Biometrics 64, 950–958
September 2008
DOI: 10.1111/j.1541-0420.2007.00954.x

Analysis of Longitudinal Data in the Presence of Informative Observational Times and a Dependent Terminal Event, with Application to Medical Cost Data

Lei Liu,[1,*] Xuelin Huang,[2] and John O'Quigley[1]

[1]Division of Biostatistics and Epidemiology, Department of Public Health Sciences, University of Virginia, Charlottesville, Virginia 22908-0717, U.S.A.
[2]Department of Biostatistics, M.D. Anderson Cancer Center, University of Texas, Houston, Texas 77030, U.S.A.
[*]*email*: liulei@virginia.edu

Summary. In longitudinal observational studies, repeated measures are often taken at informative observation times. Also, there may exist a dependent terminal event such as death that stops the follow-up. For example, patients in poorer health are more likely to seek medical treatment and their medical cost for each visit tends to be higher. They are also subject to a higher mortality rate. In this article, we propose a random effects model of repeated measures in the presence of both informative observation times and a dependent terminal event. Three submodels are used, respectively, for (1) the intensity of recurrent observation times, (2) the amount of repeated measure at each observation time, and (3) the hazard of death. Correlated random effects are incorporated to join the three submodels. The estimation can be conveniently accomplished by Gaussian quadrature techniques, e.g., `SAS Proc NLMIXED`. An analysis of the cost-accrual process of chronic heart failure patients from the clinical data repository at the University of Virginia Health System is presented to illustrate the proposed method.

Key words: Frailty model; Informative drop-out; Longitudinal medical costs; Piecewise constant baseline hazard; Proportional hazards model; Recurrent marker; Survival analysis.

1. Introduction

Modeling longitudinal medical costs is of great interest in health economics study (Baser et al., 2006). Medical care expenditures have been rising dramatically. At the current rate, the United States cannot afford the medical care of the Medicare-eligible patients within the next 20 years. Analysis of longitudinal medical costs may shed light on when and how the medical spending occurs, which is essential for budgeting. It also allows us to identify populations or groups with the greatest financial need. Modeling longitudinal medical costs plays an important role in the study of health care policy (Cotter et al., 2006). It is also necessary in the cost-effectiveness study (Heitjan, Kim, and Li, 2004).

In this article, we are interested in modeling the cost accrual process, described by the time to recurrent hospital visits and the medical cost for each visit, in the presence of potential death. Such an analysis shows a clear and complete picture of the cost-accumulation process. For example, patients in a severe disease stage often die in a shorter period, so the total medical cost may be less than that of patients in a mild disease stage. The relationship between the total cost and disease stage is sometimes misleading (Bang and Tsiatis, 2002). However, a model of the cost-accrual process in the presence of potential death can reveal a clearer and more intuitive effect of disease stage. Second, the medical cost for each visit is of primary importance in many situations. Patients need to know the actual cost of a hospital visit for budgeting purposes. They can also "shop around" to get the most cost-effective service. Third, the frequency of patients' hospital visits is of interest to hospital administrators and clinicians. They need this information to assess their work load and make corresponding arrangements. For example, hospital administrators would consider recruiting more doctors if the work load exceeds their current capacity. Clinicians would use this information to decide if they could accept new patients.

For diseases with a high mortality rate, it is important to take into account the presence of a dependent death time, often termed "informative censoring." For example, Liu, Wolfe, and Kalbfleisch (2007) showed that ignoring the dependent terminal event would result in biased estimates in modeling longitudinal medical costs. Also, Liu, Wolfe, and Huang (2004) showed disregarding the dependent death leads to biased estimates in the model of intensity of hospital visits.

Analysis of longitudinal medical costs bears much resemblance to the analysis of repeated measures and survival. Such models have been proposed by Faucett and Thomas (1996), Hogan and Laird (1997), Wulfsohn and Tsiatis (1997), Henderson, Diggle, and Dobson (2000), Tsiatis and Davidian (2001, 2004), Xu and Zeger (2001), Ratcliffe, Guo, and Ten Have (2004), Hsieh, Tseng, and Wang (2006), and Vonesh, Greene, and Schluchter (2006). In all the above literature, it is assumed that repeated measures are taken at

noninformative observation times. That is, observational times are independent of or carry no information about repeated measures. This assumption is valid in clinical trials when observation times are fixed, or some observational studies when they are random. However, in many longitudinal observational studies, observation times may be correlated with the repeated measures process. For instance, patients in a severe disease stage visit hospital more often than those in a mild disease stage. Also, their health status measured by biomarkers (e.g., PSA for prostate cancer, CD4 for AIDS, Glomerule filtration rate for end stage renal disease, etc.) at each visit is worse. Therefore, abnormal values of biomarkers are overrepresented and normal values underrepresented, resulting in a selection bias. Such data are sometimes termed "recurrent marker." Similarly, the repeated measures of health care utilization, denoted by days hospitalized or amount of medical cost, are also observed at informative times. Care must be taken to account for this correlation to avoid the selection bias.

Modeling the cost-accumulation process is also closely related to the joint modeling of recurrent hospital visits and a dependent terminal event, e.g., Lancaster and Intrator (1998), Wang, Qin, and Chiang (2001), Liu et al. (2004), Huang and Liu (2007), Rondeau et al. (2007), and Ye, Kalbfleisch, and Schaubel (2007).

On the other hand, when there is no dependent terminal event, Lin and Ying (2001) considered the semiparametric and nonparametric regression analysis of longitudinal data when observation times depend on covariates. Sun et al. (2005) proposed a semiparametric model for longitudinal data with informative observation times and applied their model to the panel count data in a bladder tumor study.

However, by our knowledge, we have not found any literature in the analysis of repeated measures, which accounts for both informative observation times and a dependent terminal event simultaneously. To model such a complicated system, we propose a joint random effects model that includes three submodels: (a) a frailty model for the intensity of recurrent hospital admission times, (b) a random effects model for markers taken at recurrent visits, and (c) a proportional hazards model for death. To take into account the dependence between the frequency of observation times and the amount of repeated measures, random effects are included in submodels (a) and (b). Furthermore, they are incorporated in submodel (c) to illustrate the association of the frequency of hospital visits and the level of repeated measures on survival.

For estimation purpose, we propose to use Gaussian quadrature techniques. The unspecified baseline hazard in the Cox model is approximated by a piecewise constant one. The details of this method were given in Liu and Huang (2007). Because we are mostly interested in the estimates of covariate effects, and the baseline hazard is often treated as a nuisance parameter, such an approximation will not lessen the value of our method. Gaussian quadrature tools, e.g., `SAS Proc NLMIXED` (Littell et al., 2006), can then be conveniently used for estimation. The advantages of this approach include: (1) the implementation (programming) is quite simple; (2) the estimation is accurate; and (3) the standard error estimates can be obtained directly.

The rest of the article is organized as follows. We propose the joint random effects model in the next section. In Section 3, we present the estimation method by Gaussian quadrature and show how to implement it. Simulation studies are conducted in Section 4 to assess the performance of the proposed estimation method. We apply our method to the medical cost-accumulation process for heart failure patients treated at the University of Virginia (UVa) Health System. Concluding remarks are given in Section 6.

2. Model

Denote by T_{ij} the jth informative recurrent observation (e.g., hospital visit) time measured from study onset for subject i, where $i = 1, 2, \ldots, n$, and $j = 1, 2, \ldots, n_i$. Write $N_{ij}(t) = I(T_{ij} \le t)$, where $I(\cdot)$ is the indicator function. Let $N_i(t) = \Sigma_j N_{ij}(t)$ denote the recurrent hospital visit process. At each hospital visit $dN_{ij}(t) = 1$, we observe a repeated measure Y_{ij}, which could be a biomarker (e.g., PSA for prostate cancer, CD4 count for AIDS), or health care utility (e.g., length of hospitalization or medical cost). The follow-up is stopped by $X_i = \min(C_i, D_i)$, the minimum of an independent censoring time C_i and a dependent terminal event (death) time D_i. Assume continuous time for all recurrent hospital visits, death, and censoring events. Denote by $\Delta_i = I(D_i \le C_i)$ the death indicator. We define by $\mathbf{z}_{ij}$ the covariate vector for repeated measure Y_{ij}. For clarification of notation, we write $\mathbf{w}_i^R$ and $\mathbf{w}_i^D$ the covariate vectors for the recurrent hospital visits and death for subject i, respectively, although time-dependent external covariates (Kalbfleisch and Prentice, 2002, p. 196) can be included. Write the intensity (hazard) for the recurrent hospital visits and the terminal event by $r_i(t)$ and $\lambda_i(t)$, respectively. Hereafter we use term "intensity" for both $r_i(t)$ and $\lambda_i(t)$ for ease of description.

We assume repeated measures and recurrent events are correlated through a common random effects u_i, e.g., higher medical costs are associated with a higher intensity of hospital admissions. Further, there is also a random effect term v_i independent of u_i, describing the heterogeneity in repeated measures due to random effects not shared with the rate of hospital visit. We assume that death intensity depends on repeated measures through v_i and recurrent events through u_i. A simple joint model is defined as:

$$r_i(t) = \exp(\mathbf{w}_i^R \beta + u_i) r_0(t), \tag{1}$$

$$y_{ij} | (dN_{ij}(t) = 1) = \mathbf{z}_{ij}\alpha + \gamma_1 u_i + v_i + e_{ij}, \tag{2}$$

$$\lambda_i(t) = \lambda_0(t) \exp(\mathbf{w}_i^D \eta + \gamma_2 u_i + \gamma_3 v_i), \tag{3}$$

where $\{\alpha, \beta, \eta, \gamma_1, \gamma_2, \gamma_3\}$ are unknown parameters, $\lambda_0(t)$ and $r_0(t)$ are baseline intensities for death and recurrent hospital admissions, respectively. Note in (2) the repeated measure is observed only at recurrent hospital visit time $dN_{ij}(t) = 1$. The effect of u_i is different on y_{ij} and $r_i(t)$ due to the inclusion of γ_1 in (2). Similarly, the effects of u_i and v_i are different on death intensity due to γ_2 and γ_3. Assume $e_{ij} \sim^{\text{iid}} N(0, \sigma_e^2)$ is independent of $\{u_i, v_i, C_i, D_i\}$.

Models (1)–(3) are defined for simplicity. More complicated random effects structure, e.g., random slope and other interactions between covariates and random effects, can be incorporated into the model with ease. For example, we can include a random slope in (2) to denote the heterogeneity in the temporal trend of repeated measures. This random effect

can then be incorporated in (3) to describe its effect on survival. Also, both $r_i(t)$ and $\mathbf{w}_i^R$ can be dependent on the order of the recurrent event. For example, an order specific baseline intensity $r_{0j}(t)$ can be used to distinguish baseline intensities of different orders of recurrent events.

The processes modeled by equations (1)–(3) are assumed to be independent of each other given random effects u_i and v_i. The likelihood for the ith subject is

$$L_i = \int\int l_i^A \quad l_i^B \quad l_i^C \quad p(u_i)p(v_i)du_i dv_i, \tag{4}$$

where $p(u_i)$ and $p(v_i)$ are density functions for u_i and v_i, respectively. In this article, we assume $u_i \sim^{\text{iid}} N(0,\, \sigma_u^2)$ and $v_i \sim^{\text{iid}} N(0,\, \sigma_v^2)$ are independent of each other. The detailed form of likelihood l_i^A for recurrent hospital admissions is

$$l_i^A = \prod_{j=1}^{n_i} \left[\exp(\mathbf{w}_i^R\beta + u_i)r_0(t_{ij})\right]^{\delta_{ij}} \times \exp\left\{-\int_0^{x_i} \exp(\mathbf{w}_i^R\beta + u_i)r_0(t)\,dt\right\},$$

where δ_{ij} is the indicator of the patient visiting hospital at time t_{ij}, and x_i is the observed follow-up time.

Likelihood l_i^B is the conditional likelihood for repeated measures at hospital visits, i.e.,

$$l_i^B = \frac{1}{(\sqrt{2\pi}\sigma_e)^{n_i}} \exp\left[-\frac{1}{2\sigma_e^2}\sum_{j=1}^{n_i} e_{ij}^2\right],$$

where

$$e_{ij} = Y_{ij} - \mathbf{z}_{ij}\alpha - \gamma_1 u_i - v_i.$$

The likelihood for death is:

$$l_i^C = \left[\lambda_0(x_i)\exp(\mathbf{w}_i^D\eta + \gamma_2 u_i + \gamma_3 v_i)\right]^{\Delta_i} \times \exp\left[-\int_0^{x_i} \exp(\mathbf{w}_i^D\eta + \gamma_2 u_i + \gamma_3 v_i)\lambda_0(t)dt\right].$$

When observation times are not informative of repeated measures, Wulfsohn and Tsiatis (1997) proposed a shared random effects model of longitudinal data and survival. On the other hand, when no repeated measure is observed at each recurrent event time, Liu et al. (2004) proposed a shared frailty model of recurrent and terminal event times. Apparently, our model generalizes the results of both papers. It is more comprehensive and includes both models as special cases.

3. Estimation

The expectation-maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) is often adopted for estimation in joint random effects models when the likelihood conditional on the random effects can be maximized in a straightforward manner. However, for our joint model, the conditional distribution of random effects given the observed data does not have a closed form. Monte Carlo methods are often needed in the E-step for approximation. However, this makes the estimation highly computationally intensive and the implementation (programming) quite difficult.

An alternative approach is through numerical integration techniques, e.g., Gaussian quadrature. However, Gaussian quadrature techniques cannot be applied directly if $\lambda_0(t)$ and $r_0(t)$ are unspecified. To circumvent this problem, we propose to replace the nonparametric baseline intensities with their piecewise constant counterparts. Lawless and Zhan (1998) and Feng, Wolfe, and Port (2005) demonstrated that models with a piecewise constant baseline intensity with 8–10 intervals often yield satisfactory estimates for fixed effects and frailty. It provides more flexibility over the a priori choices of baseline intensity distribution (e.g., Weibull), while it retains enough model structure. Andersen et al. (1997) also recommended to use a piecewise constant baseline intensity for frailty models in practical applications. Furthermore, Gray (1994) and Matsuyama, Sakamoto, and Ohashi (1998) both advocated a piecewise constant baseline intensity in Bayesian frailty models.

Liu and Huang (2007) applied this estimation method in various frailty proportional hazards models, including the simple frailty model, the joint frailty model of recurrent events and survival (Liu et al., 2004), and that of cluster events with informative dropout (Huang and Wolfe, 2002). By simulation, they showed that it performs similarly and sometimes better than the Monte Carlo expectation-maximization (MCEM) method (Liu et al., 2004) or the penalized partial likelihood (PPL) method (Therneau and Grambsch, 2000) with unspecified baseline intensity. The proposed Gaussian quadrature method is much easier to implement. Also, the standard error estimates can be obtained directly from `SAS Proc NLMIXED`, in contrast to the EM algorithm, which resorts to Louis' formula (Louis, 1982), and the PPL method that relies on resampling methods (Therneau and Grambsch, 2000). Finally, the computational time is improved greatly: for the joint frailty model (Liu et al., 2004), this method is four times faster than the MCEM method (implemented in `R` with `C` code for the Metropolis–Hastings algorithm).

The similar estimation method is adopted for the joint random effects models (1)–(3). We first divide the follow-up period into M_1 intervals (denoted by knots $Q_1^D, Q_2^D, \ldots, Q_{M_1}^D$ and $Q^D{}_0 = 0$ or the smallest failure time) for the observed death times. Denoting the piecewise constant baseline intensity by $\widetilde{\lambda}_0(t)$, we have

$$\widetilde{\lambda}_0(t) = \lambda_{0k}, \text{ for } Q_{k-1}^D < t \le Q_k^D, \text{ where } k = 1, 2, \ldots, M_1$$

or

$$\widetilde{\lambda}_0(t) = \sum_{k=1}^{M_1} \lambda_{0k} I\left(Q_{k-1}^D < t \le Q_k^D\right).$$

The cumulative baseline intensity is

$$\widetilde{\Lambda}_0(t) = \sum_{k=1}^{M_1} \lambda_{0k} \max\left(0, \min\left(Q_k^D - Q_{k-1}^D, t - Q_{k-1}^D\right)\right).$$

Similarly, we can divide the recurrent observation times into M_2 intervals by knots $Q_1^R, Q_2^R, ..., Q_{M_2}^R$, and $Q_0^R = 0$ or the smallest observation time. We have

$$\widetilde{r}_0(t) = \sum_{k=1}^{M_2} r_{0k} I\left(Q_{k-1}^R < t \le Q_k^R\right),$$

and

$$\widetilde{R}_0(t) = \sum_{k=1}^{M_2} r_{0k} \max\left(0, \min\left(Q_k^R - Q_{k-1}^R, t - Q_{k-1}^R\right)\right).$$

Likelihood (4) can then be approximated by replacing $\lambda_0(t)$ and $r_0(t)$ by the piecewise constant counterparts $\widetilde{\lambda}_0(t)$ and $\widetilde{r}_0(t)$, respectively. We thus have a parametric likelihood in the integrand of (4), which can be estimated conveniently by Gaussian quadrature tools in standard statistical packages such as `Proc NLMIXED` in `SAS` (Littell et al., 2006). An introduction to the Gaussian quadrature technique is given in the Appendix. The `SAS` code is available in the Web Supplementary Materials.

Liu and Huang (2007) showed that for a moderate sample size, satisfactory results can be obtained by taking the knots Q_s^R or Q_s^D to be every 10th quantile of recurrent and terminal events, respectively. We will use 10 intervals divided by every 10th quantile in the simulation section. In the application section, we divide the follow-up into intervals of every 2 months.

4. Simulation

In this section, we report results from a simulation study to evaluate the performance of the proposed estimation procedure. The model we adopt is

$$r_i(t) = \exp(Z_{i1}\beta + u_i) r_0(t), \tag{5}$$

$$y_{ij} | (dN_{ij}(t) = 1) = \alpha_0 + Z_{i1}\alpha_1 + t_{ij}\alpha_2 + \gamma_1 u_i + v_i + e_{ij}, \tag{6}$$

$$\lambda_i(t) = \lambda_0(t) \exp(Z_{i1}\eta + \gamma_2 u_i + \gamma_3 v_i). \tag{7}$$

For each submodel, we assume a subject level time-invariant binary covariate Z_{i1} that takes value 0 or 1 with equal probability 1/2. The coefficient β is set to be 1. For repeated measures, there is an additional covariate t_{ij} (the informative observation time) to denote the temporal trend in Y_{ij}. For each subject, we also observe y_{i0} at baseline time $t_{i0} = 0$. Coefficients are set to be $\alpha = (0,\ 1,\ 0.2)^T$. The error term $e_{ij} \sim N(0,\ \sigma_e^2)$ with $\sigma_e^2 = 1$. We include in the model a frailty term u_i from recurrent observation times and a random effect v_i from repeated measures. We assume that random effects $u_i \sim^{\text{iid}} N(0,\ \sigma_u^2)$ and $v_i \sim^{\text{iid}} N(0,\ \sigma_v^2)$ are independent, with $\sigma_u^2 = 1$ and $\sigma_v^2 = 0.5$, respectively. The correlation between recurrent hospital visits and repeated measures is introduced by the shared random effect u_i. The death intensity depends on the subject level covariate Z_{i1} with coefficient $\eta = 1$, and shared random effects u_i and v_i with coefficients γ_2 and γ_3, respectively. We assume

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = \begin{pmatrix} -1.5 \\ -0.5 \\ 1 \end{pmatrix}.$$

To demonstrate the performance of the proposed estimation procedures when the piecewise constant baseline constant intensity is indeed an approximation, the baseline intensities for both recurrent hospital visits and terminal event are chosen to be Weibull distributed with shape 2. We set $r_0(t) = 0.08t$ and $\lambda_0(t) = 0.02t$. The independent censoring time is $6 + \text{Uniform}(0, 6)$. We use sample size $n = 200$ and run 600 replicates. Each subject visits hospital on average 5.8 times; about 33% of total subjects are censored, others experience death.

For estimation, we use 10 intervals (divided by every 10th quantile) to define the piecewise baseline intensities for both recurrent and terminal event processes. We implement the adaptive Gaussian quadrature method with five quadrature points in `SAS Proc NLMIXED`. The `SAS` code is available in the supplementary materials. For each replicate, the estimation process takes around 7 minutes on a personal computer with a 1.70 GHz Intel Pentium M processor and 2 GB RAM. For all replicates, standard error estimates are obtained directly from `SAS` output. Standard error estimates are provided directly from `SAS` output for all replicates. The results are shown in Table 1.

We find that our estimation method yields satisfactory results. The empirical biases of the estimates are very small. The coverage probabilities are close to the nominal level 0.95. There exist only negligible biases for standard error estimates.

For comparison, we fit the above data sets by two reduced models. In reduced model A we assume observation times are noninformative, i.e.,

$$\begin{aligned} y_{ij} &= \alpha_0 + Z_{i1}\alpha_1 + t_{ij}\alpha_2 + v_i + e_{ij}, \\ \lambda_i(t) &= \lambda_0(t) \exp(Z_{i1}\eta + \gamma_3 v_i), \end{aligned}$$

which is the conventional joint random effects model for survival and longitudinal data (e.g., Wulfsohn and Tsiatis, 1997). This is the special case with $\gamma_1 = \gamma_2 = 0$ in our proposed model.

In reduced model B we ignore the dependent terminal event process, i.e.,

$$\begin{aligned} r_i(t) &= \exp(Z_{i1}\beta + u_i) r_0(t), \\ y_{ij} \mid (dN_{ij}(t) = 1) &= \alpha_0 + Z_{i1}\alpha_1 + t_{ij}\alpha_2 + \gamma_1 u_i + v_i + e_{ij}. \end{aligned}$$

The results from these two reduced models are shown in the middle and right panels in Table 1, respectively. We can see that falsely coercing some of γ's to 0 could lead to inadequate results. For example, larger biases and poor coverage probabilities are present for covariate effect estimates of parameters α_0, α_2, η, and γ_3 in reduced model A, and α_2 and σ_v^2 in reduced model B. Thus our model should be preferred in practical data analysis to avoid biases from insufficient models.

We also fit a model with repeated measures only, i.e., we ignore both the informative observation times and the dependent terminal event. This is a special model of reduced model A. The parameter estimates are even worse than those of reduced model A. For example, the parameter estimates for α_0 and α_2 are -0.071 with coverage probability (CP) 93.5% and 0.180 (CP: 60.0%), respectively.

Finally, similar to Liu et al. (2004), we fit a reduced model with no repeated measures, i.e.,

$$\begin{aligned} r_i(t) &= \exp(Z_{i1}\beta + u_i) r_0(t), \\ \lambda_i(t) &= \lambda_0(t) \exp(Z_{i1}\eta + \gamma_2 u_i). \end{aligned}$$

We find that the estimates for η (Bias $= -0.174$, CP $=$ 84.4%) and γ_2 (Bias $= -0.078$, CP $= 88.6\%$) are biased. This fact once again supports our proposed model.

Table 1
Simulation results: parameter estimates

	Proposed model				Reduced model A				Reduced model B			
Parameter	Est	SE	SEM	CP	Est	SE	SEM	CP	Est	SE	SEM	CP
Observation time												
$\beta = 1$	0.979	0.167	0.171	95.0%					0.966	0.168	0.171	95.2%
Repeated measures												
$\alpha_0 = 0$	−0.003	0.177	0.181	95.7%	−0.066	0.177	0.180	93.8%	−0.006	0.176	0.179	95.5%
$\alpha_1 = 1$	1.012	0.243	0.250	95.3%	1.019	0.242	0.249	95.3%	0.990	0.240	0.247	94.7%
$\alpha_2 = 0.2$	0.200	0.011	0.012	94.7%	0.191	0.012	0.012	89.8%	0.188	0.011	0.012	84.0%
$\sigma_e^2 = 1$	0.999	0.044	0.042	92.7%	1.001	0.045	0.042	93.8%	1.002	0.045	0.042	93.2%
Terminal event												
$\eta = 1$	0.991	0.227	0.235	95.8%	0.914	0.204	0.209	93.2%				
Model association												
$\gamma_1 = -1.5$	−1.501	0.106	0.104	94.7%					−1.480	0.106	0.103	93.8%
$\gamma_2 = -0.5$	−0.527	0.125	0.130	96.7%								
$\gamma_3 = 1$	0.969	0.325	0.333	95.0%	0.427	0.070	0.071	0.0%				
Covariance parameters												
$\sigma_u^2 = 1$	1.006	0.150	0.153	94.2%					1.015	0.153	0.153	93.5%
$\sigma_v^2 = 0.5$	0.471	0.129	0.128	91.7%	2.670	0.325	0.314		0.429	0.120	0.118	85.5%

Est is the mean of the parameter estimates (based on 600 replicates); SE is the standard error of the parameter estimates; SEM is the sampling mean of the standard error estimate; CP is the coverage probability of the corresponding 95% confidence interval. Note that σ_v^2 in reduced model A (true value: $\gamma_1^2\,\sigma_u^2 + \sigma_v^2 = 1.5^2 + 0.5 = 2.75$) is not comparable to that in the full model due to different model structure, so the coverage probability is not given.

To investigate the influence of knot's positions (for piecewise constant baseline intensities) on parameter estimates, we also carry out the estimation with 10 intervals of equal duration (1/10 of the longest follow-up time). The resulting parameter estimates are almost indistinguishable from those in Table 1. So the choice of (reasonable) knot's positions should have little impact on parameter estimates.

5. Application

We are interested in modeling medical costs for chronic heart failure (CHF) patients in the presence of informative observation times and a dependent death event. Five million people in the United States are living with heart failure, 70% of whom are at least 60 years old. African Americans and Hispanics have a worse outcome than white patients (Medscape report, 2005). CHF is also one of the most expensive health care problems in the United States, with a projected annual direct and indirect medical cost of $27.9 billion (American Heart Association, 2005), more than doubling that of all cancers combined (O'Connell, 2000).

Our data source is the clinical data repository (CDR) database from the UVa Health System, available online at: `http://cdr.virginia.edu/cdr`. This study includes a total of 1475 patients who were at least 60 years old and first diagnosed and treated in 2004 with heart failure (ICD9 diagnosis code beginning with 428). The follow-up ended with each patient's last hospital admission (up to July 31, 2006), or death date extracted from Death Certificate Data at the Virginia Department of Vital Statistics.

Preliminary studies showed that patients visiting hospital more often tended to pay more for each visit, i.e., the frequency of observation times is informative of the level of medical costs. These patients were also subject to a higher mortality rate. To model such a complicated system, we propose a joint random effects model for the recurrent time of hospital admission, the medical cost for each hospital visit, and death time.

We focus on the analysis of the UVa health system cost—the actual monetary expense of the hospital. To correct the right skewness, we take the log-transformed cost as the response variable. For hospital visits, we use as recurrent event times the offsets (in days) to each patient's first visit for diagnosis or treatment of heart failure.

The mean follow-up time is 18.5 months (SD: 9.0). About 20% of patients died during follow-up, others were censored. The mean age is 72 years old (SD: 8). Males account for 55% of the cohort, whites account for 73%. The medical cost of each hospital visit is highly skewed to the right (median $350 and mean $2670).

In all three parts we adjust for age (centered at 72 years old), gender (*male* = 1, *female* = 0), and race (*white* = 1, *nonwhite* = 0). From preliminary analysis we found that age has a quadratic form on the medical cost so we include the square of centered age in equation (2). Similarly, we also include the quadratic form of the visit time (in months) in the medical cost model. More covariates recorded in the CDR, e.g., insurance type, medical treatment, comorbidities, etc., can be easily added in further studies. For the Gaussian quadrature estimation, we assume piecewise constant baseline intensities for both hospital visits and death in every 2 months interval. This results in 16 intervals for recurrent hospital visits and 12 for death. We apply the Gaussian quadrature method described in Section 3. The estimation process takes around 100 minutes to converge. The parameter estimates are shown in Table 2.

The results in Table 2 show that: (1) There is a significant gender difference in the amount of medical costs: males spent 14% more than females for each hospital visit. (2) White

Table 2
Joint analysis of longitudinal medical costs and survival for heart failure patients

	Our model			Reduced model A			Reduced model B		
	Est	SE	p-value	Est	SE	p-value	Est	SE	p-value
Recurrent hospital visit									
Male	−0.013	0.032	0.67				−0.023	0.031	0.45
White	−0.099	0.036	0.006				−0.087	0.035	0.01
Age	0.126	0.021	<0.0001				0.110	0.020	<0.0001
Cost of each visit									
Intercept	6.267	0.079	<0.0001	6.335	0.079	<0.0001	6.263	0.078	<0.0001
Male	0.129	0.063	0.04	0.120	0.063	0.06	0.119	0.062	0.05
White	−0.247	0.070	0.0005	−0.255	0.070	0.0003	−0.234	0.069	0.0007
Age	−0.096	0.042	0.02	−0.099	0.042	0.02	−0.116	0.042	0.005
Age^2	−0.162	0.047	0.0006	−0.181	0.048	0.0002	−0.148	0.048	0.002
Visit time	−0.031	0.006	<0.0001	−0.031	0.006	<0.0001	−0.037	0.006	<0.0001
Visit time^2	0.0015	0.0002	<0.0001	0.0014	0.0002	<0.0001	0.0016	0.0002	<0.0001
σ_e^2	2.348	0.027		2.346	0.027		2.351	0.027	
Survival									
Male	0.279	0.141	0.05	0.256	0.128	0.05			
White	−0.243	0.155	0.12	−0.218	0.140	0.12			
Age	0.646	0.095	<0.0001	0.532	0.082	<0.0001			
Model association									
γ_1	0.988	0.079	<0.0001				0.922	0.083	<0.0001
γ_2	2.489	0.300	<0.0001						
γ_3	0.398	0.099	<0.0001	0.807	0.073	<0.0001			
Covariance matrix									
σ_u^2	0.241	0.015	<0.0001				0.216	0.013	<0.0001
σ_v^2	0.822	0.047	<0.0001	1.045	0.052	<0.0001	0.821	0.046	<0.0001

Age is centered at 72 years old and in every 10-year unit.

patients tended to visit hospital at a lower rate ($HR = 0.91$, $p = 0.006$). The cost for white patients is lower (−0.247 in log scale, or 22% lower in dollar value, $p = 0.0005$). Both results suggest that white patients had a better outcome than nonwhite patients, which is consistent with the Medscape report (2005). (3) Older patients were more likely to seek medical treatments ($HR = 1.13$ for every 10 years older in age, $p < 0.0001$). However, we find that age effect on hospital cost is quadratic ($p = 0.0006$): it peaks at 69 years old and drops thereafter. This fact may be due to (a) older patients sought medical services more frequently, diluting the medical cost for each visit; (b) older patients with heart disease were often treated less aggressively, resulting in lower medical costs (Gatsonis et al., 1995). (4) Visit time (from onset) has a quadratic effect on medical costs: there is a high initial cost (due to high initial diagnosis and treatment cost); the cost for each visit then decreases and reaches the bottom around 10 months; it increases thereafter until death. This pattern is consistent with the "bath-tub" shape in monthly outpatient erythropoietin (EPO) medical cost data for dialysis patients (Liu et al., 2007).

In the survival model, we find that age is significant: every 10-year increase in age elevates the death rate by 91%. Male patients had a higher mortality rate ($p = 0.05$), so did nonwhite patients.

For random effects, we note heterogeneity exists for both recurrent hospital visits ($\widehat{\sigma}_u^2 = 0.241, p < 0.0001$) and repeated measures of the medical cost ($\widehat{\sigma}_v^2 = 0.822, p < 0.0001$) at each visit. Note here the adjustment by Stram and Lee (1994) should be used to address the boundary testing issue for variance components. For the association between these three models, we notice that all $\widehat{\gamma}$'s are highly significant ($p < 0.0001$). This justifies our joint model of the three endpoints, instead of considering them separately. The results suggest that: (1) a significant association exists between the frequency of hospital admission and the cost of each visit—patients visiting hospital more often tended to have a higher medical cost for each visit; (2) death depends on random effects in both parts: patients who sought medical treatment more frequently and/or incurred more medical cost at each visit had a higher mortality rate.

For comparison, we also fit reduced models A and B for this data set as in the simulation section. We find that in reduced model A where we disregard the informative observation time, there are some moderate changes in parameter estimates. For example, age effect on survival is smaller (0.532 versus 0.646). In reduced model B, both the racial and gender effects on the intensity of hospital visits differ by more than 10% than those in our full model.

To investigate the potential influence of some extremely high medical costs, we also conduct an analysis capping the maximum cost to the 99% percentile (=$35,596) of all costs. The results are very close (not shown), so the right skewness of medical cost data has little effect on parameter estimates in this application.

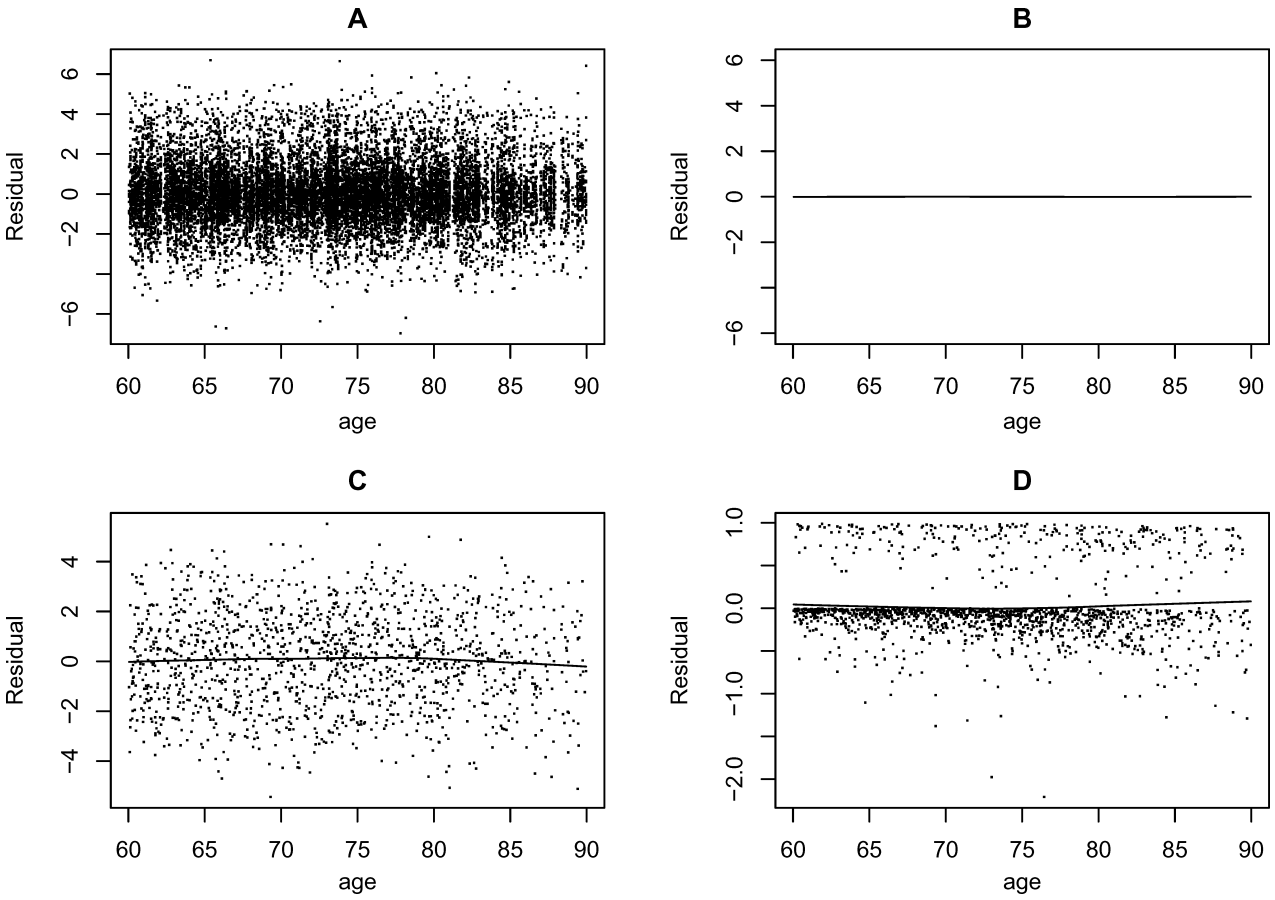

Figure 1. Residual plots for models (1)–(3). (A) Residuals for medical costs; (B) Lowess estimate in plot (A); (C) Martingale residuals for recurrent hospital visits; (D) Martingale residuals for death. Lowess estimates for (C) and (D) are shown in the corresponding plot.

For model checking, residuals from models (1)–(3) can be easily obtained by "predict" statement in `SAS Proc NLMIXED`, with empirical Bayes' estimates provided for the random effects. The empirical Bayes' estimates for random effects are computed as the mode of the posterior distribution given all the data (Y_i, T_{ij}, X_i, Δ_i). Specifically, we use Martingale residuals for submodels (2) and (3). As an example, we show in Figure 1 the residual plots for models (1)–(3) with respect to age. There is no clear pattern in any of the three residual plots as shown by the Lowess estimates.

Finally, as suggested by the associate editor, we evaluate the fit of repeated measures conditionally upon the survival status, using a method proposed by Dobson and Henderson (2003). We split the sample into two groups: censored or dead during follow-up. We then compare the predicted and observed means conditionally on the observed survival status and the follow-up time. The conditional expectation can be calculated by adaptive Gaussian quadrature method, e.g., function *integrate* or *adapt* in `R`. The plot is shown in the Web Supplementary Materials. In the upper panel we find that observed and predicted mean (log)costs agree well for censored patients, whereas in the lower panel there is a minor downward bias in later months for patients who died during follow-up. Note we show the Lowess estimates because medical costs are observed at continuous hospital visit times.

6. Discussion

In this article, we propose a joint random effects model for repeated measures in the presence of informative observation times and a dependent terminal event. The estimation can be carried out conveniently by Gaussian quadrature techniques, as implemented in `SAS Proc NLMIXED`. Simulation results demonstrate the excellent performance of our method in small to moderate samples. We apply our model to the longitudinal medical costs of 1475 CHF patients from the CDR at the UVa.

Two time scales are usually employed to analyze recurrent event data: total time scale measured from the study onset, and gap (episode, recurrence, or interevent) time scale measured between consecutive recurrent events. In this article, we adopt the total time scale. Huang and Liu (2007) proposed a joint frailty model for gap times of recurrent events and survival. It is of future interest to extend their work to a joint model of repeated measures, gap times, and survival.

Gaussian quadrature is efficient for low-dimension integration. Based on our own experience, we find that it performs well for models with $\leq$4 random effects. Importance sampling can be used for somewhat higher dimension, and MCMC for very high dimension (Ruppert, 2005). The implementation of our model in these more complicated settings merits further consideration.

Both Liu and Huang (2007) and this article show that assuming 10 intervals for piecewise constant baseline intensities yields satisfactory results for a moderate sample size. For a small sample size, it is reasonable to use a smaller number of knots (say 5) to maintain model identifiability. For a large sample size, however, if we increase the number of knots too much, we could end up with a lot of parameters to estimate for piecewise constant baseline intensities. This may cause numerical problems in the Gaussian quadrature estimation procedure.

In health services research we are also interested in the accumulative medical cost (or days of hospitalization) in a unit time, e.g., year or month. A different formulation is to model monthly (or annual) medical costs, by summing up the medical cost of each hospital visit by month (or year) (Liu et al., 2007). In this way, some months have zero cost (no hospital visit), and other months have highly skewed monthly cost. Such data are often termed "semicontinuous." A random effects two-part model (Olsen and Schafer, 2001; Tooze, Grunwald, and Jones, 2002) can be used to describe the odds of monthly cost being nonzero and the cost level if a patient seeks medical care during the month. A joint model for longitudinal semicontinuous data and survival can be proposed accordingly.

In this article, we focus on modeling the cost-accrual process. For future research, we are interested in deriving the cumulative or lifetime medical cost from our model. The comparison of our method with those modeling the lifetime medical cost directly is another topic for future research.

7. Supplementary Materials

The `SAS` code and the model checking figure (referenced in Sections 3 and 5, respectively) are available under the Paper Information link at the *Biometrics* website `http://www.biometrics.tibs.org`.

Acknowledgements

We thank the referees and associate editor for their careful reading and valuable comments. We also thank Drs Mei-Cheng Wang, John Kalbfleisch, and Rod Little for helpful discussions, and Dr Jason Lyman, Mr Mac Dent and Mr Ken Scully at the Clinical Data Repository of the University of Virginia for data preparation. This research was supported by AHRQ grant 1 R03 HS016543.

References

Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions*. New York: Dover.

American Heart Association. (2005). *Heart and Stroke Statistical Facts: 2005 Statistical Update*. Dallas, TX: American Heart Association.

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

Bang, H. and Tsiatis, A. (2002). Median regression with censored cost data. *Biometrics* **58,** 643–649.

Baser, O., Gardiner, J. C., Bradley, C. J., Yuce, H., and Given, C. (2006). Longitudinal analysis of censored medical cost data. *Health Economics* **15,** 513–525.

Cotter, D., Thamer, M., Narasimhan, K., Zhang, Y., and Bullock, K. (2006). Translating epoetin research into practice: The role of government and the use of scientific evidence. *Health Affairs* **25,** 1249–1259.

De Gruttola, V. and Tu, X. M. (1994). Modeling Progression of CD4+ lymphocyte count and its relationship to survival time. *Biometrics* **50,** 1003–1014.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data visa the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B* **39,** 1–38.

Dobson, A. and Henderson, R. (2003). Diagnostics for joint longitudinal and dropout time modeling. *Biometrics* **59,** 741–751.

Faucett, C. L. and Thomas, D. C. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Statistics in Medicine* **15,** 1663–1685.

Feng, S., Wolfe, R. A., and Port, P. K. (2005). Frailty survival model analysis of the National Deceased Donor Kidney Transplant Dataset using Poisson variance structures. *Journal of the American Statistical Association* **100,** 728–735.

Gatsonis, C., Epstein, A. M., Newhouse, J. P., Normand, S.-L., and McNeil, B. J. (1995). Variations in the utilization of coronary angiography for elderly patients with an acute myocaridal infaction: An analysis using hierarchical logistic regression. *Medical Care* **33,** 625–642.

Golub, G. H. and Welsch, J. H. (1969). Calculation of Gaussian quadrature rules. *Mathematics of Computation* **23,** 221–230.

Gray, R. J. (1994). A Bayesian analysis of institutional effects in a multicenter cancer clinical trial. *Biometrics* **50,** 244–253.

Heitjan, D. F., Kim, C. Y., and Li, H. (2004). Bayesian estimation of cost-effectiveness from censored data. *Statistics in Medicine* **23,** 1297–1309.

Henderson, R., Diggle, P., and Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics* **1**, 465–480.

Hogan, J. and Laird, N. (1997). Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine* **16,** 259–272.

Hsieh, F., Tseng, Y.-K., and Wang, J.-L. (2006). Joint modeling of survival and longitudinal data: Likelihood approach revisited. *Biometrics* **62,** 1037–1043.

Huang, X. and Liu, L. (2007). A joint frailty model for survival time and gap times between recurrent events. *Biometrics* **63,** 389–397.

Huang, X. and Wolfe, R. (2002). A frailty model for informative censoring. *Biometrics* **58,** 510–520.

Kalbfleisch, J. D. and Prentice, R. (2002). *The Statistical Analysis of Failure Time Data*, 2nd edition. New York: Wiley.

Lancaster, A. and Intrator, O. (1998). Panel data with survival: Hospitalization of HIV-positive patients. *Journal of the American Statistical Association* **93,** 46–53.

Lawless, J. F. and Zhan, M. (1998). Analysis of interval-grouped recurrent event data using piecewise constant rate function. *Canadian Journal of Statistics* **26,** 549–565.

Lin, D. Y. and Ying, Z. (2001). Semiparametric and nonparametric regression analysis of longitudinal data (with discussion). *Journal of the American Statistical Association* **96,** 103–126.

Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., and Schabernberger O. (2006). *SAS for Mixed Model*, 2nd edition. Cary, NC: SAS Institute Inc.

Liu, L. and Huang, X. (2007). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine*. In press. DOI: 10.1002/sim.3077.

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

Liu, L., Wolfe, R. A., and Kalbfleisch, J. D. (2007). A random effects model for censored medical costs. *Statistics in Medicine* **26,** 139–155.

Liu, Q. and Pierce, D. A. (1994). A note on Gauss-Hermite quadrature. *Biometrika* **81,** 624–629.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B* **44,** 226–233.

Matsuyama, Y., Sakamoto, J., and Ohashi, Y. (1998). A Bayesian hierarchical survival model for the institutional effects in a multi-center cancer clinical trial. *Statistics in Medicine* **17,** 1893–1908.

Medscape. (2005). Evidence-based medical therapy and device therapy for heart failure: Points of emphasis in the New 2005 AHA/ACC guidelines. Available at: `http://www.medscape.com/viewarticle/520123`.

O'Connell, J. B. (2000). The economic burden of heart failure. *Clinical Cardiology* **23** Suppl III III6–III10.

Olsen, M. K. and Schafer, J. L. (2001). A two-part random effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association* **96,** 730–745.

Pinheiro, J. C. and Bates, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* **4,** 12–35.

Ratcliffe, S. J., Guo, W., and Ten Have, T. R. (2004). Joint modeling of longitudinal and survival data via a common frailty. *Biometrics* **60,** 892–899.

Rondeau, V., Mathoulin-Pelissier, S., Jacqmin-Gadda, H., Brouste, V., and Soubeyran, P. (2007). Joint frailty models for recurring events and death using maximum penalized likelihood estimation: Application on cancer events. *Biostatistics* **8,** 708–721.

Ruppert, D. (2005). Discussion of "Maximization by parts in likelihood inference." *Journal of the American Statistical Association* **100,** 1161–1163.

Stram, D. O. and Lee, J. W. (1994). Variance components testing in the longitudinal mixed effects model. *Biometrics* **50,** 1171–1177. Correction: **51,** 1196.

Sun, J., Park, D.-H., Sun, L., and Zhao, X. (2005). Semiparametric regression analysis of longitudinal data with informative observation times. *Journal of the American Statistical Association* **100,** 882–889.

Therneau, T. M. and Grambsch P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. New York: Springer.

Tooze, J. A., Grunwald, G. K., and Jones, R. H. (2002). Analysis of repeated measures data with clumping at zero. *Statistical Methods in Medical Research* **11,** 341–355.

Tsiatis, A. A. and Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* **88,** 447–458.

Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica* **14,** 809–834.

Vonesh, E. F., Greene, T., and Schluchter, M. D. (2006). Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine* **25,** 143–163.

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

Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* **53,** 330–339.

Xu, J. and Zeger, S. (2001). Joint analysis of longitudinal data comprising repeated measures and times to events. *Applied Statistics* **50,** 375–387.

Ye, Y., Kalbfleisch, J. D., and Schaubel, D. E. (2007). Semiparametric analysis for correlated recurrent and terminal events. *Biometrics* **63,** 78–87.

Received March 2007. *Revised October* 2007.
Accepted October 2007.

Appendix

Gaussian Quadrature

To maximize the likelihood

$$L_i = \int \prod_{j=1}^{n_i} f(y_{ij}|X_{ij}, a_i)\ p(a_i)\, da_i$$

where a_i is the random effect with density $p(\cdot)$, we use a weighted average of the integrand assessed at Q prespecified quadrature points u_q ($q = 1, 2, \ldots, Q$) over the random effects a_i (Liu and Pierce, 1994; Pinheiro and Bates, 1995), i.e.,

$$L_i \approx \sum_{q=1}^{Q} \prod_{j=1}^{n_i} f(y_{ij}|X_{ij}, u_q) p(u_q) w_q$$

with $u_q = \sqrt{2} z_q$ and $w_q = \sqrt{2}\eta_q \exp(z_q^2)$, where η_q and z_q can be obtained from tables (Abramowitz and Stegun, 1972) or algorithms (Golub and Welsch, 1969).

The Gaussian quadrature technique has been implemented in `SAS Proc NLMIXED` (Littell et al., 2006), a user friendly procedure for estimation in nonlinear mixed models. It allows users to construct the log-likelihood function (in the integrand with respect to random effects) by regular `SAS` statements as in the DATA step. For our model, the log-likelihood function composed in `SAS` programming statements can be maximized by option "general" in the "model" statement in `Proc NLMIXED`. We adopt the adaptive Gaussian quadrature option with five quadrature points. The code is available in the Web Supplementary Materials.