Published in final edited form as:

Stat Med. 2008 June 30; 27(14): 2665–2683. doi:10.1002/sim.3077.

The Use of Gaussian Quadrature for Estimation in Frailty Proportional Hazards Models

Lei Liu,

Division of Biostatistics and Epidemiology, Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22908-0717, U.S.A.

Xuelin Huang

Department of Biostatistics, M. D. Anderson Cancer Center, University of Texas, Houston, TX 77030, U.S.A.

Abstract

In this paper we propose a novel Gaussian quadrature estimation method in various frailty proportional hazards models. We approximate the unspecified baseline hazard by a piecewise constant baseline hazard, resulting in a parametric model which can be fitted conveniently by Gaussian quadrature tools in standard software like SAS Proc NLMIXED. We first apply our method to the simple frailty model of correlated survival data (e.g., recurrent or clustered failure times), then to the joint frailty models of correlated survival data in the presence of informative dropout or a dependent terminal event like death. Simulation studies show that our method compares favorably to the well received penalized partial likelihood (PPL) method and the MCEM method, for both normal and Gamma frailty models. We apply our method to three real data examples: (1) the failure times of both eyes in a diabetic retinopathy study, (2) the joint analysis of recurrent opportunistic diseases in the presence of death for HIV infected patients, and (3) the joint modeling of local and distant tumor recurrences and patient survival in a soft tissue sarcoma study. The proposed method greatly simplifies the implementation of the (joint) frailty models and makes them much more accessible to general statistical practitioners.

Keywords

Shared random effects model; Dependent censoring; Frailty models; Random effects; Survival analysis

1 Introduction

It has been increasingly common to model correlated failure time (recurrent or clustered events) data in medical applications. Correlation of the recurrent (or clustered) failure times within the same subjects (or clusters) should be accounted for to avoid bias and improve efficiency. Moreover, over the past decade, there has been a growing research interest to model correlated failure times in the presence of informative dropout or a dependent

terminal event like death. The presence of such a dependent failure time is often termed “informative censoring”, ignorance of which could lead to biased estimates.

Models for correlated failure times were proposed by Vaupel et al. [1], Hougaard [2], Wei, Lin, and Weissfeld [3], Lee, Wei and Amato [4], and references therein. Analyses of recurrent events with informative censoring have been considered by Lancaster and Intrator [5], Wang, Qin and Chiang [6], Huang and Wang [7], Huang and Wang [8], Liu, Wolfe, and Huang [9], Huang and Liu [10]. Huang and Wolfe [11] considered the informative dropout for clustered failure times data. Biomedical examples of correlated failure times alone include the failure times of both eyes [4], the recurrent exacerbation of respiratory symptoms [12], and the recurrent asthma event data [13], etc. For correlated failure times with informative censoring, medical applications include modeling occurrence rate function for recurrent hospitalizations with informative censoring for a cohort of AIDS patients [6]; joint modeling of multiple opportunistic diseases and survival in AIDS patients [7]; and modeling survival time of dialysis patients from different clusters (dialysis centers) in the presence of informative dropout [11].

Frailty models [1] [2] have been a successful tool in the analysis of correlated failure time data. Recurrent/clustered event process can be modeled by a random effects (frailty) proportional hazards model. In the presence of informative censoring, the random effects are then incorporated in the model for the informative censoring process, resulting in a conceptually straightforward joint model ([9], [10], [11]). Such models are useful for assessing the covariate effects on both processes as well as the level of their correlation.

However, the current estimation methods for frailty proportional hazards model are not very satisfactory, which limits its further application. For the “simple frailty models” in the absence of informative censoring, there are two estimation approaches currently used in practice. Nielsen et al. [14], Klein [15] and Guo and Rodriguez [16] adopted the EM algorithm. They treated the unobserved random effects (frailties) as “missing values”, thus the EM algorithm applies naturally. However, EM algorithm converges slowly and extra effort is needed to obtain the standard errors of the estimates [17]. Another approach is the partial penalized likelihood (PPL), which was first proposed by McGilchrist and Aisbett [18] for frailty models. They considered the density of the frailty as a penalty term in the likelihood. The algorithm is fast and readily available in statistical packages such as S-plus and R (function *coxph*) by Therneau and Grambsch [19]. However, no estimate of frailty variance is given by *coxph*. Bootstrap resampling method is often needed to obtain it. These two approaches are closely related: Parner [20] and Therneau and Grambsch ([19], section 9.6) showed that for Gamma frailty, the solution from the penalized partial likelihood method is equivalent to the EM solution for any fixed value of frailty variance.

The estimation process for the “joint frailty models” in the presence of informative censoring is even more complicated. To our knowledge, PPL method is not available in this setting. The EM algorithms are often involved. However, in the E-step, the conditional distribution of random effects given observed data does not have a closed form. Monte Carlo methods (e.g., Metropolis-Hastings algorithm) are often needed to approximate the conditional expectations of the terms involving random effects, making the estimation highly

computationally intensive. An even more serious obstacle for general statistical practitioners lies in the implementation (programming) of the MCEM method, which is quite difficult and has to be treated case by case. Therefore, joint models are not yet widely adopted in practical data analysis.

In this paper we propose to use Gaussian quadrature techniques for estimation in various frailty proportional hazards models, including both the simple and the joint frailty models. To circumvent the nonparametric baseline hazard in Cox's model, we approximate it by a piecewise constant one. The resulting parametric model can be conveniently estimated by Gaussian quadrature tools like Proc NLMIXED in SAS. Since the baseline hazard is often treated as a nuisance parameter, such an approximation will not lessen the value of our method unless it affects our main interest in the estimates of covariate effects. There are several advantages of this approach: (1) our method can handle with ease both normal and non-normal parametric frailties, e.g., Gamma; (2) the implementation (programming) is relatively simple; (3) the estimate is accurate; (4) standard error is readily available, unlike the PPL method which resorts to resampling tools; (5) our method can be extended to more complicated models in a straightforward way.

The rest of the paper is organized as follows. We give the brief introduction to the three frailty models in the next section. In section 3 we introduce the estimation method by Gaussian quadrature and show how to implement it in SAS Proc NLMIXED. Simulation studies are conducted in section 4 to assess the performance of the proposed method. We apply our method to a simple frailty model for the diabetic retinopathy study, and a joint model of recurrent opportunistic diseases and survival for HIV infected patients from the CPCRA study. The proposed estimation method is also used in a bivariate frailty model to describe the relationship of local, distant tumor recurrences and patient survival in a soft tissue sarcoma study. Concluding remarks are given in section 6.

2 Frailty Model

Since the main purpose of this paper is to illustrate the estimation method, instead of the model formulation *per se*, we will give a brief introduction to some basic frailty models for recurrent/clustered events, without or with informative censoring. For a thorough review of various forms of the joint models, please refer to the afore-mentioned literatures.

2.1 Simple Frailty Model

Suppose there are $i = 1, 2, \dots, n$ centers (clusters, families, litters, etc.), each center has $j = 1, 2, \dots, n_i$ subjects. We record the follow up time T_{ij} for each center, which is the minimum of the failure time D_{ij} and censoring time C_{ij} . Here we assume the censoring time is non-informative. Denote by $\delta_{ij} = I(D_{ij} < C_{ij})$ the event indicator, where $I(\cdot)$ is the indicator function. For each subject there are covariate vectors Z_{ij} for fixed effects. Denote by v_i the unobserved random effect (frailty) with a density $f_{\theta}(\cdot)$, which is often assumed to follow a log-Gamma (Hereafter we simply refer to as "Gamma") or normal distribution, although other distributions have been proposed, e.g., positive stable [21]. Note here we put the frailty term in the exponential term for simplicity of notation. The frailty model can be written as

$$h_{ij}(t | v_i) = h_0(t) \exp(\beta^T Z_{ij} + v_i) \quad (1)$$

where $h_0(t)$ is the unspecified baseline hazard, β is the regression coefficient. The observed data $\mathbf{O}_i \equiv \{T_{ij}, \delta_{ij}, Z_{ij}, j = 1, 2, \dots, n_i\}$ are i.i.d for distinct subject i . The likelihood for $\mathbf{O}_i$ is

$$L(\mathbf{O}_i) = \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} [h_{ij}(t_{ij} | v_i)]^{\delta_{ij}} \exp \left[- \int_0^{t_{ij}} h_{ij}(t | v_i) dt \right] f_{\theta}(v_i) dv_i \quad (2)$$

2.2 Joint modeling of recurrent events and survival

Define $T_{i0} = 0$. Let $T_{i1}, T_{i2}, \dots, T_{i, n_i}$ be the observed recurrent event times measured from the study onset for subject i , where n_i is the (random) number of recurrent events before $X_i = \min(C_i, D_i)$, the minimum of censoring time C_i and death time D_i . Write $\delta_{ij} = \mathbb{I}(D_i > C_i)$. Heterogeneity is introduced by observed covariate Z_i and unobserved frailty v_i with a density $f_{\theta}(\cdot)$. Define the hazards for recurrent events by $r_i(t)$ and terminal event by $\lambda_i(t)$ Liu, Wolfe and Huang [9] proposed a shared random effects (frailty) model for recurrent events and survival:

$$r_i(t) = \exp(\beta^T Z_i + v_i) r_0(t) \quad (3)$$

$$\lambda_i(t) = \exp(\alpha^T Z_i + \gamma v_i) \lambda_0(t) \quad (4)$$

where β, α, γ are coefficients, $r_0(t)$ and $\lambda_0(t)$ are baseline hazards for recurrent and terminal event processes, respectively. The correlation between these two processes is introduced by the shared random effect v_i , which can have different impact on $r_i(t)$ and $\lambda_i(t)$ due to γ .

Denote by δ_{ij} the indicator of the recurrent event at time t_{ij} . Let x_i and t_{ij} be the observed follow-up and recurrent event time, respectively. The likelihood for subject i is

$$L_i = \int \prod_{j=1}^{n_i} \left[\exp(\beta^T Z_i + v_i) r_0(t_{ij}) \right]^{\delta_{ij}} \exp \left\{ - \int_0^{x_i} \exp(\beta^T Z_i + v_i) r_0(t) dt \right\} \\ \left[\exp(\alpha^T Z_i + \gamma v_i) \lambda_0(x_i) \right]^{\delta_i} \exp \left\{ - \int_0^{x_i} \exp(\alpha^T Z_i + \gamma v_i) \lambda_0(t) dt \right\} f_{\theta}(v_i) dv_i \quad (5)$$

2.3 Modeling clustered failure times with informative dropout

A similar joint frailty model was proposed by Huang and Wolfe [11] for clustered failure times to account for the correlation between failure and informative censoring. Denote by $T_{ij}, j = 1, 2, \dots, n_i$ the clustered event times, within cluster i . T_{ij} could be censored by informative dropout C_{ij} so we only observe $X_{ij} = \min(C_{ij}, T_{ij})$. Given covariate Z_{ij} and frailty v_i , the hazards for failure $\lambda_{ij}^T(t)$ and informative censoring $\lambda_{ij}^C(t)$ are given by:

$$\lambda_{ij}^T(t) = \exp(\beta^T Z_{ij} + v_i) \lambda_0^T(t) \quad (6)$$

$$\lambda_{ij}^C(t) = \exp(\alpha^T Z_{ij} + \gamma v_i) \lambda_0^C(t) \quad (7)$$

Write $\delta_{ij} = I(D_{ij} = C_{ij})$. The likelihood for subject i is

$$L_i = \int \left[\prod_{j=1}^{n_i} L_{ij}^T L_{ij}^C \right] f_{\theta}(v_i) dv_i \quad (8)$$

where

$$L_{ij}^T = \left[\exp(\beta^T Z_{ij} + v_i) \lambda_0^T(x_{ij}) \right]^{\delta_{ij}} \exp \left[- \int_0^{x_{ij}} \exp(\beta^T Z_{ij} + v_i) \lambda_0^T(t) dt \right]$$

and

$$L_{ij}^C = \left[\exp(\alpha^T Z_{ij} + \gamma v_i) \lambda_0^C(x_{ij}) \right]^{1 - \delta_{ij}} \exp \left[- \int_0^{x_{ij}} \exp(\alpha^T Z_{ij} + \gamma v_i) \lambda_0^C(t) dt \right]$$

3 Estimation methods

3.1 Available approaches

Currently, two estimation approaches are generally adopted in frailty models: the EM algorithm [22] and the penalized partial likelihood (PPL) [18]. The likelihood conditional on the random effects can be maximized straightforwardly, which makes the EM algorithm a natural choice. For “simple frailty model” in section 2.1 with Gamma frailty, the expectation terms involving frailties have a closed form in the E-step [14], making the estimation process fairly straightforward. However, for “simple frailty model” with normal frailty, and the joint frailty models (model (3) and (4) in section 2.2, and model (6) and (7) in section 2.3) with both Gamma and normal frailty, the expectation terms of frailty do not have a closed form. Monte Carlo Markov Chain (MCMC) method, like Metropolis-Hastings algorithm [23], could be applied to generate M random numbers $v_i^{(m)} (m = 1, \dots, M)$ for the estimation of the expectation of such terms in the E-step. Nevertheless, the EM algorithm converges slowly, especially in the case of MCEM for normal frailty model. Up to now, there is no readily available generic EM algorithm for estimation in frailty models. Researchers have to rely on their own programming code. The implementation of such methods (especially MCEM) usually requires high level of programming skills, which greatly hinders its application in practice.

For simple frailty model (1), another available option is the penalized partial likelihood (PPL). McGilchrist and Aisbett [18] considered the density of the frailty as a penalty term in the likelihood. The log penalized likelihood is given as:

$$PPL = \sum_{i=1}^n \log L(\mathbf{O}_i | v_i) - g(v_i, \theta) \quad (9)$$

where $g(v_i, \theta)$ is the penalty term. For Gamma($1/\theta, 1/\theta$) frailty with mean 1 and variance θ ,

$$g(v_i, \theta) = (v_i - \exp v_i)/\theta$$

For normal frailty with corresponding normal distribution $N(0, \theta)$, the penalty term is

$$g(v_i, \theta) = -\frac{1}{2\theta} v_i^2$$

Frailty for each subject was treated as fixed effect and estimated accordingly. The score equation of the PPL with respect to β and v_i is first obtained both as functions of θ . The log profile penalized partial likelihood $PPL(\theta) = PPL(\hat{\beta}(\theta), \hat{v}(\theta), \theta)$ is then maximized over θ . The estimates for β and θ can proceed by iterating between these two steps until convergence. It has been showed ([19], section 9.6) that for Gamma frailty, the solution to the penalized partial likelihood model with the above frailty function is equivalent to the EM solution for any fixed value of θ .

The PPL algorithm is fast and has been incorporated in statistical packages such as S-plus and R (in function *coxph*) by Therneau and Grambsch [19]. However, the estimate $\hat{\theta}$ for frailty variance could be biased. Also, it does not provide standard error for $\hat{\theta}$ [24]. In this sense it is not a “complete” method. Therneau and Grambsch ([19], p. 249) suggested resampling methods, such as bootstrap be used to obtain the standard error estimates.

Other methods have been proposed by Ripatti and Palmgren [25] by penalized partial likelihood with a Laplace approximation, Ha, Lee and Song [26] and Ha and Lee [27] by hierarchical likelihood approach. However, the implementation of these approaches is challenging. Also, these methods were proposed for the simple frailty model and it is hard to extend them to the joint frailty models. Therefore, we will disregard these methods in later presentation.

3.2 Proposed method

At first look, likelihoods (2), (5) and (8) all involve an integral with respect to random effects. Numerical integration techniques, e.g., Gaussian quadrature, thus can be applied for the estimation purpose. However, if the baseline hazards is left unspecified, we will have a non-parametric term in the integrand which can not be estimated directly by Gaussian quadrature techniques.

As shown by Lawless and Zhan [28] and Feng and Wolfe [29], models with a piecewise constant baseline hazard with 8–10 intervals often yield excellent estimates in fixed effects and frailty. The piecewise baseline hazard offers an extra degree of flexibility over the *a priori* choices of baseline hazard distribution (e.g., Weibull), while retaining enough model structure. The resulting parametric likelihood can then be maximized conveniently by

Gaussian quadrature tools in standard statistical packages such as Proc NLMIXED in SAS. An introduction to Gaussian quadrature techniques is given in the appendix. It is noteworthy that Gaussian quadrature can be used for numerical integration when the random effect is non-normal, e.g., Gamma distributed, as illustrated by Nelson et al. [30] by probability integral transformation.

We discuss this approach by taking joint frailty model of recurrent events and death (model (3) and (4) in section 2.2) as an example. The application to the simple frailty model (1) and joint frailty model (6) and (7) is very similar and thus ignored. We start by dividing the follow up for recurrent events into 10 intervals by every 10th quantile (denoted by $Q_1^R, Q_2^R, \dots, Q_{10}^R$ and $Q_0^R = 0$ or the smallest recurrent event time) of the observed recurrent events times. Denote the piecewise constant baseline hazard by $\tilde{r}_0(t)$, we have

$$\tilde{r}_0(t) = r_{0k}, \text{ for } Q_{k-1}^R < t \leq Q_k^R, \text{ where } k = 1, 2, \dots, 10$$

or

$$\tilde{r}_0(t) = \sum_{k=1}^{10} r_{0k} I(Q_{k-1}^R < t \leq Q_k^R)$$

The cumulative baseline hazard for recurrent event is

$$\tilde{R}_0(t) = \sum_{k=1}^{10} r_{0k} \max(0, \min(Q_k^R - Q_{k-1}^R, t - Q_{k-1}^R))$$

Following the same procedure, we can divide the follow up for the terminal events into 10 quantiles. We denote the quantiles for the terminal event times by $Q_1^D, Q_2^D, \dots, Q_{10}^D$ and $Q_0^D = 0$ or the smallest death time. Similarly, we have

$$\tilde{\lambda}_0(t) = \sum_{k=1}^{10} \lambda_{0k} I(Q_{k-1}^D < t \leq Q_k^D)$$

and

$$\tilde{\Lambda}_0(t) = \sum_{k=1}^{10} \lambda_{0k} \max(0, \min(Q_k^D - Q_{k-1}^D, t - Q_{k-1}^D))$$

Likelihood (5) can then be approximated by

$$L_i = \int \prod_{j=1}^{n_i} \left[\exp(\beta^T Z_i + v_i) \tilde{r}_0(t_{ij}) \right]^{\delta_{ij}} \exp \left\{ - \int_0^{x_i} \exp(\beta^T Z_i + v_i) \tilde{r}_0(t) dt \right\} \left[\exp(\alpha^T Z_i + \gamma v_i) \tilde{\lambda}_0(x_i) \right]^{d_i} \exp \left\{ - \int_0^{x_i} \exp(\alpha^T Z_i + \gamma v_i) \tilde{\lambda}_0(t) dt \right\} f_\theta(v_i) dv_i \quad (10)$$

In the case that only time independent covariates are considered, the above likelihood can be simplified as

$$L_i = \int \prod_{j=1}^{n_i} \exp(\tilde{l}_{ij}) f_\theta(v_i) dv_i \quad (11)$$

where

$$\tilde{l}_{ij} = \delta_{ij} \left[\beta^T Z_i + v_i + \log \tilde{r}_0(t_{ij}) \right] - \exp(\beta^T Z_i + v_i) \tilde{R}_0(x_i) + d_i \left[\alpha^T Z_i + \gamma v_i + \log \tilde{\lambda}_0(x_i) \right] - \exp(\alpha^T Z_i + \gamma v_i) \tilde{\Lambda}_0(x_i)$$

The resulting likelihood (11) can then be estimated conveniently by Gaussian Quadrature techniques, as SAS Proc NLMIXED in SAS 9.1 [31]. SAS Proc NLMIXED is a user friendly procedure for estimation in nonlinear mixed models. It has an attractive new feature which allows users to construct the log likelihood function (in the integrand with respect to random effects) by the regular SAS statements, similarly to those in the DATA step. Please see SAS manual [31] for details on this feature. In our case, the log likelihood function form $\tilde{l}_{ij}$ in the integrand of equation (11) can be conveniently specified by SAS programming statements in Proc NLMIXED. Likelihood (11) can then be maximized by option “general” in the “Model” statement in Proc NLMIXED. The code is available from the first author.

4 Simulation

In this section we will perform separate simulation studies for each of the three models in section 2. The results are listed in the corresponding subsection as follows. We assume both normal and Gamma frailty in the simple frailty model in subsection 4.1. Gamma frailty is assumed for the joint model of recurrent events and survival in subsection 4.2. Normal frailty is assumed in subsection 4.3.

4.1 Simple Frailty Model

For both normal and Gamma random effects, we generate data from two settings. In setting I we generate data with an exponential baseline hazard, so the piecewise constant baseline hazard is a proper assumption. In setting II data are generated from a Weibull baseline hazard with shape=2, in which case the piecewise constant baseline hazard is an approximation to the true distribution.

In both settings we assume there are 100 centers with 6 subjects in each center. There are two covariates of interest: (1) a center-level Z_1 which takes value 0 or 1 each with

probability 1/2; and (2) a subject specific Z_2 which is uniformly distributed in (0, 1) for each subject. Fixed effect $\beta = (1, -1)^T$ in both settings. The exponential baseline hazard in setting I is 1. The Weibull baseline hazard in setting II has scale 1 and shape 2, i.e. $h_0(t) = 2t$. Censoring time is taken to be Uniform(0, 2).

We generate the normal frailty from $N(0, \theta)$ with $\theta = 1$, and the Gamma density from Gamma($1/\theta, 1/\theta$) with $\theta = 1$ (mean 1 and variance 1). The results for normal frailty are given in Table 1 and those for Gamma frailty in Table 2.

It can be seen that for both normal and Gamma frailty models, the magnitudes of the empirical biases of the estimates are very small. The coverage probabilities are close to the nominal level 0.95, too. We observe only minor, if not negligible, biases for standard error estimates.

For comparison, we also fit the data by penalized partial likelihood method in R function *coxph*. Since the penalized *coxph* model can not provide the standard error estimates of the frailty variance, we use a bootstrap sample of size 100 to calculate it ([19], p. 249). The results are shown on the right hand side of Tables 1 and 2. We can see the Gaussian quadrature with a piecewise constant baseline hazard perform as well as or often better than the PPL method with an unspecified baseline hazard.

We plot in Figure 1 the estimates of the cumulative baseline hazards by averaging the cumulative baseline hazard estimates from 600 replicates in each case. We can see that the resulting estimates from our method almost overlap the true baseline hazards. In contrast, the cumulative baseline hazards estimated by the PPL method have an upward bias for normal frailty (panels (A) and (B)) no matter what the distribution of the true baseline hazard is. For Gamma frailty, the PPL method yield an unbiased estimate when the underlying baseline hazard is exponentially distributed (panel (C)), but with downward bias when it is Weibull distributed (panel (D)).

4.2 Joint modeling of recurrent events with survival

We generate the data similarly to that in [9]. A single binary covariate Z is considered, which takes value 0 or 1 each with probability 1/2. We set regression coefficients $\alpha = 1$ and $\beta = 1$. Frailties are generated from a Gamma distribution with mean 1 and variance $\theta = 1$. We assume a random censoring time $C = \text{Uniform}(0, 2)$. We choose sample size $n = 200$ and run 600 replicates.

In contrast to [9], we take both $r_0(t)$ and $\lambda_0(t)$ to be Weibull distributed with shape 2. The scales are 2 and 1 for recurrent and terminal events, respectively, resulting in $r_0(t) = 8t$ and $\lambda_0(t) = 2t$. To generate recurrent failures from Weibull distribution with shape 2, we take $T_{j+1} = \sqrt{T_j^2 - \frac{\log U}{\text{scale}^2}}$, where T_j is the j -th recurrent event time ($T_0 = 0$) and $U \sim \text{Uniform}(0, 1)$.

On average, each subject has 2.1 recurrent events. Fifty three percent subjects experience the terminal event, others are censored. 40% of subjects do not have any recurrent event. The results are shown in Table 3.

It can be seen that the estimates have only slight biases. The coverage probabilities are close to the nominal level 0.95. Only small biases are present for standard error estimates.

For comparison, we also fit the data by the EM algorithm with Metropolis-Hastings method in the E-step as in [9]. We assume the baseline hazard is unspecified in this method. The results are shown on the right hand side of Table 3. We can see our method performs as well as the MCEM method with an unspecified baseline hazard. However, the proposed Gaussian quadrature method is much easier to implement. Also, our method runs much faster than the MCEM method: it takes around 4 minutes to converge for the Gaussian quadrature method (by SAS Proc NLMIXED), but 15 minutes for the MCEM method (implemented in R with C code for Metropolis-Hastings algorithm).

We also find that the resulting estimates for cumulative baseline hazards from our method agree very well for most part to those from the MCEM algorithm, as well as the true baseline hazards (results not shown).

4.3 Joint modeling of clustered failure times with informative censoring

We adopt the same setting as in [11], i.e.,

$$\lambda_{ij}^T(t) = 0.2t \exp(0.1\text{Age}_{ij} - 1.4\text{TR}_{ij} + v_i) \lambda_0^T(t)$$

$$\lambda_{ij}^C(t) = 0.04t \exp(0.2\text{Age}_{ij} + 1.2\text{TR}_{ij} + \gamma v_i) \lambda_0^C(t)$$

Covariates include a binary variable TR for treatment which takes value 0 or 1 each with probability 1/2, and a continuous variable for centered age sampled from Uniform(−10, 10). We assume $v \sim N(0, \theta)$ with $\theta = 1$. We set $\gamma = 1$ in setting I and $\gamma = -1$ in setting II. Independent censoring time is set to be Uniform(0, 20) in setting I and Uniform(5, 15) in setting II, to induce roughly 35% informative censoring and 20% administrative censoring in both settings. We choose 40 clusters ($n = 40$) and five subjects ($n_i = 5$) in each cluster and run 500 replicates for each setting.

The results are shown in Table 4. For comparison, we also list in the right hand side the results from MCEM as in Table 1 of [11]. For most parameter estimates, our method is slightly worse than MCEM in terms of biases when the correlation is positive ($\gamma = 1.0$) and slightly better when the correlation is negative ($\gamma = -1.0$). However, the differences between these two methods are very minor. Both yield estimates with small biases. The coverage probabilities are close to the nominal level 0.95. We observe small biases for standard error estimates.

Similar to that in subsection 4.2, we observe that the baseline hazard estimates from both Gaussian quadrature and MCEM algorithm are unbiased (results not shown).

5 Application

5.1 Simple frailty model

We apply our method to the analysis of diabetic retinopathy data, which is available online at <http://www.mayo.edu/hsr/people/therneau/book/data/diabetes.html>. Diabetic retinopathy is the most common and most serious eye complication of diabetes, which may lead to poor vision or even blindness. The Diabetic Retinopathy Study (DRS) was started in 1971 to investigate the efficacy of laser photocoagulation to delay the onset of severe vision loss. This study enrolled 1742 patients. One eye of each patient was randomly chosen for treatment and the other eye serving as an untreated control. Huster, Brookmeyer, and Self [32], Therneau and Grambsch [19] both analyzed a small part of the total data, which is a 50% sample of the high risk patients ($N = 197$). We use this subset too in our analysis.

Since the the failure times of the left eye and right eye from the same patients are correlated, Therneau and Grambsch ([19], section 9.4) fitted the data by frailty models and carried out the estimation by the penalized partial likelihood method. In their study they included two covariates: treatment indicator and a binary variable for juvenile/adult onset of disease. We apply our estimation method for the same frailty models. In particular, we use a piecewise constant baseline hazard with 10 intervals. We then fit the data by Proc NLMIXED assuming both normal and Gamma frailty. The results from our method and that of Therneau and Grambsch ([19], p. 240–1) are shown in Table 5.

We find that the estimates of covariate effects from both methods are close. There is a noticeable difference between these methods in frailty variance estimate (0.961 vs. 0.774) in the normal frailty case. This is consistent with the simulation studies (Table 1) where PPL method has a larger downward bias in $\hat{\theta}$ for normal frailty, especially for the exponential baseline hazard.

We next plot the cumulative baseline hazard estimates from both models in Figure 2. The baseline hazards appear to be exponential in both panels. We find the estimates of cumulative baseline hazard from these methods match well for the gamma frailty. However, there is a moderate difference for the normal frailty, which may be due to the overestimate by the PPL method as shown in simulation.

5.2 Joint modeling of recurrent opportunistic diseases and survival

We applied our method to the Terry Beinr Community Programs for Clinical Research on AIDS (CPCRA) data ([34] [35]). This study compared the effect of didanosine (ddI) and zalcitabine (ddC) on HIV-infected patients who previously received zidovudine and had 300 or fewer CD4 cells per cm^3 . Totally 467 patients were enrolled in this study. Among them 230 were randomized to receive ddI and the others received ddC. The median follow up time is 13 months (range: 1–21 months). There were 100 deaths in the didanosine group and 88 in the zalcitabine group. Patients may experience opportunistic diseases, which were defined in table I in [35] or Table 3 in [34]. Totally there were 172 confirmed or probable opportunistic diseases in the ddI group and 191 in the ddC group. Each patient had between 0 and 5 opportunistic diseases.

We are interested in the joint model of recurrent opportunistic disease and survival. Huang and Wang [7] analyzed the frequency of recurrent opportunistic disease at failure time by semiparametric joint accelerated failure time model. In this paper we analyze the dataset by the shared frailty model as in [9]. We apply the Gaussian quadrature technique to the estimation process, assuming separate piecewise constant baseline hazards for recurrent and terminal event processes.

As in [7], we include treatment indicator ($\text{trt}=1$ for ddC, $\text{trt}=0$ for ddI) as the only covariate.

The resulting model is

$$r_i(t) = \exp(\beta \text{trt} + v_i) r_0(t)$$

$$\lambda_i(t) = \exp(\alpha \text{trt} + \gamma v_i) \lambda_0(t)$$

We assume frailty follows a Gamma distribution with mean 1 and variance θ . We then carry out the estimation by the proposed Gaussian quadrature method with a piecewise constant baseline hazard with 10 intervals. The estimation results are shown in Table 6. We find that there is no significant difference between ddC and ddI in the hazards of recurrent opportunistic diseases or death. ddC reduced the death hazard by 32% ($p = 0.13$), while had almost no effect on the hazard rate of opportunistic diseases ($p = 0.96$). These results are in agreement with those from [7] qualitatively. We also observe that rate of recurrent opportunistic diseases and death are positively associated ($\gamma = 2.526$, $p = 0.0001$), justifying the inclusion of a shared random effect. Higher rate of recurrent opportunistic diseases is associated with higher death rate.

5.3 A bivariate frailty model for a soft tissue sarcoma study

Previous examples consider only one type of recurrent events. In this subsection, we will analyze a data set with two types of recurrent events, namely local and distant recurrences in a soft tissue sarcoma study ([36] [37]). All 674 patients received surgery to remove the tumor. Among them, 218 received radiation alone, 121 received chemotherapy alone, 215 received both, and 120 received none of them. We are interested in the effects of adjuvant radiation and chemotherapy. The medium follow-up time was 6.87 years. One hundred and twelve patients had at least one local recurrence, 300 patients had at least one distant recurrence. The total numbers of local and distant recurrences were 135 and 413 respectively. The maximum numbers of local and distant recurrences were both three. Two hundred and twenty-six patients died of the disease or treatment complications.

Radiation is used to shrink the tumor and prevent local recurrences. Chemotherapy is utilized to kill cancer cells and prevent distant recurrences. While the effect of radiation is well-accepted, many physicians are suspicious about the effects of chemotherapy for soft tissue sarcoma patients. Moreover, local and distant tumor recurrences are mostly correlated. This correlation and the influences of local and distant recurrences on survival are also of great research interest. It is clear that distant recurrences (i.e. metastasis) cause death.

However, it is under debate whether local recurrences increase the risks of metastasis and death.

To answer the above research questions, we propose a joint model of survival, local and distant recurrences. First we assume that unobserved frailty variables v_{1i} and v_{2i} for subject i have the following distribution,

$$\begin{pmatrix} v_{1i} \\ v_{2i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \right) i = 1, \dots, 674, \text{ independently.}$$

For subject i , denote his/her local and distant recurrence intensity function by $r_{1i}(t)$ and $r_{2i}(t)$, and death hazard function by $\lambda_i(t)$, then the joint model is,

$$\begin{cases} r_{1i}(t | Z_i, v_{1i}, v_{2i}) = r_{10}(t) \exp(\beta_1^T Z + v_{1i}) \\ r_{2i}(t | Z_i, v_{1i}, v_{2i}) = r_{20}(t) \exp(\beta_2^T Z + v_{2i}) \\ \lambda_i(t | Z_i, v_{1i}, v_{2i}) = \lambda_0(t) \exp(\alpha^T Z + \gamma_1 v_{1i} + \gamma_2 v_{2i}) \end{cases}$$

Patients with high local (distant) recurrence risks have large v_{1i} (v_{2i}). Three covariates are included in modeling each of these three intensity (hazard) functions: (1) indicator of receiving radiation treatment, (2) indicator of receiving chemotherapy, and (3) the maximum tumor size at baseline (measured in centimeters, ranging from 5cm to 41cm). Their effects on recurrences and death are denoted by vectors β_1 , β_2 and α respectively. The effects of local (distant) recurrences on death are specified by γ_1 (γ_2). The correlation between local and distant recurrences is characterized by the correlation parameter $\rho = \sigma_{21} / \sqrt{\sigma_{11}\sigma_{22}}$.

We carry out the estimation by the Gaussian quadrature technique, assuming piecewise constant baseline hazards for $r_{10}(t)$, $r_{20}(t)$ and $\lambda_0(t)$. The estimation results are shown in Table 7. Radiation is significant in reducing the risks of local, distant recurrences and death. Chemotherapy does not appear to have significant benefit. The frailties for local and distant recurrences are positively correlated ($\rho = 0.84$ with 95% confidence interval (0.74, 0.94)). Local recurrence frailty is not significant associated with death ($p = 0.19$).

6 Discussion

In this paper we propose a unified Gaussian quadrature approach for estimation in various frailty models. We adopt a piecewise constant baseline hazard which gives a close approximation to the unspecified baseline hazard. The resulting parametric model is then conveniently maximized by Gaussian quadrature tools as in SAS Proc NLMIXED. Our method yields satisfactory results as far as parameter estimation is concerned. Compared to the PPL method, our method has a smaller bias and can provide the standard error estimate (for frailty variance) directly. Furthermore, our method can be easily applied to the joint frailty models as in sections 2.2 and 2.3, while such extension for the PPL method is not currently available. Compared to the MCEM approach, an appealing feature of our method is the ease of implementation, which can greatly enhance the applications of joint frailty models in practice. The proposed approach also reduces computational time greatly. Our

estimation method provides a practical solution to the analysis of very complicated scenarios which were highly challenging before, e.g., when there are two correlated recurrent event processes as well as a dependent terminal event, as shown in the soft tissue sarcoma study.

In [28], Lawless and Zhan considered a sub-theme of their paper to advocate the usage of piecewise constant hazard function to “avoid many problems associated with non- and semiparametric methods for incomplete survival and event history data, while giving a high degree of robustness”. We feel that the piecewise constant hazard is specially useful in the frailty model setting due to its great advantage in computation. By transforming the semiparametric frailty model into a parametric one through the piecewise constant baseline hazard approximation, various widely adopted estimation methods can be directly applied to maximize the resulting likelihood, e.g., penalized quasi-likelihood (PQL), Gaussian quadrature, and Laplace approximation. This will greatly expedite the estimation process of random effects models for survival data. For example, Feng and Wolfe [29] used the PQL (SAS Macro GLIMMIX) for estimation. In comparison, we use Gaussian quadrature (SAS Proc NLMIXED), which is more accurate and also easier to implement. Another advantage for such a parametric model is that asymptotic theory is readily available, while the asymptotic theory for the semiparametric frailty model is still under investigation [20].

Two approximations exist in our proposed method. First, we approximate the arbitrary baseline hazard by a piecewise constant counterpart. Second, we approximate the resulting likelihood by Gaussian quadrature techniques. We show that for normal frailty model, our method works well with 10 intervals in the first approximation and 5 points adaptive Gaussian quadrature in the second. However, both approximations can be improved with ease. For example, we can divide the follow up period into finer intervals, say 20 instead of 10, and increase the number of quadrature points to 10 from 5. However, our experience shows that such further improved approximation yields almost identical results as our current choice. On the other hand, with a large number of intervals, we will have too many parameters to estimate. As a result, the stability of computation is in jeopardy.

Good starting points are very important to the success of Gaussian quadrature estimation method. They can generally be obtained by fitting the data by a simpler model. For example, for the simple frailty model (1), we can ignore the random effects and fit the data by the traditional Cox model. In our simulations and applications, however, we tried different starting values and the resulting estimates were very close. So our method is fairly robust to the choice of (reasonable) starting values.

In the Gamma frailty model, similar to [30], we use non-adaptive Gaussian quadrature option in our SAS code. We tried the adaptive Gaussian quadrature option in SAS 9.1 but it failed. However, personal communications with Dr. Kerrie Nelson revealed that other researchers may have successfully implemented the adaptive Gaussian quadrature technique in other software (see <http://otter-rsch.com/admbre/examples/gamma/gamma.html> for details). So the problem may be associated with the specific software.

Probability integral transformation works well for Gaussian quadrature with respect to an independent non-normal random effect. However, it does not apply to the case of correlated

non-normal random effects. We are currently exploring new approaches to handle correlated non-normal random effects in the Gaussian quadrature estimation method.

Our approach can be easily adapted to other settings. For example, a similar bivariate frailty model as in section 5.2 can be used to describe the relationship of outpatient and inpatient services and a dependent terminal event such death. Also, since CD4 value is an important biomarker in AIDS study, we might be interested in the interplay of CD4 repeated measures, recurrent opportunistic diseases, and survival, which is an extension to Guo and Carlin [38] and Vonesh, Greene and Schluchter [39]. Both applied Gaussian quadrature tools to the joint modeling of repeated measures and survival. We are also interested in the estimation of the joint models with crossed or nested random effects. For example, when random effects are crossed, Feng and Wolfe [29] proposed to use penalized quasi-likelihood for estimation with a piecewise constant baseline hazard. With new computational methods emerging for complicated multilevel generalized linear mixed models (e.g., [40]), we expect our method can be easily adapted to these more complex settings in the near future.

Acknowledgements

The authors thank Dr. Kerrie Nelson for helpful discussions.

Appendix: Gaussian Quadrature

Gaussian quadrature is a useful numerical integration tool. Suppose we want to maximize the following 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 normal random effect with density $p(\cdot)$. L_i can be approximated by a weighted average of the integrand assessed at Q predetermined quadrature points u_q ($q = 1, 2, \dots, Q$) over the random effects a_i ([41], [42]), 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 [43] or algorithms [44].

With normal random effect, the Gaussian quadrature technique can be applied directly. However, some adaption is needed when the random effect has other distribution, e.g., Gamma. Nelson et al. [30] presented a solution by probability integral transformation. For $a \sim N(0, 1)$, $u = \Phi(a) \sim \text{Uniform}(0, 1)$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function (CDF). Likewise, the CDF of a non-normal random variable b is $\text{Uniform}(0, 1)$, i.e., $F(b) \sim \text{Uniform}(0, 1)$. Denote by the inverse CDF of b by $F^{-1}(\cdot)$, we can use

$$b = F^{-1}(\Phi(a)) \quad (12)$$

to generate the non-normal random variable of interest.

The implementation of Gaussian quadrature techniques is incorporated in Proc NLMIXED in SAS 9.1 [31]. In particular, a general log likelihood function constructed by SAS programming statements could be conveniently specified and maximized, by option “general” in the “Model” statement in Proc NLMIXED. Please see SAS manual [31] for details on this feature. The code is available from the first author.

References

- [1]. Vaupel JW, Manton KG and Stallard E. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 1979; 16, 439–454. [PubMed: 510638]
- [2]. Hougaard P *Analysis of Multivariate Survival Data*. New York: Springer-Verlag, 2000.
- [3]. Wei LJ, Lin DY and Weissfeld L. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association* 1989; 84, 1065–1073.
- [4]. Lee EW, Wei LJ, and Amato DA. Cox-type regression analysis for large numbers of small groups of correlated failure time observations *Survival Analysis: State of the Art 1992* Eds. Klein JP and Goel PK, 237–247. Dordrecht: Kluwer Academic.
- [5]. Lancaster A and Intrator O. Panel data with survival: hospitalization of HIV-positive patients. *Journal of the American Statistical Association* 1998; 93, 46–53.
- [6]. Wang M, Qin J, and Chiang C. Analyzing recurrent event data with informative censoring. *Journal of the American Statistical Association* 2001; 96, 1057–1065.
- [7]. Huang Y and Wang MC. Frequency of recurrent events at failure time: modeling and inference. *Journal of the American Statistical Association* 2003; 98, 663–670.
- [8]. Huang C-Y and Wang MC. Joint modeling and estimation for recurrent event processes and failure time data. *Journal of the American Statistical Association* 2004; 99, 1153–1165. [PubMed: 24068850]
- [9]. Liu L, Wolfe RA, and Huang X. Shared frailty models for recurrent events and a terminal event. *Biometrics* 2004; 60, 747–756. [PubMed: 15339298]
- [10]. Huang X and Liu L. A joint frailty model for survival time and gap times between recurrent events *Biometrics* 2007; In press.
- [11]. Huang X and Wolfe RA. A frailty model for informative censoring. *Biometrics* 2002; 58, 510–520. [PubMed: 12229985]
- [12]. Therneau TM and Hamilton SA. rhDNase as an example of recurrent event analysis. *Statistics in Medicine* 1997; 16, 2029–2047. [PubMed: 9308130]
- [13]. Duchateau L, Janssen P, Kezic I. and Fortpied C. Evolution of recurrent asthma event rate over time in frailty models. *Applied Statistics* 2003; 52, 355–363.
- [14]. Nielsen G, Gill R, Andersen P and Sorensen T. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics* 1992; 19, 25–43.
- [15]. Klein J Semiparametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* 1992; 48, 795–806. [PubMed: 1420842]
- [16]. Guo G and Rodriguez G. Estimating a multivariate proportional hazards model for clustered data using the EM algorithm, with an application to child survival in Guatemala. *Journal of the American Statistical Association* 1992; 87, 969–976.
- [17]. Louis TA. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B* 1982; 44, 226–233.
- [18]. McGilchrist CA and Aisbett CW. Regression with frailty in survival analysis. *Biometrics* 1991; 47, 461–466. [PubMed: 1912255]

- [19]. Therneau TM and Grambsch PM. Modeling Survival Data: Extending the Cox Model. New York: Springer, 2000.
- [20]. Parner E Asymptotic theory for the correlated Gamma-frailty model. *Annals of Statistics* 1998; 28, 183–214.
- [21]. Hougaard P A class of multivariate failure time distributions. *Biometrika* 1986; 73, 671–678.
- [22]. Dempster AP, Laird NM, and Rubin DB. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B* 1977; 39, 1–38.
- [23]. Chib S and Greenberg E. Understanding the Metropolis-Hastings algorithm. *The American Statistician* 1995; 49, 327–335.
- [24]. Therneau TM, Grambsch PM, and Pankratz VS. Penalized survival models and frailty. *Journal of Computational and Graphical Statistics* 2003; 12, 156–175.
- [25]. Ripatti S and Palmgren J. Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics* 2000; 56, 1016–1022. [PubMed: 11129456]
- [26]. Ha ID, Lee Y and Song J-K. Hierarchical likelihood approach for frailty models. *Biometrika* 2001; 88, 233–243.
- [27]. Ha ID and Lee Y. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics* 2003; 12, 663–681.
- [28]. Lawless JF and Zhan M. Analysis of interval-grouped recurrent event data using piecewise constant rate function. *Canadian Journal of Statistics* 1998; 26, 549–565.
- [29]. Feng S and Wolfe RA. Frailty survival model analysis of the National Deceased Donor Kidney Transplant Dataset using Poisson variance structures. *Journal of the American Statistical Association* 2005; 100, 728–735.
- [30]. Nelson KP, Lipsitz SR, Fitzmaurice GM, Ibrahim J, Parzen M and Strawderman R. Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *Journal of Computational and Graphical Statistics* 2006; 15, 39–57.
- [31]. Littell RC, Milliken GA, Stroup WW, Wolfinger RD, and Schabenberger O. *SAS for Mixed Model*, Second Edition 2006; Cary, NC: SAS Institute Inc.
- [32]. Huster WJ, Brookmeyer R, and Self SG. Modeling Paired Survival Data with Covariate. *Biometrics* 1989; 45, 145–156. [PubMed: 2655727]
- [33]. Stram DO and Lee JW. Variance components testing in the longitudinal mixed effects model. *Biometrics* 1994; 50, 1171–1177. Correction: 51, 1196. [PubMed: 7786999]
- [34]. Abrams DI, Goldman AI, Launer C, Korvick JA, Neaton JD, Crane LR, Grodesky M, Wake-field S, Muth K, Kornegay S, Cohn DL, Harris A, Luskin-Hawk R, Markowitz N, Sampson JH, Thompson M, Deyton L. A comparative trial of Didanosine or Zalcitabine after treatment with Zidovudine in patients with human immunodeficiency virus infection. *New England Journal of Medicine* 1994; 330, 657–662. [PubMed: 7906384]
- [35]. Neaton JD, Wentworth DN, Rhame F, Hogan C, Abrams DI, and Deyton L. Considerations in choice of a clinical endpoint for AIDS clinical trials. *Statistics in Medicine* 1994; 13, 2107–2125. [PubMed: 7846414]
- [36]. Cormier JN, Huang X, Xing Y, Thall PF, Wang X, Benjamin RS, Pollock RE, Antonescu CR, Maki RG, Brennan MF, and Pisters PWT. Cohort analysis of 674 patients with localized high-risk extremity soft tissue sarcoma treated at two cancer centers: chemotherapy-associated outcomes. *Journal of Clinical Oncology* 2004; 22, 4567–4574. [PubMed: 15542808]
- [37]. Huang X, Cormier JN, and Pisters PWT. Estimation of the Causal Effects on Survival of Two-Stage Nonrandomized Treatment Sequences for Recurrent Diseases. *Biometrics* 2006; 62, 901–909. [PubMed: 16984334]
- [38]. Guo X and Carlin BP. Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician* 2004; 58, 16–24.
- [39]. Vonesh EF, Greene T, and Schluchter MD. Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine* 2006; 25, 143–163. [PubMed: 16025541]
- [40]. Pinheiro JC and Chao EC. Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. *Journal of Computational and Graphical Statistics* 2006; 15, 58–81.

- [41]. Pinheiro JC and Bates DM. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 1995; 4, 12–35.
- [42]. Liu Q and Pierce DA. A note on Gauss-Hermite quadrature. *Biometrika* 1994; 81, 624–629.
- [43]. Abramowitz M and Stegun I. *Handbook of Mathematical Functions*, 1972; New York: Dover.
- [44]. Golub GH and Welsch JH. Calculation of Gaussian quadrature rules. *Mathematics of Computation* 1969; 23, 221–230.

Figure 1:

The estimates of cumulative baseline hazards. —: Estimate from our method; - - -: Estimate from the PPL method; ···: True value. (A): Normal frailty with exponential baseline hazard; (B): Normal frailty with Weibull baseline hazard; (C): Gamma frailty with exponential baseline hazard; (D): Gamma frailty with Weibull baseline hazard. The true exponential baseline hazard is $h_0(t) = 1$ and the true Weibull baseline hazard is $h_0(t) = 2t$. Both x and y axes are in log scale in panels (B) and (D).

Figure 2:

The estimates of cumulative baseline hazards. —: Estimate from our method; - - -: Estimate from PPL method. (a): Normal frailty; (b): Gamma frailty.

Table 1:

Simulation Results for the Simple Frailty Model: Normal Frailty

Parameter	Gaussian quadrature				Penalized partial likelihood			
	Bias	SE	SEM	CP	Bias	SE	SEM	CP
<i>Exponential baseline hazard</i>								
$\beta_1 = 1$	-0.008	0.234	0.235	94.5%	-0.050	0.223	0.223	93.8%
$\beta_2 = -1$	0.012	0.206	0.219	96.5%	0.019	0.204	0.217	96.2%
$\theta = 1$	-0.032	0.221	0.217	91.0%	-0.094	0.195	0.189	85.0%
<i>Weibull baseline hazard</i>								
$\beta_1 = 1$	-0.019	0.231	0.233	94.8%	-0.043	0.225	0.225	94.0%
$\beta_2 = -1$	0.013	0.211	0.226	96.3%	0.006	0.212	0.225	96.3%
$\theta = 1$	-0.068	0.217	0.216	90.7%	-0.079	0.203	0.198	88.3%

Bias is the mean of the parameter estimates (based on 600 replicates) minus the true value; SE is the sampling standard error of the parameter estimate; SEM is the sampling mean of the standard error estimate; CP is the coverage probability of the corresponding 95% confidence interval. Gaussian quadrature is carried out assuming piecewise constant baseline hazard.

Penalized partial likelihood is carried out assuming unspecified baseline hazard.

Table 2:

Simulation Results for the Simple Frailty Model: Gamma Frailty

Parameter	Gaussian quadrature				Penalized partial likelihood			
	Bias	SE	SEM	CP	Bias	SE	SEM	CP
<i>Exponential baseline hazard</i>								
$\beta_1 = 1$	-0.005	0.255	0.247	94.2%	-0.009	0.246	0.241	93.5%
$\beta_2 = -1$	0.008	0.240	0.241	95.7%	0.009	0.239	0.240	95.5%
$\theta = 1$	-0.013	0.201	0.199	92.0%	-0.010	0.203	0.210	93.8%
<i>Weibull baseline hazard</i>								
$\beta_1 = 1$	-0.016	0.260	0.248	92.5%	-0.077	0.257	0.245	92.3%
$\beta_2 = -1$	0.008	0.249	0.248	93.7%	0.018	0.246	0.247	94.7%
$\theta = 1$	-0.021	0.207	0.200	93.2%	0.064	0.228	0.223	94.5%

Table 3:

Simulation Results for the Joint Frailty Model of Recurrent Events and Survival

Parameter	Gaussian quadrature				MCEM			
	Bias	SE	SEM	CP	Bias	SE	SEM	CP
$\beta = 1.0$	-0.013	0.208	0.200	93.7%	-0.002	0.206	0.211	95.2%
$\alpha = 1.0$	0.007	0.277	0.278	95.8%	0.023	0.281	0.295	96.7%
$\gamma = 1.0$	-0.002	0.211	0.205	95.2%	0.023	0.215	0.219	96.7%
$\theta = 1.0$	-0.036	0.158	0.157	92.2%	-0.027	0.157	0.161	92.7%

Gaussian quadrature is carried out assuming piecewise constant baseline hazard. MCEM is carried out assuming unspecified baseline hazard.

Table 4:

Simulation Results for the Joint Frailty Model of Clustered Events and Survival

Parameter	Gaussian quadrature				MCEM			
	Bias	SE	SEM	CP	Bias	SE	SEM	CP
$\gamma = 1$								
$10\beta_{\text{age}}^T = 1.0$	-0.029	0.235	0.242	96.6%	0.023	0.251	0.256	95.2%
$\beta_{\text{TR}}^T = -1.4$	-0.030	0.280	0.278	94.4%	-0.023	0.288	0.291	95.4%
$10\beta_{\text{age}}^C = 2.0$	-0.030	0.282	0.294	94.8%	0.012	0.311	0.302	94.2%
$\beta_{\text{TR}}^C = 1.2$	0.027	0.305	0.312	97.2%	0.025	0.322	0.331	96.0%
$\gamma = 1.0$	0.014	0.292	0.295	93.8%	0.027	0.291	0.322	96.2%
$\theta = 1$	-0.048	0.392	0.398	91.0%	-0.023	0.405	0.417	93.0%
$\gamma = -1$								
$10\beta_{\text{age}}^T = 1.0$	-0.002	0.240	0.242	95.2%	-0.004	0.251	0.243	94.0%
$\beta_{\text{TR}}^T = -1.4$	-0.009	0.284	0.283	94.8%	-0.015	0.276	0.289	95.4%
$10\beta_{\text{age}}^C = 2.0$	0.008	0.307	0.300	96.4%	0.052	0.305	0.308	95.6%
$\beta_{\text{TR}}^C = 1.2$	0.004	0.300	0.303	95.2%	-0.001	0.306	0.312	95.2%
$\gamma = -1.0$	-0.021	0.302	0.318	95.8%	-0.034	0.337	0.328	93.0%
$\theta = 1$	0.034	0.456	0.444	92.0%	0.001	0.458	0.497	92.8%

Table 5:

Analysis of Diabetic Retinopathy Data (Huster, et al. 1989, Therneau and Grambsch 2000)

Parameter	Gauss quadrature method			Penalized partial likelihood		
	Est	SE	P value	Est	SE	P value
<i>Normal Frailty</i>						
Treatment	−0.934	0.183	< .0001	−0.899	0.174	< .0001
Adult Onset	0.067	0.225	0.77	0.060	0.209	0.77
Frailty Variance	0.961	0.355	0.003	0.774	0.284	0.003
<i>Gamma Frailty</i>						
Treatment	−0.917	0.180	< .0001	−0.911	0.174	< .0001
Adult Onset	0.039	0.221	0.86	0.041	0.221	0.85
Frailty Variance	0.858	0.315	0.003	0.851	0.380	0.013

The frailty variance in the PPL method is estimated by bootstrap method with a sample size of 100. The P values for frailty variance have been adjusted to address the issue of testing nonzero variance components [33].

Table 6:

Parameter Estimates for Joint Modeling of Recurrent Opportunist Diseases and Survival

Parameter	Est	SE	P value
β	-0.006	0.132	0.96
α	-0.390	0.255	0.13
γ	2.526	0.649	0.0001
θ	0.638	0.127	< .0001

Table 7:

Soft Tissue Sarcoma Study: Parameter Estimates for Bivariate Frailty Model

Parameter	Effect	Est	SE	P-value
On local recurrence				
β_{11}	Tumor size	0.046	0.0179	0.011
β_{12}	Radiation	-0.373	0.216	0.08
β_{13}	Chemotherapy	-0.157	0.207	0.45
On distant recurrence				
β_{21}	Tumor size	0.0857	0.019	< 0.0001
β_{22}	Radiation	-0.479	0.224	0.033
β_{23}	Chemotherapy	-0.095	0.213	0.65
On survival				
α_1	Tumor size	0.194	0.0486	< 0.0001
α_2	Radiation	-1.33	0.497	0.008
α_3	Chemotherapy	-0.242	0.472	0.61
γ_1	Local recurrences	0.295	0.225	0.19
γ_2	Distant recurrences	2.188	0.249	< 0.0001
Variance matrix for bivariate frailty				
σ_{11}	Variance of local recurrences	1.690	0.435	0.0001
σ_{21}	Covariance	2.368	0.3746	< 0.0001
σ_{22}	Variance of distant recurrences	4.750	0.613	< 0.0001