Published in final edited form as:
Biometrics. 2008 September ; 64(3): 762–771. doi:10.1111/j.1541-0420.2007.00952.x.

A Joint Model for Longitudinal Measurements and Survival Data in the Presence of Multiple Failure Types

Robert M. Elashoff[1,2], **Gang Li**[1], and **Ning Li**[2,*]

[1]Department of Biostatistics, School of Public Health, University of California at Los Angeles, Los Angeles, California 90095, U.S.A.

[2]Department of Biomathematics, University of California at Los Angeles, 10833 Leconte Avenue, Box 951766, Los Angeles, California 90095-1766, U.S.A.

Summary

In this article we study a joint model for longitudinal measurements and competing risks survival data. Our joint model provides a flexible approach to handle possible nonignorable missing data in the longitudinal measurements due to dropout. It is also an extension of previous joint models with a single failure type, offering a possible way to model informatively censored events as a competing risk. Our model consists of a linear mixed effects submodel for the longitudinal outcome and a proportional cause-specific hazards frailty submodel (Prentice et al., 1978, *Biometrics* **34**, 541-554) for the competing risks survival data, linked together by some latent random effects. We propose to obtain the maximum likelihood estimates of the parameters by an expectation maximization (EM) algorithm and estimate their standard errors using a profile likelihood method. The developed method works well in our simulation studies and is applied to a clinical trial for the scleroderma lung disease.

Keywords

Cause-specific hazard; Competing risks; EM algorithm; Joint modeling; Longitudinal data; Mixed effects model

1. Introduction

In many biomedical studies it is common that both longitudinal measurements of a response variable and the time to some event of interest are recorded during follow-up. A typical example is the AIDS study where CD4 count and viral load are collected longitudinally and the time to AIDS or death is also monitored. Another example is the Scleroderma Lung Study (Tashkin et al., 2006), a double-blinded, randomized clinical trial to evaluate effectiveness of oral cyclophosphamide (CYC) versus placebo in the treatment of lung disease due to scleroderma. In this study the primary outcome is forced vital capacity (FVC, as % predicted) determined at 3-month intervals from the baseline. The event of interest is the time-to-treatment failure or death. A treatment failure occurs when %FVC of a patient in either group falls by ≥15% after 3 months into the treatment. In both examples the two endpoints are known to be correlated, which may introduce nonignorable nonresponse missing values for the longitudinal outcome after event times (Schluchter, 1992; Hogan and Laird, 1997). This type of missing data cannot

*email: ningli@ucla.edu.

6. Supplementary Materials A Web Appendix referenced in Section 2.2 is available under the Paper Information link at the *Biometrics* website http://www.biometrics.tibs.org.

be handled correctly by standard methods such as mixed effects models (Harville, 1977; Laird and Ware, 1982; Saha and Jones, 2005) and generalized estimating equations (Liang and Zeger, 1986; Zeger, Liang, and Albert, 1988; Robins, Rotnizky, and Zhao, 1995). In the Scleroderma Lung Study, dependence between the two endpoints is further complicated by informatively censored events due to dropout during follow-up. We note that both death and dropout could cause nonignorable missing data in the measurements of %FVC.

Joint analysis of longitudinal measurements and event time data has been proposed to adjust inferences on longitudinal measurements in the presence of nonignorable missing values (Schluchter, 1992; DeGruttola and Tu, 1994; Little, 1995; Hogan and Laird, 1997; Henderson, Diggle, and Dobson, 2000). Joint models can also be used to assess effects of factors of interest on both endpoints simultaneously (Zeng and Cai, 2005b). Yet in other studies joint models have been proposed to solve diffculties in Cox proportional hazards model with time-dependent covariates, which are possibly missing at some event times or subject to substantial measurement error (Faucett and Thomas, 1996; Wulfsohn and Tsiatis, 1997; Wang and Taylor, 2001; Xu and Zeger, 2001; Song, Davidian, and Tsiatis, 2002; Brown and Ibrahim, 2003; Tseng, Hsieh, and Wang, 2005). A common assumption used by all these authors is noninformative censorship in the submodel for survival data with a single failure type, which is no longer applicable in the presence of competing risks or informative censoring. In the scleroderma lung study, dropout cannot always be treated as noninformative censoring because it may be correlated with treatment failure or death and could also cause nonignorable missing values in %FVC.

This article considers joint analysis of repeated measurements and survival data in the presence of multiple failure types. Our method is a natural extension of Henderson et al. (2000). The new aspect in this article is the extension from a single-type failure to the competing risks at the survival endpoint, which enables one to handle informative censoring. Moreover, the profile likelihood approach is used for making inference. In our joint model, a linear mixed effects submodel is used to characterize the distribution of the longitudinal measurements, together with a cause-specific hazards frailty model for competing risks survival data (Prentice et al., 1978). The association between the two aspects is modeled via the linkage of the latent random effects. An expectation maximization (EM) algorithm is derived to estimate the parameters in both submodels, and inverse of the empirical Fisher information from the profile likelihood is used to approximate the variance-covariance matrix of the estimators. We note that standard analysis of competing risks using the causespecific proportional hazards model without random effects can be done by treating one risk as the event of interest, and the others as noninformatively censored events (Kalbfleisch and Prentice, 2002, p. 251-259). However, when cause-specific hazards contain random effects (frailty), such an approach is no longer valid. A new estimation and inference procedure is indeed needed, or the analysis could lead to biased results.

We note that the cause-specific hazards model is not aimed at evaluating the effects of risk factors on the marginal probabilities of occurrence of different risks. If it becomes the research interest, a mixture model approach can be used (Larson and Dinse, 1985; Ng and McLachlan, 2003). We study in another paper joint analysis of longitudinal outcome and competing risks survival data based on a mixture submodel for the competing risks, where additional hidden variables need to be introduced to simplify the EM algorithm.

The rest of this article is organized as follows. The model and the inference procedure are described in Section 2. Section 3 contains application of the joint model to the scleroderma lung study. Section 4 includes simulation studies in which the joint model is compared to separate analyses in the presence of different strengths of associations between the longitudinal

measurements and the event times. Section 5 contains some concluding remarks and possible future directions.

2. The Model, Estimation, and Inference Procedure

2.1 The Model

Let $Y_i(t)$ be the longitudinal outcome measured at time t for subject i, where $i = 1, 2, \ldots, n$, and n is the total number of subjects in study. Each subject may experience one of g distinct failure types or could be right censored during follow-up. Let $C_i = (T_i, D_i)$ denote the competing risks data on subject i, where T_i is the failure/censoring time, and D_i takes value in $\{0, 1, \ldots, g\}$, with $D_i = 0$ indicating a censored event and $D_i = k$ showing that subject i fails from the kth type of failure, where $k = 1, \ldots, g$. Throughout, the censoring mechanism is assumed to be independent of the survival time. As mentioned before, dependent (or informative) censoring can be treated as one of the g types of failures.

The joint model is specified in terms of the following two linked components:

$$\begin{cases} Y_i(t) = X_i^{(1)}(t)^T\beta + \tilde{X}_i^{(1)}(t)^T b_i + \epsilon_i(t), \\ \lambda_k\left(t; X_i^{(2)}(t), u_i, \gamma_k, \nu_k\right) \quad \text{for} \quad k = 1, \cdots, g \\ \quad = \lim_{h\to 0} \frac{p\left\{t \le T_i < t+h, D_i = k \mid T_i \ge t, X_i^{(2)}(t), u_i\right\}}{h} \\ \quad = \lambda_{0k}(t) \exp\left\{X_i^{(2)}(t)^T \gamma_k + \nu_k u_i\right\}. \end{cases} \tag{1}$$

In the linear mixed effects submodel (1), $X_i^{(1)}(t)$ and $\tilde{X}_i^{(1)}(t)$ are vectors of covariates associated with the longitudinal trajectory $Y_i(t)$ and are allowed to change over time. Note that $\tilde{X}_i^{(1)}(t)$ may or may not be the same as $X_i^{(1)}(t)$. The parameter β represents the fixed effects of $X_i^{(1)}(t)$, the vector b_i is a latent random variable that can be interpreted as subject-specific effects of $\tilde{X}_i^{(1)}(t)$, and $\varepsilon_i(t) \sim N(0, \sigma^2)$ for all $t \ge 0$ is the measurement error. We assume that b_i is independent of $\varepsilon_i(t)$ and that $\varepsilon_i(t_1)$ is independent of $\varepsilon_i(t_2)$ for any $t_1 \ne t_2$.

Submodel (2) specifies the distribution of the competing risks survival data with $\lambda_k\left(t; X_i^{(2)}(t), u_i, \gamma_k, \nu_k\right)$ being the instantaneous rate for failures of type k at time t given the vector of covariates $X_i^{(2)}(t)$ and the frailty u_i in the presence of all other failure types. Here $\lambda_{0k}(t)$ is a completely unspecified baseline hazard function for risk k, where $k = 1, \ldots, g$. We assume that b_i and u_i jointly have a multivariate normal distribution:

$$\theta_i = \begin{pmatrix} b_i \\ u_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{bb} & \Sigma_{bu}^T \\ \Sigma_{bu} & \sigma_u^2 \end{pmatrix}\right).$$

We denote $\gamma = \left(\gamma_1^T, \ldots, \gamma_g^T\right)^T$ and $\nu = (\nu, \ldots, \nu_g)^T$. The parameter ν_1 is set to 1 to ensure identifiability. There are three sources of correlations in our joint model. The correlation among the longitudinal measurements within the same subject is modeled by random effects b_i, which is similar to standard mixed effects models. Submodel (2) is an extension of the cause-specific hazards model for competing risks survival data described by Prentice et al. (1978) with subject-specific random effects u_i to model the correlation between different failure types. The linkage between the longitudinal measurements and the competing risks is flexibly modeled

through the association between b_i and u_i via a joint multivariate normal distribution. Therefore, the association between the two endpoints can be tested by the hypothesis that all the elements in Σ_{bu} equal zero. Finally, we assume the longitudinal measurements are independent of the competing risks survival data, conditional on all the covariates and random effects.

2.2 Likelihood and the EM Algorithm

Suppose the longitudinal outcome $Y_i(t)$ is observed at time points t_{ij} for $j = 1,\ldots, n_i$, and denote $Y_i = (Y_{i1},\ldots,Y_{ini})$. Note that the set $(t_{i1},\ldots,t_{ini})$ can be different among subjects, due to different event times and the fact that some patients may miss one or more visits. We assume that the missing values in the longitudinal measurements caused by reasons other than occurrence of the events are missing at random. Recall that the competing risks data on subject i are $C_i = (T_i, D_i)$. It is important to note that the joint distribution of (Y, C) is completely determined by $f(Y \mid \theta, \Psi)$, $f(C \mid \theta, \Psi)$, and $f(\theta \mid \Psi)$ as specified in Section 2.1, where $\Psi = (\beta, \sigma^2, \gamma, \nu, \Sigma, \lambda_{01}(t), \ldots, \lambda_{0g}(t))$ is the vector containing all the unknown parameters from (1) to (2). The full likelihood function for Ψ, conditional on the observed data (Y_i, C_i) for $i = 1,\ldots, n$ and the covariates, is thus

$$
\begin{aligned}
L\ (\Psi; Y, C) &\propto \Pi_{i=1}^{n} f(Y_i, C_i|\Psi) \\
&= \Pi_{i=1}^{n} \int_{\theta} f(Y_i|C_i, \theta\Psi)\, f(C_i|\theta, \Psi)\, f(\theta|\Psi)\, d\theta \\
&= \Pi_{i=1}^{n} \int_{\theta} f(Y_i|\theta, \Psi)\, f(C_i|\theta, \Psi)\, f(\theta|\Psi)\, d\theta \\
&= \Pi_{i=1}^{n} \int_{\theta} \left[\Pi_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} \left(Y_{ij} - X_i^{(1)}(t_{ij})^T \beta - \tilde{X}_i^{(1)}(t_{ij})^T b \right)^2 \right\} \right] \left\{ \Pi_{k=1}^{g} \lambda_k \left(T_i; X_i^{(2)}(T_i), u, \gamma_k, \nu_k \right)^{I(D_i=k)} \right\} \\
&\quad \times \exp\left[-\int_0^{T_i} \left\{ \sum_{k=1}^{g} \lambda_k \left(t; X_i^{(2)}(t), u, \gamma_k, \nu_k \right) \right\} dt \right] \\
&\quad \times \frac{1}{\sqrt{2\pi|\Sigma|}} \exp\left(-\frac{1}{2} \theta^T \Sigma^{-1} \theta \right) d\theta,
\end{aligned}
\tag{3}
$$

where the second equality follows from the assumption that Y and C are independent conditional on all the covariates and the random effects. The cumulative hazards of the baseline functions in λ_k are chosen to be step functions with jumps at observed event times.

The observed data likelihood is diffcult to maximize in the presence of integration. Below we propose to obtain the maximum likelihood estimate of Ψ through an EM algorithm. Given θ_i, the complete data likelihood is

$$
\begin{aligned}
L\ (\Psi; Y, C, \theta) &\propto \Pi_{i=1}^{n} \left[\Pi_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} \left(Y_{ij} - X_i^{(1)}(t_{ij})^T \beta - \tilde{X}_i^{(1)}(t_{ij})^T b_i \right)^2 \right\} \right] \left\{ \Pi_{k=1}^{g} \lambda_k \left(T_i; X_i^{(2)}(T_i), u_i, \gamma_k, \nu_k \right)^{I(D_i=k)} \right\} \\
&\quad \times \exp\left[-\int_0^{T_i} \left\{ \sum_{k=1}^{g} \lambda_k \left(t; X_i^{(2)}, (t), u_i, \gamma_k, \nu_k \right) \right\} dt \right] \\
&\quad \times \frac{1}{\sqrt{2\pi|\Sigma|}} \exp\left(-\frac{1}{2} \theta_i^T \Sigma^{-1} \theta_i \right).
\end{aligned}
\tag{4}
$$

In the E-step of the $(m + 1)$th iteration, we need to calculate the expected values of all the functions of θ_i, say $h(\theta_i)$, that appear in $l(\Psi;\theta) = \log L(\Psi; Y, C, \theta)$ conditional on (Y, C) and $\Psi^{(m)}$. The expectation can be derived by

$$
E_{\theta_i|Y_i,C_i,\Psi^{(m)}} (h(\theta_i)) = \int h(\theta_i)\, f\left(\theta_i|Y_i, C_i, \Psi^{(m)} \right) d\theta_i.
\tag{5}
$$

In the M-step, we update Ψ using

$$\Psi^{(m+1)}=\text{argmax}_{\Psi} Q\left(\Psi;\Psi^{(m)}\right), \tag{6}$$

where $Q(\Psi;\Psi^{(m)})= E_{\theta|Y,C,\Psi}(m)\ (l(\Psi;\theta))$. Recall that $\Psi = (\beta, \sigma^2, \gamma, \nu, \Sigma, \lambda_{01}(t),\ldots, \lambda_{0g}(t))$. See Web Appendix A for the derivation of the EM algorithm. The integrals can be evaluated using Gauss-Hermite quadrature (Press et al., 1992). We found that 20 quadrature points approximated the integrals satisfactorily, and hence were used for all expectations required in the calculation. We show that β, σ^2, Σ, and the cumulative baseline hazard functions $H_{0k}(t)$ can be updated with closed forms, where $H_{0k}(t)$ is a step function with jumps at observed event times due to risk k, $k = 1,\ldots,g$. No closedform solutions exist for γ and ν, which need to be updated using a one-step Newton-Raphson algorithm in each iteration. The algorithm stops when the convergence criteria are satisfied. We point out that the model construction and the EM algorithm derivations are natural extensions of the work originally proposed by Wulfsohn and Tsiatis (1997) and then extended by Henderson et al. (2000).

2.3 Standard Error Estimation

The parameter vector Ψ can be split into two components, the parametric component $\Omega = (\beta, \sigma^2, \gamma, \nu, \Sigma)$ and the collection of nonparametric baseline hazard functions $\Lambda = (\lambda_{01}(t), \ldots,\lambda_{0g}(t))$. The dimension of Λ is $O(n)$ which makes the method by Louis (1982) infeasible. On the other hand, because we are only interested in making inference on Ω, calculating the entire information matrix with the baseline functions is unnecessary. We propose to approximate the variance- covariance matrix of Ω by inverting the empirical Fisher information obtained from the profile likelihood where the baseline hazards functions are profiled out (Lin, McCulloch, and Rosenheck, 2004). Let $l^{(i)}\left(\widehat{\Omega};Y,C\right)$ denote the observed score vector from the profile likelihood on the ith subject evaluated at $\widehat{\Omega}$. The observed information matrix of Ω can be approximated by

$$\sum_{i=1}^{n} l^{(i)}\left(\widehat{\Omega};Y,C\right) l^{(i)}\left(\widehat{\Omega};Y,C\right)^{T}. \tag{7}$$

3. Example

In the scleroderma lung study, the primary outcome is FVC (% predicted), which was measured every 3 months from the baseline. We are interested in evaluating whether oral CYC can either improve %FVC scores or decrease the risk of treatment failure or death. The study enrolled 158 patients with scleroderma-related interstitial lung disease, who were randomized to receive either CYC (2 mg/kg; 79 patients) or identical-appearing placebo (79 patients) for 12 months. A second year of follow-up was performed to determine whether CYC effects persisted after stopping treatment. The patients may drop out or die before the completion of the study and the average number of visits per patient is 7.3. For illustration purposes, we considered two factors in our joint model when assessing the CYC treatment effects: baseline %FVC (FVC_0) and baseline lung fibrosis (FIB_0). The latter was included because it was an important risk factor and it had an interaction effect with the treatment as suggested by our preliminary analyses. Comparison of baseline- and fibrosisadjusted %FVC scores between the two groups revealed overlapping values prior to 6 months due to the dose escalation protocol for CYC and the anticipated delay in treatment effect. In addition, the beneficial effects of CYC on pulmonary function continued to increase after stopping treatment, but eventually dissipated at month 24. Therefore we analyzed 6-21 months' %FVC scores at the longitudinal endpoint.

We excluded one patient who had a suspicious %FVC measurement at month 9, and eventually there were 140 patients in the final analysis. Table 1 summarizes the number of treatment failures, deaths, and censored events due to withdrawal in these patients. We classified the censored events into two types: (1) informative, if the event was known to be disease related or treatment related in which the patient withdrew due to worsening disease, adverse event (AE), or serious adverse event (SAE). However, there may be %FVC measurements available after treatment withdrawal because these patients were encouraged to continue with scheduled visits and procedures. Therefore, this category can be further divided into two sub-types depending on whether there were measurements after the events. (2) Noninformative, if there was no evidence showing that the event was related to the disease or the treatment. We observed 14 treatment failures or deaths, 32 informatively censored events, and 5 noninformatively censored events. The remaining 89 patients (45 in CYC group and 44 in the placebo group) completed the study. Because the informatively censored events were related to the patients' disease condition, this category not only correlates with treatment failure or death, but could also cause nonignorable missing data in %FVC scores.

Figure 1 (a) and (b) displays the longitudinal profile of %FVC over time for the two groups. There exists a large variation in the baseline %FVC and the measurements at the subsequent visits are apparently correlated with the baseline values. The three types of events, treatment failure or death, informatively censored events, and noninformatively censored events, are labeled by different symbols. It seems that occurrence of some events of the first two types is related to low %FVC scores. However, whether it would cause a problem for the analysis of %FVC measurements remains unclear before we apply the joint model.

We considered a random slope model that provided a better model fit than a random intercept model, and the improvement of a random intercept and random slope model from this simpler one was marginal. With the baseline covariates adjusted totheir means, we have, for subject *i* at visit *j*,

$$\begin{aligned} \%FVC_{ij} = \; & \beta_0+\beta_1 t_{ij}+\beta_2 FVC_{0i}+\beta_3 FIB_{0i}+\beta_4 CYC_i \\ & +\beta_5 FVC_{0i}\times CYC_i+\beta_6 FIB_{0i}\times CYC_i \\ & +\beta_7 t_{ij}\times CYC_i+b_i t_{ij}+\epsilon_{ij}, \end{aligned} \tag{8}$$

where b_i is the random slope and $\in_{ij} \sim N(0, \sigma^2)$ is the mutually independent measurement error; the cause-specific hazards for treatment failure or death (risk 1) and informatively censored events (risk 2) are specified as

$$\lambda_1(t) = \lambda_{01}(t)\exp(\gamma_{11}FVC_{0i}+\gamma_{12}FIB_{0i}+\gamma_{13}CYC_i+\gamma_{14}FVC_{0i}\times CYC_i+\gamma_{15}FIB_{0i}\times CYC_i+u_i), \tag{9}$$

$$\lambda_2(t) = \lambda_{02}(t)\exp(\gamma_{21}FVC_{0i}+\gamma_{22}FIB_{0i}+\gamma_{23}CYC_i+\gamma_{24}FVC_{0i}\times CYC_i+\gamma_{25}FIB_{0i}\times CYC_i+\nu_2 u_i), \tag{10}$$

respectively. The joint distribution of the random slope b_i in (8) and the frailty u_i in (9) and (10) are assumed to be a zero-mean bivariate normal with variance-covariance

$$\Sigma = \begin{pmatrix} \sigma_b^2 & \sigma_{bu} \\ \sigma_{bu} & \sigma_u^2 \end{pmatrix}.$$

The results of such a joint analysis (model (8)-(10)) are summarized in Table 2. For comparison purposes, we also carried out separate analyses of the two endpoints, which were done by fitting

a random slope linear model (8) for %FVC scores using SAS Proc Mixed and a cause-specific hazards frailty model (9)-(10) for the competing risks failure time data. We used the same approach as proposed in Section 2.3 to compute the variance-covariance matrix for the parameters at the survival endpoint. It is observed that the two methods produce similar point estimates and standard errors for most parameters and identify the same set of significant effects at the longitudinal endpoint: FVC_0, FIB_0, and the interaction between FIB_0 and CYC. Lung fibrosis is negatively correlated with %FVC, but its interaction with CYC suggests that CYC increases %FVC more effectively for patients with a higher degree of fibrosis. The overall effects of treatment CYC on %FVC scores could be evaluated by testing the null hypothesis $\beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$, which yields a *p*-value 0.0271 for the joint model, and 0.0429 for the separate model. No significant overall effects of CYC are identified for the time-to-treatment failure or death by testing $\gamma_{13} = \gamma_{14} = \gamma_{15} = 0$. In the joint model, with negative covariance σ_{bu} and positive ν_2, there tends to be a lower risk for both treatment failure or death and informatively censored events due to dropout for patients with higher than average increasing rate of %FVC over time. However, there is not enough evidence to show that either σ_{bu} or ν_2 is significantly different from zero, which explains why we observe similar estimates in the joint analysis and the separate analysis. We note that, as seen in our first simulation study in the next section, the estimates for ν_2 are not reliable under the current sample size and event rates. Hence we would not overinterpret the quantities in this application. In addition, the simulation suggests that the bias of ν_2 does not seem to affect the estimation of other parameters in the joint model. Finally, we would like to point out that it is necessary to perform the joint analysis to avoid possible invalid inference from the separate analyses, although the two approaches produced similar results for the scleroderma lung study.

4. Simulation Studies

We note that the association between the longitudinal measurements and the survival data is modeled by the following parameters: the parameter ν_2, the variances of random effects b_i and u_i, and their correlation. In the simulation experiments we show how to adjust the association between the two end points by manipulating one or all of the above parameters. We conducted two simulations. In the first simulation the association between the two endpoints is negligible, so we expect similar results between the joint model and the separate analyses. In the second study there is a strong association between the two endpoints, so the competing risks could cause considerable missing values in the longitudinal measurements. Consequently, the separate analyses are expected to produce biased estimates.

We conducted the first simulation by generating data with structures similar to the scleroderma lung study, in which the variance σ^2_u was set to 0.05, so the linkage between the two endpoints was expected tobe negligible. The longitudinal measurements and the competing risks event times were simulated from model (8)-(10), where the covariates were sampled from distributions close to what we observed in the real data. We approximated the nonparametric baseline hazards by Weibull distributions, and the other parameters were set to their estimated values from the joint model. The results of the joint model and the separate analyses are compared in Table 3 using 200 simulated data sets with sample size $n = 140$. In the table we label the mean of point estimates as Est, the empirical standard error as SE, the median of estimated standard error as Est. SE, and the confidence interval coverage probability as CP.

It is shown that the two methods produce comparable point estimates and empirical standard errors for most parameters. However, the random effects ν_2 and σ^2_u and their standard errors are poorly estimated by the separate analyses. Even the joint model is not able to provide sound estimates for ν_2. It suggests that with $n = 140$, low event rates (around 10% for risk 1 and 23% for risk 2), and a relatively small variance of u_i $\left(\sigma^2_u=0.05\right)$, the frailty at the survival endpoint

could be hard to estimate. Also because of the small variance of u_i, the association between the longitudinal measurements and the survival data is not sufficiently strong, so we observed similar results from the two methods. At last, we note that under these settings the joint model could be conservative in the sense that the estimated standard errors tend to be larger than the empirical ones.

In the second simulation, we adjusted all the four parameters (ν_2, the variances of random effects b_i and u_i, and their correlation) to increase association between the two endpoints and to show that the separate analyses could lead to invalid inference. The longitudinal measurements were simulated from the following random slope model:

$$Y_{ij}=\beta_0+\beta_1 t_{ij}+\beta_2 X_{2i}+\beta_3 X_{ij}\times X_{2i}+b_i t_{ij}+\epsilon_{ij}, \tag{11}$$

where t_{ij} = 0, 0.5, 1, …, 5 was the scheduled visiting time, X_{2i} ~ Bernoulli(0.5) acting as a treatment group indicator in randomized trials, and the measurement error $\in_{ij}$ ~N(0, 5). The distribution for the random slope b_i is specified below. Again, we simulated two competing risks for event times, say risk 1 and risk 2, with the following cause-specific hazards:

$$\lambda_1\left(t;X_{1j},X_{2i},u_i,\gamma_1\right)=\lambda_{01}(t)\exp\left(\gamma_{11}X_{1i}+\gamma_{12}X_{2i}+u_i\right), \tag{12}$$

$$\lambda_2(t;X_{1i},X_{2i},u_i,\gamma_2,\nu_2)=\lambda_{02}(t)\exp\left(\gamma_{21}X_{1i}+\gamma_{22}X_{2i}+\nu_2 u_i\right), \tag{13}$$

where the covariates X_{1i} ~ N(2, 1.0) and X_{2i} was from (11). The random effects b_i and u_i have a zero-mean bivariate normal distribution with variance-covariance matrix

$$\Sigma=\begin{pmatrix}\sigma_b^2 & \sigma_{bu}\\ \sigma_{bu} & \sigma_u^2\end{pmatrix},$$

where σ_b^2=0.5, σ_u^2=0.5, and σ_{bu} = 0.45, to reach a correlation of -0.9. The baseline hazards were held constant at 0.1 and 0.15 for risk 1 and risk 2, respectively, so the time to each risk formed an exponential distribution. We simulated the censoring time from an exponential distribution with mean 25 and set the maximum follow-up time as 5. The longitudinal measurements were missing after the observed or censored event times. Overall, we obtained censoring rate at about 20%, risk 1 rate at about 45%, and risk 2 rate at about 35%. The Monte Carlo samples were analyzed in two ways: the joint model as specified in (11)-(13), and the separate analyses of the two endpoints that were done by fitting a random slope linear model (11) for the longitudinal outcome and a cause-specific hazards frailty model (12)-(13) for the competing risks failure time data. The simulation was based on 200 Monte Carlosamples, and two sample sizes, 200 and 500, were considered (Tables 4 and 5). We label the mean square error as MSE_J and MSE_S for the joint model and the separate analyses, respectively.

The results can be summarized as follows. First of all, the joint model is able to get almost unbiased estimates for all the parameters and the estimated standard errors from profile likelihood are close to the empirical ones. The simulated coverage probabilities of 95% confidence intervals constructed with these estimates are all around the nominal value 0.95. However, the time trend β_1 (the time trend for the control group) and β_3 (change of the time trend comparing the treatment group with the placebo group) of the longitudinal measurements are severely biased in the separate analysis, even for a large sample size (say, 500). The variance

of the time trend, σ_b^2, is also underestimated. These result in much lower confidence interval coverage probabilities for the parameters. The biases are the consequence of the informative dropout process, in which with a negative correlation between b_i and u_i and the positive coefficient ν_2, we observe a higher risk of dropout for those subjects with lower than average increasing rates over time. The nonignorable missing values after dropout cannot be accounted for in the linear mixed effects model alone and biases in the estimated time trend and its variance are observed. This also results in attenuated slope change comparing the treatment group with the placebo group (β_3). The biases will not vanish by increasing the sample size as shown in Table 4. In the joint analysis the informative dropout process has been modeled together with the longitudinal measurements so that we are able to obtain almost unbiased estimates of these quantities. In addition, both the empirical and the estimated standard errors of β_1, β_3, and σ_b^2 tend to be smaller in the separate model than in the joint model, which could also be due to violation of the missing-at-random assumption in the separate analysis. Our results are consistent with the findings of Henderson et al. (2000). Second, it is observed that the joint model provides a more accurate estimate for ν_2 than the cause-specific hazards model alone for both sample sizes. This indicates that we can improve efficiency of frailty estimation in the survival endpoint by combining the information of the longitudinal outcome, if the two endpoints are correlated and the correlation is correctly modeled. Third, the joint model, which utilizes information from both endpoints, tends to produce smaller empirical standard errors of the parameters for the competing risks endpoint than the separate model. Over- all, the joint model performs better than the separate analyses because the *MSE*s from the joint model are almost always smaller. At last, we point out that estimation of the standard errors at the survival endpoint seems to require a larger sample size than the longitudinal model. As a consequence, the empirical standard errors and the estimated standard errors are fairly similar for parameters in the longitudinal submodel, but more varied for those in the competing risks.

5. Discussion

In this article, we have proposed a joint model for longitudinal measurements and competing risks failure time data, which can be used to handle informative censoring at the survival endpoint by treating the informatively censored events as a competing risk for the event of interest. This is an extension of previous joint models with a single failure type for the event times. The assumption of multivariate normal provides a flexible approach to link together the random effects from the two aspects. We have developed an EM algorithm to obtain the MLEs of the parameters and proposed a profile likelihood method to estimate their standard errors. Our joint model not only offers a framework for joint inference on longitudinal outcome and time-to-event data with competing risks, but also a means to analyze longitudinal outcome with nonignorable missing mechanisms. However, because the dimension of the parameter space increases with the sample size due to nonparametric baseline hazards, a rigorous treatment of the asymptotic properties of the MLEs under our model warrants future research. We note that recently Zeng and Cai (2005a) derived the asymptotic distributions of the maximum likelihood estimators from their joint model in which they considered a single failure type for the time to event.

We assume that correlations among different competing risks are driven by the same random effect u_i. This assumption can be further extended so that the risks are linked by different random effects that are potentially correlated. We note that this model would require considerably more computation efforts, especially when g is moderate, and a larger sample size so that there are enough events for each risk for sound inference.

We know that in the scleroderma study influential points and outliers plausibly occur, so robustizing the joint model to handle outlying observations in the longitudinal outcome can be

investigated in the future. The robustness obtained using the *t* distribution in linear models have been studied by Sutradhar and Ali (1986), Lange, Little, and Taylor (1989), and Taylor, Yu, and Sandler (2005). In our joint model we may replace the normal distribution assumption for measurement errors with a *t* distribution to take into account longer-thannormal tails.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Dr Donald P. Tashkin in UCLA David Geffen School of Medicine for providing the Scleroderma Lung Study data. We also thank the associate editor and the referees for their valuable comments.

References

Brown ER, Ibrahim JG. A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. Biometrics 2003;59:221–228. [PubMed: 12926706]

DeGruttola V, Tu XM. Modeling progression of CD-4 lymphocyte count and its relationship to survival time. Biometrics 1994;50:1003–1014. [PubMed: 7786983]

Faucett CL, Thomas DC. Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. Statistics in Medicine 1996;15:1663–1685. [PubMed: 8858789]

Harville DA. Maximum likelihood approaches to variance component estimation and to related problems. Journal of the American Statistical Association 1977;72:320–338.

Henderson R, Diggle P, Dobson A. Joint modeling of longitudinal measurements and event time data. Biostatistics 2000;4:465–480. [PubMed: 12933568]

Hogan JW, Laird NM. Model-based approaches to analysing incomplete longitudinal and failure time data. Statistics in Medicine 1997;16:259–272. [PubMed: 9004396]

Kalbfleisch, JD.; Prentice, RL. The Statistical Analysis of Failure Time Data. Vol. 2nd edition. John Wiley and Sons; Hoboken, NJ: 2002.

Laird NM, Ware JH. Random-effects models for longitudinal data. Biometrics 1982;38:963–974. [PubMed: 7168798]

Lange KL, Little RJA, Taylor JMG. Robust statistical modeling using the *t* distribution. Journal ofthe American Statistical Association 1989;84:881–896.

Larson MG, Dinse GE. A mixture model for the regression analysis of competing risks data. Applied Statistics 1985;34:201–211.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13–22.

Lin H, McCulloch CE, Rosenheck RA. Latent pattern mixture models for informative intermittent missing data in longitudinal studies. Biometrics 2004;60:295–305. [PubMed: 15180654]

Little RJA. Modeling the drop-out mechanism in repeated-measures studies. Journal of the American Statistical Association 1995;90:1112–1121.

Louis TA. Finding the observed information matrix when using the EM algorithm. Journal of the Royal Statistical Society 1982;44:226–233.Series B

Ng SK, McLachlan GJ. An EM-based semiparametric mixture model approach to the regression analysis of competing-risks data. Statistics in Medicine 2003;22:1097–1111. [PubMed: 12652556]

Prentice RL, Kalbfleisch JD, Peterson AV, Flournoy N, Farewell VT, Breslow NE. The analysis of failure times in the presence of competing risks. Biometrics 1978;34:541–554. [PubMed: 373811]

Press, WH.; Teutolsky, SA.; Vetterling, WT.; Flannery, BP. Numerical Recipes in FORTRAN. The Art ofScientific Computing. Vol. 2nd edition. Cambridge University Press; New York: 1992.

Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated out-comes in the presence of missing data. Journal of the American Statistical Association 1995;90:106–121.

Saha C, Jones MP. Asymptotic bias in the linear mixed effects model under nonignorable missing data mechanisms. Journal ofthe Royal Statistical Society 2005;67:167–182.Series B

Schluchter MD. Methods for the analysis of informatively censored longitudinal data. Statistics in Medicine 1992;11:1861–1870. [PubMed: 1480878]

Song X, Davidian M, Tsiatis AA. A semi-parametric likelihood approach to joint modeling of longitudinal and time-to-event data. Biometrics 2002;58:742–753. [PubMed: 12495128]

Sutradhar BC, Ali MM. Estimation of parameters of a regression model with a multivariate *t* error variable. Communications in Statistics—Theory and Methods 1986;15:429–450.

Tashkin DP, Elashoff RM, Clements PJ, et al. Cyclophosphamide versus placebo in scleroderma lung disease. The New England Journal ofMedicine 2006;354:2655–2666.

Taylor JMG, Yu M, Sandler HM. Individualized predictions of disease progression following radiation therapy for prostate cancer. Journal of Clinical Oncology 2005;23:816–825. [PubMed: 15681526]

Tseng YK, Hsieh F, Wang JL. Joint modelling of accelerated failure time and longitudinal data. Biometrika 2005;92:587–603.

Wang Y, Taylor JMG. Joint modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. Journal of the American Statistical Association 2001;96:895–905.

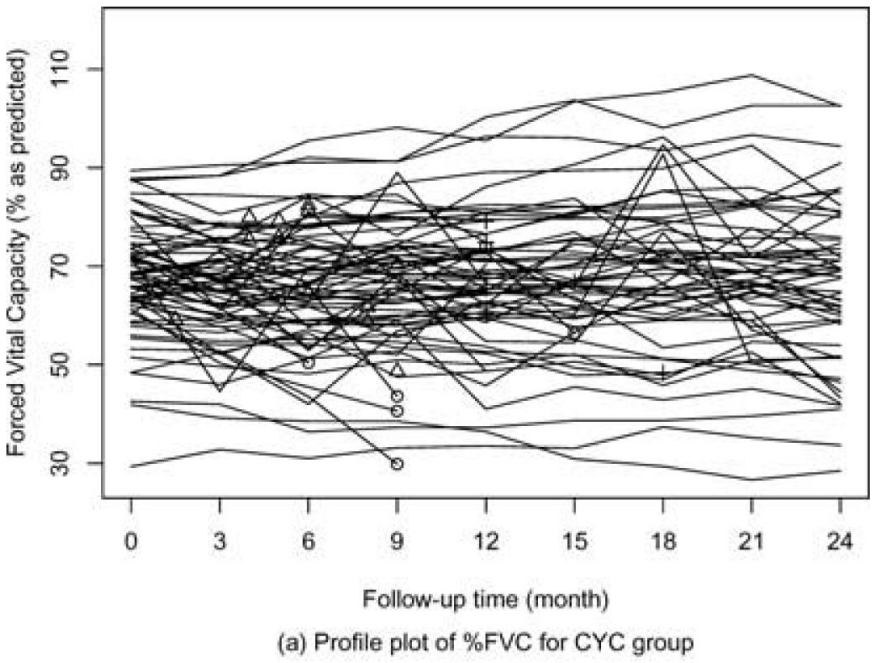
Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics 1997;53:330–339. [PubMed: 9147598]

Xu J, Zeger SL. Joint analysis of longitudinal data comprising repeated measures and times to events. Applied Statistics 2001;50:375–387.

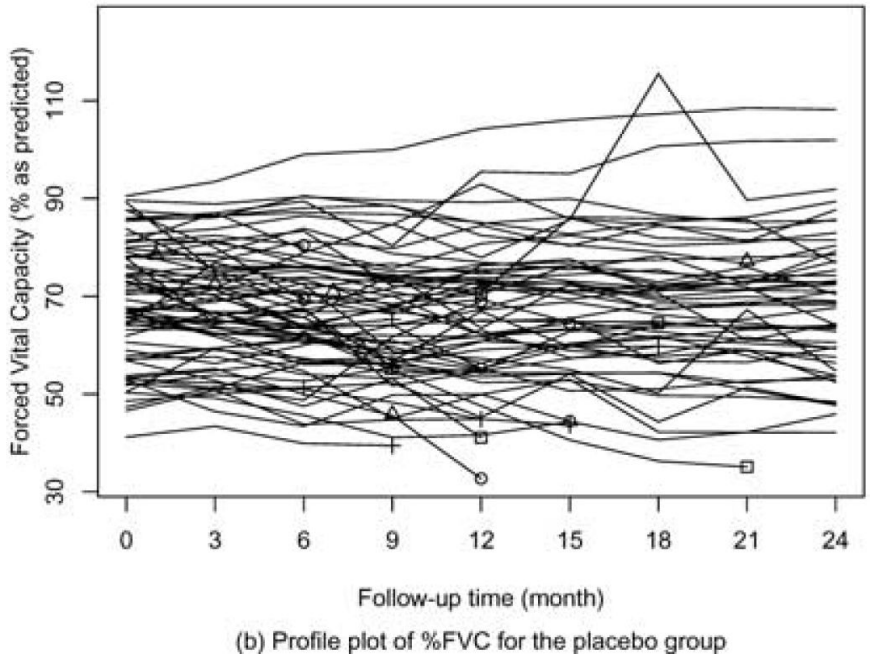
Zeger SL, Liang KY, Albert PS. Models for longitudinal data: A generalized estimating equation approach. Biometrics 1988;44:1049–1060. [PubMed: 3233245]

Zeng D, Cai J. Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time. The Annals ofStatistics 2005a;33:2132–2163.

Zeng D, Cai J. Simultaneous modelling of survival and longitudinal data with an application to repeated quality of life measures. Lifetime Data Analysis 2005b;11:151–174. [PubMed: 15940822]



(a) Profile plot of %FVC for CYC group



(b) Profile plot of %FVC for the placebo group

Figure 1.

(a)-(b) Profile plots of %FVC for CYC group versus the placebo group: ○ for treatment failure or death; + for informative censoring without %FVC measurements after the events; △ for informative censoring with %FVC measurements after the events; □for noninformatively censored events.

Table 1

Three types of events for subjects with %FVC measurements at 6 months or later in the scleroderma lung study

Category	CYC	Placebo	Total
(1) Treatment failure or death	5	9	14
(2) Informatively censored events (worsening disease, AE, SAE)			
No measurements after withdrawal	6	7	13
With measurements after withdrawal	12	7	19
(3) Noninformatively censored events (other reasons)	1	4	5

Table 2

Analysis of 6-21 months' scleroderma lung study data

	Joint analysis estimate (SE)	Separate analyses estimate (SE)
Longitudinal outcome %FVC		
Time (β_1)	-0.05 (0.08)	-0.04 (0.08)
FVC_0(β_2)	0.91 (0.04)*	0.90 (0.05)*
FIB_0(β_3)	-1.83 (0.61)*	-1.90 (0.54)*
CYC (β_4)	0.11 (1.00)	0.15 (1.06)
$FVC_0 \times CYC$ (β_5)	0.13 (0.07)	0.13 (0.07)
$FIB_0 \times CYC$ (β_6)	1.71 (0.80)*	1.85 (0.79)*
Time × *CYC* (β_7)	0.13 (0.11)	0.12 (0.11)
σ^2	24.78 (0.67)	24.72 (1.47)
σ_b^2	0.22 (0.03)	0.22 (0.03)
p-value for $H_0 : \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$	0.0271	0.0429
Cause-specific hazards (treatment failure or death)		
FVC_0(γ_{11})	0.02 (0.03)	0.02 (0.03)
FIB_0 (γ_{12})	0.18 (0.35)	0.17 (0.32)
CYC (γ_{13})	-0.69 (0.77)	-0.70 (0.70)
$FVC_0 \times CYC$ (γ_{14})	-0.05 (0.07)	-0.05 (0.07)
FIB_0 × CYC (γ_{15})	-0.47 (1.12)	-0.51 (0.98)
p-value for $H_0 : \gamma_{13} = \gamma_{14} = \gamma_{15} = 0$	0.7711	0.7292
Cause-specific hazards (informatively censored events)		
FVC_0(γ_{21})	-0.07 (0.03)*	-0.07 (0.03)*
FIB_0(γ_{22})	0.20 (0.33)	0.18 (0.32)
CYC (γ_{23})	0.31 (0.46)	0.31 (0.45)
$FVC_0 \times CYC$ (γ_{24})	0.11 (0.04)*	0.10 (0.04)*
$FIB_0 \times CYC$ (γ_{25})	0.12 (0.41)	0.10 (0.39)
Random effects for survival endpoint		
ν_2	1.93 (2.23)	-0.24 (> 100)
σ_u^2	0.05 (0.09)	0.20 (0.52)
Covariance of b_i and u_i		
σ_{bu}	-0.10 (0.10)	

* p-value < 0.05.

Table 3

Comparison of the joint model and the separate analyses when the association is negligible (sample size = 140)

Parameter	True	Separate				Joint			
		Est	SE	Est.SE	CP	Est	SE	Est.SE	CP
Longitudinal Fixed effects									
β_0	65.9	65.83	1.056	1.067	0.955	65.81	1.017	1.157	0.960
β_1	-0.05	-0.033	0.126	0.121	0.935	-0.041	0.125	0.131	0.950
β_2	0.9	0.901	0.052	0.055	0.940	0.901	0.052	0.059	0.965
β_3	-1.8	-1.786	0.598	0.602	0.955	-1.709	0.586	0.660	0.945
β_4	0.1	0.124	1.316	1.529	0.975	0.164	1.375	1.615	0.985
β_5	0.1	0.093	0.069	0.077	0.985	0.102	0.082	0.086	0.970
β_6	1.7	1.668	0.801	0.867	0.965	1.670	0.859	0.950	0.965
β_7	0.1	0.085	0.164	0.177	0.965	0.100	0.171	0.186	0.980
Random effects									
σ^2	24.8	24.75	1.877	2.069	0.950	24.57	2.032	2.213	0.945
σ_b^2	0.2	0.188	0.040	0.039	0.895	0.192	0.038	0.043	0.940
Survival Fixed effects									
γ_{11}	0.02	0.021	0.058	0.051	0.955	0.023	0.047	0.057	0.980
γ_{12}	0.2	0.297	0.797	0.566	0.975	0.282	0.514	0.593	0.990
γ_{13}	-0.7	-0.814	1.389	1.235	0.995	-0.835	1.316	1.503	0.990
γ_{14}	-0.05	-0.058	0.089	0.115	0.985	-0.057	0.113	0.143	0.980
γ_{15}	-0.5	-0.662	1.185	1.198	1.000	-0.674	1.105	1.552	0.995
γ_{21}	-0.07	-0.073	0.038	0.032	0.925	-0.078	0.031	0.035	0.995
γ_{22}	0.2	0.187	0.290	0.312	0.980	0.211	0.307	0.344	1.000
γ_{23}	0.3	0.288	0.471	0.450	0.960	0.308	0.455	0.495	0.990
γ_{24}	0.1	0.103	0.046	0.043	0.945	0.113	0.044	0.047	0.990
γ_{25}	0.1	0.093	0.411	0.422	0.980	0.085	0.434	0.476	0.990
Random effects									
ν_2	1.9	-0.129	3.567	> 100	0.995	0.180	4.162	2.775	0.810
σ_u^2	0.05	0.010	0.001	0.864	0.865	0.049	0.009	0.139	1.000

Parameter	True	Separate				Joint			
		Est	SE	Est.SE	CP	Est	SE	Est.SE	CP
Covariance									
σ_{bu}	-0.095	-	-	-	-	-0.093	0.018	0.144	0.995

Table 4

Comparison of the joint model and the separate analyses when there is a strong association (sample size = 200)

		Separate				Joint				
Parameter	**True**	**Bias**	**SE**	**Est. SE**	**CP**	**Bias**	**SE**	**Est. SE**	**CP**	**MSE_S/MSE_J**
Longitudinal										
Fixed effects										
β_0	10	-0.072	0.182	0.180	0.920	-0.020	0.184	0.190	0.960	1.118
β_1	-1	**0.357**	0.197	0.189	**0.520**	0.017	0.245	0.242	0.950	2.757
β_2	1.5	-0.003	0.244	0.240	0.935	0.003	0.244	0.250	0.960	1.000
β_3	0.6	**-0.131**	0.243	0.227	**0.875**	-0.009	0.247	0.248	0.945	1.248
Random effects										
σ^2	5	0.024	0.261	0.263	0.955	0.012	0.258	0.277	0.970	1.030
σ_b^2	0.5	**-0.073**	0.101	0.099	**0.790**	-0.006	0.125	0.133	0.950	0.992
Survival										
Fixed effects										
γ_{11}	0.8	-0.023	0.172	0.164	0.950	0.018	0.159	0.159	0.950	1.176
γ_{12}	-1	0.016	0.295	0.281	0.945	-0.046	0.267	0.286	0.975	1.189
γ_{21}	0.5	0.012	0.162	0.170	0.960	0.032	0.159	0.165	0.960	1.003
γ_{22}	-1	-0.059	0.341	0.337	0.955	-0.088	0.357	0.301	0.955	0.886
Random effects										
ν_2	0.5	**-0.185**	0.912	1.215	0.975	0.088	0.582	0.612	0.965	2.499
σ_u^2	0.5	-0.051	0.456	0.443	0.925	0.067	0.427	0.358	0.940	1.127
Covariance										
σ_{bu}	-0.45	-	-	-	-	0.006	0.206	0.210	0.950	-

Note: The large bias or large error in CP are highlighted in boldface.

Table 5

Comparison of the joint model and the separate analyses when there is a strong association (sample size = 500)

		Separate				Joint				
Parameter	**True**	**Bias**	**SE**	**Est. SE**	**CP**	**Bias**	**SE**	**Est. SE**	**CP**	**MSE_S/MSE_J**
Longitudinal										
Fixed effects										
β_0	10	-0.073	0.105	0.114	0.935	-0.018	0.107	0.117	0.955	1.389
β_1	-1	**0.369**	0.122	0.119	**0.150**	-0.011	0.159	0.150	0.940	5.946
β_2	1.5	0.005	0.147	0.151	0.955	0.010	0.147	0.154	0.975	0.997
β_3	0.6	**-0.137**	0.157	0.144	**0.815**	0.004	0.159	0.152	0.950	1.716
Random effects										
σ^2	5	0.009	0.174	0.165	0.935	0.002	0.173	0.169	0.945	1.014
σ_b^2	0.5	**-0.063**	0.066	0.065	**0.750**	-0.002	0.082	0.082	0.950	1.237
Survival										
Fixed effects										
γ_{11}	0.8	0.009	0.125	0.127	0.940	0.002	0.103	0.098	0.960	1.480
γ_{12}	-1	0.019	0.206	0.177	0.915	-0.023	0.169	0.176	0.965	1.471
γ_{21}	0.95	0.018	0.118	0.120	0.955	0.013	0.106	0.099	0.935	1.249
γ_{22}	-1	-0.034	0.207	0.219	0.955	-0.055	0.205	0.182	0.945	0.977
Random effects										
ν_2	0.5	**-0.124**	0.878	0.956	0.920	0.092	0.380	0.359	0.940	5.144
σ_u^2	0.5	0.057	0.541	0.497	0.915	0.032	0.286	0.225	0.920	3.573
Covariance										
σ_{bu}	-0.45	-	-	-	-	0.008	0.139	0.127	0.950	-

Note: The large bias or large error in CP are highlighted in boldface.