hazard functions, we use piecewise constant functions with 10 intervals by every 1/10th quantile. The nonparametric functions of age in the hazard functions of the three types of recurrent events are estimated by Bayesian B-splines with 10 knots. Good mixing properties of the MCMC chains are observed in the trace plots.

Previous studies suggest a number of risk factors for the cardiovascular disease, including older age, male gender, hypertension, diabetes, tobacco use, excessive alcohol consumption, excessive sugar consumption, family history of cardiovascular disease, obesity, lack of physical activity, air pollution, among others[26,51]. Based on the data availability, we include the following covariates in the models for recurrent events and all-cause mortality: lipid-lowering trial randomization group (llt=0 for pravastatin, 1 for usual care), gender (gender=0 for male, 1 for female), race (race=0 for white and others, 1 for black), diabetes (diabetes=0 for yes, 1 for no), history of CHD (lchd=0 for yes, 1 for no), cigarette smoker (cursmoke=0 for current smoker, 1 for past smoker, 2 for never smoker), aspirin (aspirin=0 for yes, 1 for no), antihypertensive treatment before trial (blmeds=0 for yes, 1 for no), antihypertensive randomization group (att=0 for doxazosin, 1 for chlorthalidone, 2 for amlodipine, 3 for lisinopril), age, body mass index (bmi), systolic blood pressure (sbp), diastolic blood pressure (dbp), high-density lipoprotein cholesterol (hdl), and low-density lipoprotein cholesterol (ldl).

To investigate the linearity of the age effects on the multi-type recurrent events and all-cause mortality, we conduct a preliminary analysis for all these events separately using coxph function with option cluster (to analyze the recurrent events) and pspline (to use a penalized spline basis) in R package survival [31]. Table 3 compares the models with either linear or nonparametric age functions using the Akaike information criterion (AIC). All recurrent event hazard models with nonparametric age functions have smaller AIC than their counterparts with linear age functions, suggesting that nonparametric age functions are preferable in modeling all three types of recurrent events. However, the mortality hazard model with linear age function has smaller AIC compared to the counterpart with nonparametric age function, suggesting that linear age function is more appropriate in modeling the hazard of mortality. Therefore, we include nonparametric functions of age in the hazard models of the multi-type recurrent events.

To assess the performance of all three models (the joint, reduced, and parametric models), we use the model selection criteria discussed in Section 3.3. Table 4 presents DIC_3, EAIC, and EBIC, LPML values for all three models. The results suggest that the joint model outperforms the other two models in all criteria with smaller DIC_3, EAIC, EBIC, and larger LPML, and it is selected as the final model. Table 5 compares the posterior mean, standard deviation (SD), and 95% equal-tail credible intervals from all three models. We find that there is no significant difference between usual care and pravastatin treatment in the risks of CHD (HR= exp(0.128) = 1.137; 95% CI [0.957, 1.353]), stroke (HR= 1.106; 95% CI [0.894, 1.376]), heart failure (HR= 0.953; 95% CI [0.766, 1.196]), and all-cause mortality (HR= 1.025; 95% CI [0.815, 1.303]). The results are consistent with the final ALLHAT-LLT article[25]. The results in Table 5 also indicate the significant effects of diabetes. For example,

diabetes significantly increases the rates of CHD (HR= 1.721; 95% CI [1.406, 2.081]), stroke (HR= 1.697; 95% CI [1.323, 2.162]), heart failure (HR= 2.162; 95% CI [1.718, 2.716]), and all-cause mortality (HR= 2.275; 95% CI [1.721, 3.013]), which is also consistent with the literature[52]. The complete lists of covariate effects are given in Web Tables 1–4.

Moreover, Table 5 displays the estimates of the entries in the covariance matrix of random effects $\boldsymbol{b}_i$. Specifically, the random effects variances are $\hat{\sigma}_1 = 2.117$ (95% CI [1.969, 2.269]), $\hat{\sigma}_2 = 2.036$ (95% CI [1.852, 2.229]), and $\hat{\sigma}_3 = 2.456$ (95% CI [2.278, 2.643]), for the recurrent events of CHD, stroke, and heart failure, respectively. Conditional on the observed risk factors, we observe significant positive correlation between the risks of CHD and stroke events ($\hat{\rho}_{21} = 0.864$, 95% CI [0.792, 0.923]), between the risks of CHD and heart failure events ($\hat{\rho}_{31} = 0.921$, 95% CI [0.881, 0.952]), and between the risks of stroke and heart failure events ($\hat{\rho}_{32} = 0.708$, 95% CI [0.625, 0.786]). This phenomena suggests that subjects with one type of cardiovascular disease events (CHD, stroke, or heart failure) are very likely to experience another type of cardiovascular disease events than those without. Moreover, the risk of having recurrent CHD events is positively associated with the hazard of all-cause mortality ($\hat{\delta}_1 = 1.671$, 95% CI [0.943, 2.680]), conditional on the observed risk factors. It suggests that subjects with CHD events tend to die earlier. However, the other two types of recurrent events (stroke and heart failure) do not show significant correlation with all-cause mortality in this study.

To visualize the nonparametric covariate functions, Figure 5 displays the estimated covariate functions of age (solid lines) in the risks of CHD (left panel), stroke (middle panel), and heart failure (right panel) from the joint model, along with the 95% pointwise credible intervals (dashed lines). The results suggest that risks of cardiovascular disease events (CHD, stroke, and heart failure) increase as age increases, but not in a linear fashion. In general, age increases the risks of cardiovascular disease events more rapidly at older ages than at younger ages. On average, the hazard ratio of CHD for one year increase in age is 1.048 (95% CI [1.029, 1.064]) when age is less than 80 years. After that, the hazard ratio increases to 1.124 (95% CI [1.043, 1.190]). For stroke, the hazard ratio for one year increase in age increases rapidly when age is larger than 65 years (HR= 1.025, 95% CI [0.993, 1.080] for age ≤ 65; HR= 1.106, 95% CI [1.046, 1.124] for age > 65). For heart failure, the hazard ratio for one year increase in age increases rapidly when age is larger than 85 years (HR= 1.076, 95% CI [1.052, 1.094] for age ≤ 85; HR= 1.315, 95% CI [1.082, 1.520] for age > 85). The phenomena of cardiovascular disease event risks increasing nonlinearly as age growth is also reported in some epidemiology studies[20,28].

6 Discussion

In this article, we propose a multivariate frailty model to jointly analyze the multi-type recurrent events and dependent terminal event with nonparametric covariate functions. The recurrent events and dependent terminal event are linked together via shared random effects, which represent the subject-specific heterogeneity in the hazard functions. To model the nonparametric covariate functions, we use the B-spline smooth functions with the penalty terms being replaced by random walk prior distributions. Through these model specification

with fairly small number of parameters in piece-wise constant functions and B-spline basis function, satisfactory results can be obtained for a moderate to large data sets. The simulation study indicates that in the scenario of independent termination, the joint model provides results comparable to the reduced model, and both outperform the parametric model. In the presence of a dependent terminal event and nonparametric covariate functions, the joint model provides more accurate parameter estimates than the reduced model and the parametric model. In the analysis of the ALLHAT-LLT study, the joint model has a better fit than the reduced model and parametric model. We have identified insignificant pravastatin treatment effects, significant diabetes effects, and nonlinear age effects on the risks of CHD, stroke, and heart failure, but insignificant pravastatin treatment effect, significant diabetes effect, and linear age effect on the hazard of all-cause mortality. Moreover, we conclude that the risk of one type of cardiovascular disease event is positively correlated to other types of cardiovascular disease events, and the risk of CHD is positively correlated to the all-cause mortality. For simplicity of illustration, we include only one nonparametric covariate function for each type of recurrent events. Additional nonparametric covariate functions can be incorporated using the same approach. The Bayesian B-spline we adopt is capable of modeling a moderate to reasonably large number of nonparametric regression functions simultaneously[41].

Our proposed model has some limitations that we view as future research directions. We have chosen a multivariate normal distribution for the random effects vector because it is flexible in modeling the covariance structure within and between various types of recurrent events and it has meaningful interpretation on correlation. Due to these reasons, it has been used in modeling multi-type recurrent event data in several articles[12,19]. In generalized linear mixed models, misspecification of random effects distribution has little impact on the parameters that are not associated with the random effects[53–55]. The impact of random effects misspecification in our joint model framework warrants further investigation. One limitation is the shared random effects assumption, under which the model of dependent terminal event shares the frailty terms with the models of multi-type recurrent events. While this shared random effects model is easy to implement and has low dimension of random effects distribution, it makes a relatively strong assumption about the association between subject-specific heterogeneity. Molenberghs and Verbeke[56] pointed out that this kind of shared parameter model is a special case of the multivariate frailty model, which in the current context assumes the multi-type recurrent events and terminal events are linked by a four-dimensional vector of correlated random effects with a multivariate normal distribution. Moreover, we have excluded the subjects with missing data in covariates. Missing covariate is a common issue in longitudinal data analysis and is an active research area. Ibrahim et al.[57] gives excellent review of common approaches for inference in generalized linear models with missing covariate data. How to address the missing covariate issue in the proposed modeling framework is an interesting future research direction.

Dynamic prediction of risk of death using the information from recurrent events has recently become a research topic of significant clinical interest[58]. How to conduct the dynamic prediction in the framework of multi-type recurrent events is an interesting further research topic. As a last note, the risks of future disease events may change after some events occur. For example, the incidence of first stroke may elevate the risks of future stroke and other

cardiovascular disease events. Several multi-state models have been studied in the semi-competing risk data to model the effect of nonfatal event occurrence on the risk of terminal event[59]. In the context of recurrent events, each event can be considered as one state and the number of states increases as the number of events increases. We would like to assess the event effects in our proposed model as a future research endeavor.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sheng Luo's research was supported in part by the National Institute of Neurological Disorders and Stroke under Award Number R01NS091307 and by the National Center for Advancing Translational Sciences under Award Number KL2-TR000370. 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.

References

1. Fuchs HJ, Borowitz DS, Christiansen DH, Morris EM, Nash ML, Ramsey BW, et al. Effect of aerosolized recombinant human DNase on exacerbations of respiratory symptoms and on pulmonary function in patients with cystic fibrosis. New England Journal of Medicine. 1994; 331(10):637–642. [PubMed: 7503821]
2. Tuli S, Drake J, Lawless J, Wigg M, Lamberti-Pasculli M. Risk factors for repeated cerebrospinal shunt failures in pediatric patients with hydrocephalus. Journal of Neurosurgery. 2000; 92(1):31–38. [PubMed: 10616079]
3. Sacco R, Wolf PA, Kannel W, McNamara P. Survival and recurrence following stroke. The Framingham study. Stroke. 1982; 13(3):290–295. [PubMed: 7080120]
4. Aalen OO. Heterogeneity in survival analysis. Statistics in medicine. 1988; 7(11):1121–1137. [PubMed: 3201038]
5. Lawless JF, Nadeau C. Some simple robust methods for the analysis of recurrent events. Technometrics. 1995; 37(2):158–168.
6. Kelly PJ, Lim LLY. Survival analysis for recurrent event data: an application to childhood infectious diseases. Statistics in Medicine. 2000; 19(1):13–33. [PubMed: 10623910]
7. Vaida F, Xu R. Proportional hazards model with random effects. Statistics in Medicine. 2000; 19(24):3309–3324. [PubMed: 11122497]
8. Ibrahim, JG., Chen, MH., Sinha, D. Bayesian Survival Analysis. Springer; 2005.
9. Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. Journal of the American Statistical Association. 1993; 88(423):811–820.
10. Lin D, Wei L, Yang I, Ying Z. Semiparametric regression for the mean and rate functions of recurrent events. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2000; 62(4):711–730.
11. Velagaleti RS, Pencina MJ, Murabito JM, Wang TJ, Parikh NI, D'Agostino RB, et al. Long-term trends in the incidence of heart failure after myocardial infarction. Circulation. 2008; 118(20): 2057–2062. [PubMed: 18955667]
12. Chen BE, Cook RJ, Lawless JF, Zhan M. Statistical methods for multivariate interval-censored recurrent events. Statistics in Medicine. 2005; 24(5):671–691. [PubMed: 15558580]
13. Cai J, Schaubel DE. Marginal means/rates models for multiple type recurrent event data. Lifetime Data Analysis. 2004; 10(2):121–138. [PubMed: 15293628]
14. Cook, RJ., Lawless, JF. The Statistical Analysis of Recurrent Events. Springer; 2007.

15. Huang X, Wolfe RA. A frailty model for informative censoring. Biometrics. 2002; 58(3):510–520. [PubMed: 12229985]
16. Ghosh D, Lin D. Marginal regression models for recurrent and terminal events. Statistica Sinica. 2002; 12(3):663–688.
17. Chen BE, Cook RJ. Tests for multivariate recurrent events in the presence of a terminal event. Biostatistics. 2004; 5(1):129–143. [PubMed: 14744832]
18. Zhu L, Sun J, Srivastava DK, Tong X, Leisenring W, Zhang H, et al. Semiparametric transformation models for joint analysis of multivariate recurrent and terminal events. Statistics in Medicine. 2011; 30(25):3010–3023. [PubMed: 21786280]
19. Mazroui Y, Mathoulin-Pélissier S, MacGrogan G, Brouste V, Rondeau V. Multivariate frailty models for two types of recurrent events with a dependent terminal event: Application to breast cancer data. Biometrical Journal. 2013; 55(6):866–884. [PubMed: 23929494]
20. Finegold JA, Asaria P, Francis DP. Mortality from ischaemic heart disease by country, region, and age: Statistics from World Health Organisation and United Nations. International Journal of Cardiology. 2013; 168(2):934–945. [PubMed: 23218570]
21. Mensah, GA. The Atlas of Heart Disease and Stroke. World Health Organization; 2004.
22. Yu Z, Liu L. A joint model of recurrent events and a terminal event with a nonparametric covariate function. Statistics in Medicine. 2011; 30(22):2683–2695. [PubMed: 21751230]
23. Ripatti S, Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. Biometrics. 2000; 56(4):1016–1022. [PubMed: 11129456]
24. Davis BR, Cutler JA, Gordon DJ, Furberg CD, Wright JT Jr, Cushman WC, et al. Rationale and design for the antihypertensive and lipid lowering treatment to prevent heart attack trial (ALLHAT). American Journal of Hypertension. 1996; 9(4):342–360. [PubMed: 8722437]
25. ALLHAT O, et al. 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). Journal of the American Medical Association. 2002; 288(23):2998–3007. [PubMed: 12479764]
26. Fuster, V., Kelly, BB., et al. Promoting Cardiovascular Health in the Developing World: a Critical Challenge to Achieve Global Health. National Academies Press; 2010.
27. Mendis, S., Puska, P., Norrving, B., et al. Global Atlas on Cardiovascular Disease Prevention and Control. World Health Organization; 2011.
28. Cheung AK, Sarnak MJ, Yan G, Dwyer JT, Heyka RJ, Rocco MV, et al. Atherosclerotic cardiovascular disease risks in chronic hemodialysis patients. Kidney International. 2000; 58(1): 353–362. [PubMed: 10886582]
29. Grambsch PM, Therneau TM, Fleming TR. Diagnostic plots to reveal functional form for covariates in multiplicative intensity models. Biometrics. 1995; 51(4):1469–1482. [PubMed: 8589234]
30. Hastie T, Tibshirani R. Exploring the nature of covariate effects in the proportional hazards model. Biometrics. 1990; 46(4):1005–1016. [PubMed: 1964808]
31. Therneau, TM. Modeling Survival Data: Extending the Cox Model. Springer; 2000.
32. O'Sullivan F. Fast computation of fully automated log-density and log-hazard estimators. SIAM Journal on Scientific and Statistical Computing. 1988; 9(2):363–379.
33. Kelly C, Rice J. Monotone smoothing with application to dose-response curves and the assessment of synergism. Biometrics. 1990; 46(4):1071–1085. [PubMed: 2085626]
34. Ruppert, D., Wand, MP., Carroll, RJ. Semiparametric Regression. Cambridge University Press; 2003.
35. Keele, LJ. Semiparametric Regression for the Social Sciences. John Wiley & Sons; 2008.
36. Gray RJ. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. Journal of the American Statistical Association. 1992; 87(420):942–951.
37. Luo S. A Bayesian approach to joint analysis of multivariate longitudinal data and parametric accelerated failure time. Statistics in Medicine. 2014; 33(4):580–594. [PubMed: 24009073]

38. Feng S, Wolfe R, Port F. Frailty survival model analysis of the national deceased donor kidney transplant dataset using Poisson variance structures. Journal of the American Statistical Association. 2005; 100(471):728–735.
39. Lawless J, Zhan M. Analysis of interval-grouped recurrent-event data using piecewise constant rate functions. Canadian Journal of Statistics. 1998; 26(4):549–565.
40. Eilers PH, Marx BD. Flexible smoothing with B-splines and penalties. Statistical Science. 1996; 11(2):89–102.
41. Lang S, Brezger A. Bayesian P-splines. Journal of Computational and Graphical Statistics. 2004; 13(1):183–212.
42. O'Malley AJ, Zaslavsky AM. Domain-level covariance analysis for multilevel survey data with structured nonresponse. Journal of the American Statistical Association. 2008; 103(484):1405–1418.
43. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing. 2000; 10(4):325–337.
44. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statistical Science. 1992; 7(4):457–472.
45. Carlin, B., Louis, T. Bayesian Methods for Data Analysis. Chapman & Hall/CRC; 2009.
46. Spiegelhalter D, Best N, Carlin B, Van der Linde A. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B:Statistical Methodology. 2002; 64(4):583–639.
47. Celeux G, Forbes F, Robert C, Titterington D. Deviance information criteria for missing data models. Bayesian Analysis. 2006; 1(4):651–673.
48. Dey DK, Chen MH, Chang H. Bayesian approach for nonlinear random effects models. Biometrics. 1997; 53(4):1239–1252.
49. Magee L. Nonlocal behavior in polynomial regressions. The American Statistician. 1998; 52(1): 20–22.
50. Gelman, A., Imbens, G. Why high-order polynomials should not be used in regression discontinuity designs. National Bureau of Economic Research; 2014.
51. Howard BV, Wylie-Rosett J. 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. 2002; 106(4):523–527. [PubMed: 12135957]
52. Manson JE, Colditz GA, Stampfer MJ, Willett WC, Krolewski AS, Rosner B, et al. A prospective study of maturity-onset diabetes mellitus and risk of coronary heart disease and stroke in women. Archives of Internal Medicine. 1991; 151(6):1141–1147. [PubMed: 2043016]
53. Jacqmin-Gadda H, Sibillot S, Proust C, Molina JM, Thiébaut R. Robustness of the linear mixed model to misspecified error distribution. Computational Statistics & Data Analysis. 2007; 51(10): 5142–5154.
54. Rizopoulos D, Verbeke G, Molenberghs G. Shared parameter models under random effects misspecification. Biometrika. 2008; 95(1):63–74.
55. McCulloch CE, Neuhaus JM, et al. Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. Statistical Science. 2011; 26(3):388–402.
56. Molenberghs, G., Verbeke, G. Models for Discrete Longitudinal Data. Springer; 2005.
57. Ibrahim JG, Chen MH, Lipsitz SR, Herring AH. Missing-data methods for generalized linear models: A comparative review. Journal of the American Statistical Association. 2005; 100(469): 332–346.
58. Mauguen A, Rachet B, Mathoulin-Pélissier S, MacGrogan G, Laurent A, Rondeau V. Dynamic prediction of risk of death using history of cancer recurrences in joint frailty models. Statistics in Medicine. 2013; 32(30):5366–5380. [PubMed: 24030899]
59. Xu J, Kalbfleisch JD, Tai B. Statistical analysis of illness–death processes and semicompeting risks data. Biometrics. 2010; 66(3):716–725. [PubMed: 19912171]

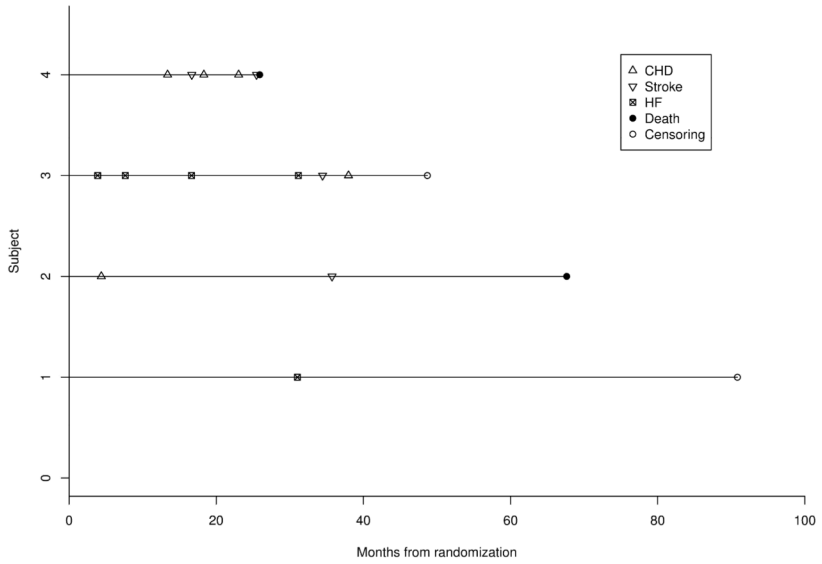

Figure 1.
Time plots of four selected subjects with occurrences of CHD, stroke, heart failure, and death or censoring.

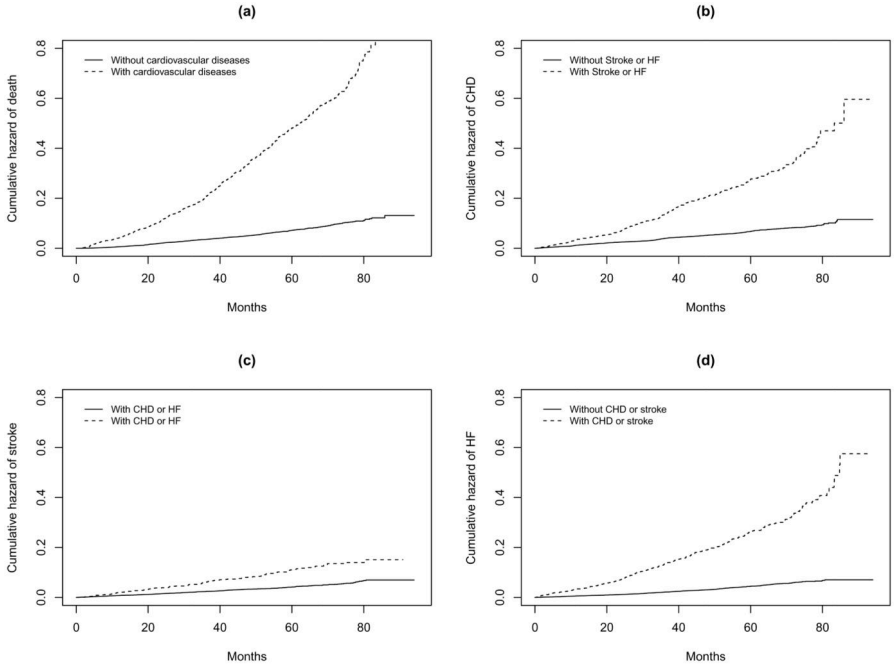

Figure 2.

Cumulative hazard curves. (a) Cumulative hazard of death for subjects with and without the occurrence of cardiovascular disease events (CHD, stroke, and heart failure). (b) Cumulative hazard of CHD for subjects with and without the occurrence of stroke or heart failure. (c) Cumulative hazard of stroke for subjects with and without the occurrence of CHD or heart failure. (d) Cumulative hazard of heart failure for subjects with and without the occurrence of CHD or stroke.

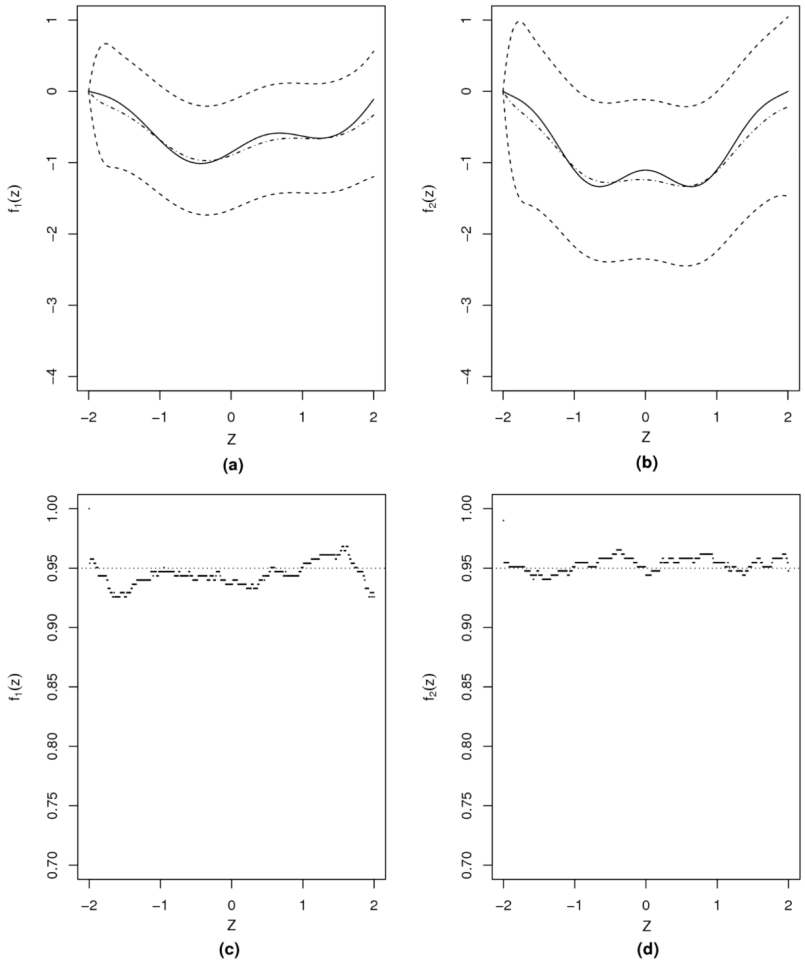

Figure 3.

The joint model's estimates of nonparametric covariate functions in simulation setting I. Penalized spline estimates of $f_1(Z)$ (panel a) and $f_2(Z)$ (panel b). True functions, solid; estimated functions, dotdash; 95% pointwise credible intervals, dashed. Coverage probabilities of the 95% pointwise credible intervals for $f_1(Z)$ (panel c) and $f_2(Z)$ (panel d) with a reference line (dotted) at 0.95.

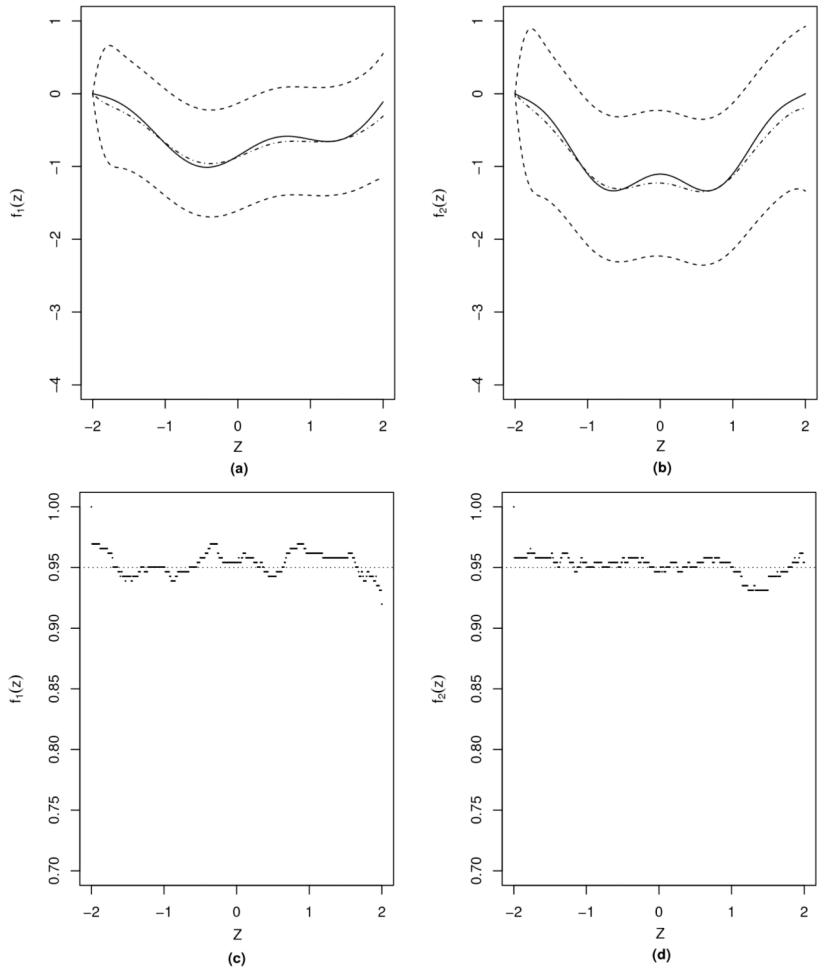

Figure 4.

The joint model's estimates of nonparametric covariate functions in simulation setting II. Penalized spline estimates of $f_1(Z)$ (panel a) and $f_2(Z)$ (panel b). True functions, solid; estimated functions, dotdash; 95% pointwise credible intervals, dashed. Coverage probabilities of the 95% pointwise credible intervals for $f_1(Z)$ (panel c) and $f_2(Z)$ (panel d) with a reference line (dotted) at 0.95.

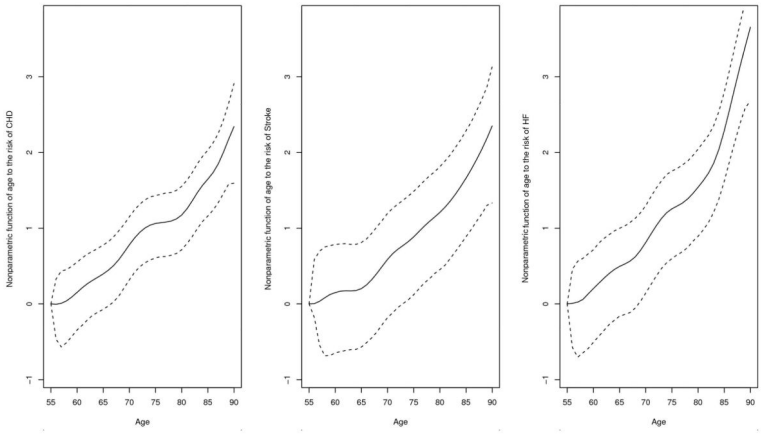

Figure 5.

Estimated nonparametric covariate functions of age in the risks of CHD (left panel), stroke (middle panel), and heart failure (right panel). Estimated functions, solid; 95% pointwise credible intervals, dashed line.

Table 1

Numbers of subjects with CHD, stroke, and heart failure events.

	CHD		stroke		heart failure	
	pravastatin	**usual care**	**pravastatin**	**usual care**	**pravastatin**	**usual care**
1 event	331	349	176	184	154	183
2+ events	29	48	25	30	78	58

Table 2

Results of the simulation study. Setting I: no correlation between the multi-type recurrent events and terminal event ($\delta_1 = \delta_2 = 0$); Setting II: positive correlation between the recurrent events and terminal event ($\delta_1 = 0.3$, $\delta_2 = 0.5$).

Setting I	**Joint Model**				**Reduced model**				**Parametric Model**			
	Bias	**SD**	**SE**	**CP(%)**	**Bias**	**SD**	**SE**	**CP(%)**	**Bias**	**SD**	**SE**	**CP(%)**
For recurrent events												
$\beta_1 = 1$	0.018	0.071	0.072	94.3	0.009	0.084	0.081	95.5	0.017	0.102	0.096	84.4
$\beta_2 = 1.5$	−0.032	0.104	0.110	92.5	−0.030	0.121	0.129	92.3	−0.019	0.138	0.128	87.8
$\rho = 0.316$	0.002	0.043	0.044	93.5	0.006	0.051	0.048	94.3	0.058	0.052	0.049	74.0
$\sigma_1 = 0.8$	0.002	0.030	0.029	96.5	0.000	0.035	0.037	93.0	0.050	0.039	0.037	68.8
$\sigma_2 = 1.265$	0.004	0.044	0.042	94.5	0.002	0.052	0.054	93.8	0.082	0.054	0.055	66.8
For terminal event												
$\alpha_0 = 0.5$	0.001	0.040	0.041	94.8	0.003	0.045	0.046	95.5	0.005	0.046	0.048	94.4
$\alpha_1 = -1$	−0.010	0.090	0.092	93.8	−0.006	0.106	0.103	96.8	−0.020	0.104	0.107	96.0
$\alpha_2 = 2$	0.006	0.109	0.102	96.8	0.016	0.128	0.135	93.5	0.012	0.130	0.129	94.0
$\delta_1 = 0$	−0.036	0.068	0.069	92.0					−0.039	0.079	0.078	92.8
$\delta_2 = 0$	−0.013	0.040	0.042	93.5					−0.012	0.046	0.046	94.4
Setting II	Joint Model				Reduced model				Parametric Model			
	Bias	SD	SE	CP(%)	Bias	SD	SE	CP(%)	Bias	SD	SE	CP(%)
For recurrent events												
$\beta_1 = 1$	0.012	0.090	0.092	94.0	0.035	0.091	0.092	88.8	0.011	0.105	0.091	87.5
$\beta_2 = 1.5$	−0.060	0.121	0.121	92.8	−0.108	0.124	0.132	83.8	−0.067	0.140	0.126	87.3
$\rho = 0.316$	−0.016	0.056	0.054	94.0	−0.020	0.052	0.051	92.5	0.048	0.055	0.052	83.0
$\sigma_1 = 0.8$	−0.005	0.038	0.038	94.0	0.003	0.039	0.038	93.8	0.043	0.040	0.039	78.3
$\sigma_2 = 1.265$	−0.030	0.048	0.046	91.8	0.001	0.051	0.053	93.5	0.062	0.049	0.051	77.3
For terminal event												
$\alpha_0 = 0.5$	−0.017	0.053	0.053	93.0	−0.110	0.043	0.046	30.0	−0.050	0.052	0.048	80.3
$\alpha_1 = -1$	0.007	0.108	0.108	93.5	0.209	0.099	0.109	44.3	0.057	0.105	0.105	91.0
$\alpha_2 = 2$	−0.051	0.138	0.144	91.8	−0.465	0.112	0.121	1.8	−0.147	0.137	0.131	79.0

Setting I	Joint Model				Reduced model				Parametric Model			
	Bias	SD	SE	CP(%)	Bias	SD	SE	CP(%)	Bias	SD	SE	CP(%)
$\delta_1 = 0.3$	−0.022	0.083	0.084	94.0					−0.097	0.081	0.079	73.5
$\delta_2 = 0.5$	−0.032	0.052	0.049	91.8					−0.117	0.048	0.048	32.3

Table 3

AIC for the models with either linear or nonparametric functions of age.

Event type	Linear	Nonparametric
CHD	9822.6	9819.1
Stroke	4937.8	4934.0
Heart failure	7604.4	7592.3
All-cause mortality	20195.4	20224.5

Table 4

Model selection criteria for the ALLHAT-LLT study. DIC_3: deviance information criterion. EAIC: expected Akaike information criterion; EBIC: expected Bayesian information criterion; LPML: log pseudo-marginal likelihood.

Criteria	Joint model	Reduced model	Parametric model
DIC_3	28356.9	33445.8	28443.5
EAIC	26654.5	32261.6	26765.0
EBIC	27525.7	33111.2	27636.2
LPML	−16417.4	−17708.7	−16509.6

Table 5

Parameter estimations from the joint model, reduced model, and parametric model. PM: posterior mean.

	Joint Model				Reduced model				Parametric Model			
	PM	**SD**	**95% CI**		**PM**	**SD**	**95% CI**		**PM**	**SD**	**95% CI**	
For CHD events												
age (years)									0.078	0.007	0.065	0.091
llt (usual care)	0.128	0.088	−0.044	0.302	0.141	0.080	−0.019	0.297	0.130	0.090	−0.044	0.308
gender (female)	−0.599	0.108	−0.811	−0.387	−0.486	0.097	−0.684	−0.302	−0.613	0.107	−0.821	−0.410
diabetes (no)	−0.543	0.099	−0.733	−0.341	−0.411	0.088	−0.578	−0.236	−0.562	0.103	−0.761	−0.358
For stroke events												
age (years)									0.088	0.008	0.072	0.104
llt (usual care)	0.101	0.110	−0.112	0.319	0.105	0.110	−0.109	0.318	0.094	0.111	−0.121	0.315
gender (female)	−0.215	0.129	−0.468	0.029	−0.133	0.130	−0.381	0.125	−0.230	0.126	−0.482	0.009
diabetes (no)	−0.529	0.123	−0.771	−0.280	−0.438	0.122	−0.674	−0.201	−0.551	0.119	−0.778	−0.320
For heart failure events												
age (years)									0.117	0.008	0.101	0.134
llt (usual care)	−0.048	0.113	−0.267	0.179	−0.046	0.119	−0.284	0.189	−0.045	0.115	−0.264	0.174
gender (female)	−0.213	0.135	−0.474	0.052	−0.171	0.140	−0.455	0.106	−0.219	0.140	−0.506	0.046
diabetes (no)	−0.771	0.120	−0.999	−0.541	−0.699	0.125	−0.939	−0.458	−0.772	0.128	−1.022	−0.518
For all-cause mortality												
age (years)	0.155	0.011	0.134	0.177	0.079	0.004	0.071	0.087	0.159	0.012	0.138	0.182
llt (usual care)	0.025	0.119	−0.205	0.265	0.035	0.059	−0.084	0.153	0.020	0.118	−0.208	0.254
gender (female)	−0.726	0.146	−1.013	−0.444	−0.295	0.068	−0.432	−0.163	−0.728	0.146	−1.012	−0.440
diabetes (no)	−0.822	0.141	−1.103	−0.543	−0.381	0.064	−0.507	−0.253	−0.817	0.138	−1.091	−0.547
σ_1	2.117	0.076	1.969	2.269	1.400	0.065	1.276	1.528	2.117	0.078	1.972	2.277
σ_2	2.036	0.096	1.852	2.229	1.667	0.103	1.473	1.869	2.033	0.099	1.834	2.228
σ_3	2.456	0.092	2.278	2.643	2.286	0.094	2.112	2.475	2.467	0.089	2.296	2.644
ρ_{21}	0.864	0.034	0.792	0.923	0.576	0.074	0.424	0.708	0.875	0.035	0.797	0.937
ρ_{31}	0.921	0.018	0.881	0.952	0.986	0.014	0.949	1.000	0.929	0.018	0.893	0.961
ρ_{32}	0.708	0.041	0.625	0.786	0.547	0.062	0.422	0.665	0.712	0.041	0.630	0.788
δ_1	1.671	0.420	0.943	2.680					0.812	0.667	0.748	2.942

	Joint Model				Reduced model				Parametric Model			
	PM	SD	95% CI		PM	SD	95% CI		PM	SD	95% CI	
δ_2	0.364	0.226	−0.176	0.754					0.749	0.790	−0.348	1.656
δ_3	−0.135	0.218	−0.661	0.235					0.102	0.654	−1.023	1.322