Published in final edited form as:
J Appl Stat. 2014 January 1; 41(10): 2192–2205. doi:10.1080/02664763.2014.909784.

Prediction of transplant-free survival in idiopathic pulmonary fibrosis patients using joint models for event times and mixed multivariate longitudinal data

Jiin Choi, Ph.D.[1], **Stewart J. Anderson, Ph.D.**[2], **Thomas J. Richards, Ph.D.**[3], and **Wesley K. Thompson, Ph.D.**[4]

[1]Biostatistics Collaboration Unit, Yonsei University College of Medicine, Seoul, Korea

[2]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania

[3]The Dorothy P. & Richard P. Simmons Center for Interstitial Lung Disease, Division of Pulmonary, Allergy, & Critical Care Medicine, Department of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania

[4]Department of Psychiatry, University of California San Diego, San Diego, California

SUMMARY

We implement a joint model for mixed multivariate longitudinal measurements, applied to the prediction of time until lung transplant or death in idiopathic pulmonary fibrosis. Specifically, we formulate a unified Bayesian joint model for the mixed longitudinal responses and time-to-event outcomes. For the longitudinal model of continuous and binary responses, we investigate multivariate generalized linear mixed models using shared random effects. Longitudinal and time-to-event data are assumed to be independent conditional on available covariates and shared parameters. A Markov chain Monte Carlo (MCMC) algorithm, implemented in OpenBUGS, is used for parameter estimation. To illustrate practical considerations in choosing a final model, we fit 37 different candidate models using all possible combinations of random effects and employ a Deviance Information Criterion (DIC) to select a best fitting model. We demonstrate the prediction of future event probabilities within a fixed time interval for patients utilizing baseline data, post-baseline longitudinal responses, and the time-to-event outcome. The performance of our joint model is also evaluated in simulation studies.

Keywords

Idiopathic Pulmonary Fibrosis; Joint model; Mixed continuous and binary data; Multivariate longitudinal data; Prediction model; Shared parameter model; Survival analysis

Current Address for correspondence: Stewart J. Anderson, Ph.D., Professor, Department of Biostatistics, University of Pittsburgh Graduate School of Public Health, 312 Parran Hall, 130 Desoto Street, Pittsburgh, PA 15213, Phone: (412) 6244-3026, Fax: (412) 624-2183, sja@pitt.edu.

1. Introduction

While longitudinal data from biomarkers or intermediate clinical assessments are often collected in Phase II and III clinical trials, the primary clinical trial outcome typically consists of time until the occurrence of a specified event, e.g., time to onset of illness or death. An example from the field of chronic lung disease research is the study of transplant-free survival in idiopathic pulmonary fibrosis (IPF). IPF is a chronic, progressive lung disease characterized by fibrosis of unknown etiology [24]. Here, our primary concern is the prediction of failure times by modeling the relationship between longitudinally measured indicators of pulmonary function and time to either lung transplant or death. Specifically, we utilize longitudinally assessed forced vital capacity (FVC), a continuous measure of pulmonary function, as a predictor of event times. We also utilize a second, binary, disease progression indicator defined as a decline of 5% or more in FVC from baseline FVC. Since these measures are potentially conjointly informative of time to lung transplant or death, it would be desirable to implement a modeling approach that allows prediction of event times using two or more longitudinal responses simultaneously.

Several approaches have been developed for handling multivariate longitudinal outcomes [4,20,16,23,29]. For example, some approaches factorize the joint distribution as the product of the marginal distribution of one variable times the conditional distribution of the second variable given the first [3,11]. Other approaches use random effects to model relationships between outcomes [8,13]. To date, however, most researchers have implemented joint models in which a univariate continuous longitudinal response is related to the time-to-event outcome [7,14,15,27]. In this paper, we develop a joint model for mixed continuous and binary longitudinal responses and a time-to-event outcome. For longitudinal responses we investigate mixed models with shared random effects. To build a joint model with a time-to-event outcome, we assume shared parameters exist between the longitudinal mixed models and the time-to-event outcome. The longitudinal responses and the time-to-event outcome are assumed to be independent conditional on the shared parameters and measured covariates.

We apply the proposed joint model to the dynamic prediction of mortality occurring within a fixed time for a subject still at risk just before time t [30]. A number of authors have proposed the prediction of future event probabilities for subjects based on the joint modeling of longitudinal measurements, time-to-event outcomes, and other covariates [9,25]. Fieuws *et al*. [9] investigated predicting renal graft failure using multiple longitudinal outcomes of biochemical and physiological markers utilizing a multivariate mixed model [22] and a pattern-mixture approach [19]. Rizopoulos [25] provided individualized prediction models of survival in AIDS patients who also had longitudinal CD4 cell count measurements. Our prediction approach is most closely related to that of Rizopoulos [25]. However, as is outlined in Section 2, we use a fully parametric model and a Bayesian approach similar to that of Guo and Carlin [14].

The remainder of the paper is organized as follows. In Section 2, we present our joint model assuming shared random effects for the longitudinal outcomes. In Section 3, we describe parameter estimation approach and model selection methods. In Section 4, we illustrate the

dynamic prediction of event probabilities within a fixed time window. Section 5 presents simulation studies using the proposed model, whereas Section 6 applies the joint models to prediction of failure times in idiopathic pulmonary fibrosis data. Section 7 concludes with a brief discussion.

2. Joint modeling approach

2.1 Notation and definitions

For brevity, we focus the development of the model on the case of bivariate longitudinal responses, one continuous and normally distributed and the other binary. The generalization to higher dimensions and other members of the exponential family of distributions is conceptually straightforward (see Discussion). For subject i, $i = 1,\cdots, n$, let y_{1ij} and y_{2ij} denote the j^{th} outcome at time point t_{ij} consisting of continuous and binary components, respectively. Further, let $\boldsymbol{y}_i=(\boldsymbol{y}_{1i}^\top, \boldsymbol{y}_{2i}^\top)^\top$ denote the bivariate longitudinal outcome vector for subject i, where $\boldsymbol{y}_{hi} = (y_{hi1},\cdots, y_{hij})^\top$, $h = 1,2$, $j = 1,\cdots, n_{hi}$ is an n_{hi}-dimensional column vector giving the h^{th} longitudinal outcomes for subject i. For the longitudinal bivariate response vector, $\boldsymbol{y}_i$, with different data types, we assume a generalized linear mixed effects model

$$E(\boldsymbol{y}_{hi}|\boldsymbol{b}_{hi})=g_h(\mathbf{X}_{hi}\boldsymbol{\beta}_h+\boldsymbol{Z}_{hi}\boldsymbol{b}_{hi}),\ \ h=1,2 \quad (1)$$

where $g_h(\cdot)$ denotes known bijective link functions that differ across data types, $\boldsymbol{X}_{hi}$ and $\boldsymbol{\beta}_h$ denote an $n_{hi} \times p_h$ design matrix of covariate values and a p_h-dimensional vector of fixed effects, respectively, and $\boldsymbol{Z}_{hi}$ and $\boldsymbol{b}_{hi}$ denote $n_{hi} \times q_1$ design matrix of covariates and a q_h-dimensional vector of normally distributed random effects with zero mean and covariance matrix as described below. We assume that the elements of $\boldsymbol{y}_{hi}$ are independent conditional on $\boldsymbol{b}_{hi}$ [10,12].

Here, we choose the identity link for the continuous response and the logit link for the binary response. Thus the generalized linear mixed effects model (1) can be written in the form

$$E(\boldsymbol{y}_{1i}|\mathbf{b}_{1i})=\boldsymbol{X}_{1i}\beta_1+\boldsymbol{Z}_{1i}\mathbf{b}_{1i} \quad (2a)$$

$$\text{logit}(\Pr(\boldsymbol{y}_{2i}=1)|\mathbf{b}_{2i})=\boldsymbol{X}_{2i}\boldsymbol{\beta}_2+\boldsymbol{Z}_{2i}\mathbf{b}_{2i}. \quad (2b)$$

We assume that $\boldsymbol{b}_{1i}$ follows a normal distribution with a mean vector of zeros and variance-covariance matrix, $\boldsymbol{\Sigma}$, and that $\boldsymbol{b}_{2i}$ is proportional to $\boldsymbol{b}_{1i}$, i.e., $\boldsymbol{b}_{2i} = \boldsymbol{A}_0\boldsymbol{b}_{1i}$, where $\boldsymbol{A}_0$ is a diagonal matrix of unknown constants. A joint model with this assumption is termed a *shared random effects* joint model. We briefly outline another alternative, also implemented in our code, in the Discussion section.

2.2 Joint model of multivariate longitudinal outcomes and a time-to-event outcome

Let T_i denote the true event time for subject *i*, C_i be the censoring time, and $\delta_i = I(T_i \leq C_i)$ be the event indicator. Let $T_i^* = \min(T_i, C_i)$ be the observed event time for subject *i*. We assume that censoring is non-informative [6]. A proportional hazards model is given by

$$\lambda_i(t|\boldsymbol{x}_{3i}, U_{3i}) = \lambda_0(t)\exp\{\boldsymbol{x}_{3i}^\top \boldsymbol{\beta}_3 + U_{3i}\}, \quad (3)$$

where $\boldsymbol{x}_{3i}$ is a p_3-dimensional vector of covariates with regression coefficients $\boldsymbol{\beta}_3$, and $\lambda_0(t)$ is the baseline hazard function, which can be assumed to be of a parametric form or left unspecified. To express the effects of longitudinal outcomes on the time-to-event outcome, the shared parameters, U_{3i}, are associated with the random effects of longitudinal outcomes $\boldsymbol{b}_{1i}$ and $\boldsymbol{b}_{2i}$. The *joint model* connects the longitudinal response submodels (2a,b) and the event time outcome submodel (3)

$$U_{3i} = \boldsymbol{a}_1^\top \boldsymbol{b}_{1i} + \boldsymbol{a}_2^\top \boldsymbol{b}_{2i} + b_{3i}, \quad (4)$$

where $\boldsymbol{a} = (\boldsymbol{a}_1^\top, \boldsymbol{a}_2^\top)^\top$ is a set of unknown constants and b_{3i} is a normally distributed frailty term with mean zero and variance σ_3^2, independent of the $\boldsymbol{b}_i = (\boldsymbol{b}_{1i}^\top, \boldsymbol{b}_{2i}^\top)^\top$. The hazard function of the time-to-event depends on the longitudinal outcomes through the shared $\boldsymbol{b}_{1i}$ and $\boldsymbol{b}_{2i}$. Thus, the parameter $\boldsymbol{a}$ quantifies the degree of association explained by the random effects in (2a,b). In terms of a *shared random effects* joint model, Eq. (4) can be reduced in the form

$$U_{3i} = \boldsymbol{a}_1^\top \boldsymbol{b}_{1i} + b_{3i}.$$

The components of $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$ may or may not all be the same, allowing the longitudinal continuous and binary responses, and the time-to-event outcome to depend on different and/or overlapping covariate information.

3. Estimation and model selection

3.1 Parameter estimation

We assume that the longitudinal responses vectors $\boldsymbol{y}_i$ and the time-to-event outcome T are independent conditional on covariates $\boldsymbol{X}$, $\boldsymbol{Z}$, and the random effects vector, $\boldsymbol{b}$. The observed data for the i^{th} subject, with n_i repeated measurements, for the h^{th} response are denoted by

$$\{y_{hi}(t_{ij}),\ X(t_{ij}),\ Z(t_{ij}),\ T_i^*,\ \delta_i; h=1,2,\ i=1,\cdots,n,\ j=1,\cdots,n_{hi}\}.$$

Based on a full conditional independence assumption, we can express the joint distribution of the observed data for the i^{th} subject as

$$f(\boldsymbol{y}_i, T_i^*, \delta_i|\boldsymbol{b}_i;\boldsymbol{\theta}) = f(\boldsymbol{y}_{1i}|\boldsymbol{b}_{1i};\boldsymbol{\beta}_1, \sigma^2_{\varepsilon_1}) f(\boldsymbol{y}_{2i}|\boldsymbol{b}_{2i};\boldsymbol{\beta}_2) f(T_i^*, \delta_i|\boldsymbol{b}_i;\boldsymbol{\beta}_3, \boldsymbol{a}, \sigma^2_3)$$

where $\boldsymbol{\theta}$ denotes the complete parameter vector and $f(\cdot)$ denotes a generic probability density function. Thus, the log-likelihood for the observed data is given by

$$\begin{aligned}
&\sum_{i=1}^{n} \log \int_{\boldsymbol{b}_i} \prod_{j=1}^{n_{1i}} \left[\frac{1}{\sqrt{2\pi\sigma^2_{\varepsilon_1}}} \exp\left\{ -\frac{1}{2\sigma^2_{\varepsilon_1}} (y_{1ij} - \boldsymbol{x}_{1ij}\boldsymbol{\beta}_1 - \boldsymbol{z}_{1ij}\boldsymbol{b}_{1i})^2 \right\} \right] \\
&\times \prod_{j=1}^{n_{2i}} \left[\left\{ \frac{1}{1+\exp(-\boldsymbol{x}_{2ij}\boldsymbol{\beta}_2 - \boldsymbol{z}_{2ij}\boldsymbol{b}_{2i})} \right\}^{y_{2ij}} \left\{ \frac{\exp(-\boldsymbol{x}_{2ij}\boldsymbol{\beta}_2 - \boldsymbol{z}_{2ij}\boldsymbol{b}_{2i})}{1+\exp(-\boldsymbol{x}_{2ij}\boldsymbol{\beta}_2 - \boldsymbol{z}_{2ij}\boldsymbol{b}_{2i})} \right\}^{1-y_{2ij}} \right] \\
&\times [\lambda_0(T_i^*)\exp\{\boldsymbol{x}_{3i}^\top\boldsymbol{\beta}_3 + U_i\}]^{\delta_i} \exp\left\{ -\int_0^{T_i^*} \lambda_0(s)\exp\{\boldsymbol{x}_{3i}^\top\boldsymbol{\beta}_3 + U_i\} ds \right\} \\
&\times f(\boldsymbol{b}_i|\textstyle\sum) d\boldsymbol{b}_i
\end{aligned}$$

where $\boldsymbol{x}_{hij}$ and $\boldsymbol{z}_{hij}$ are j^{th} row of matrices $\boldsymbol{x}_{hi}$ and $\boldsymbol{z}_{hi}$, respectively, and $f(\boldsymbol{b}_i|\boldsymbol{\Sigma})$ is the multivariate normal density function of $\boldsymbol{b}_i$ conditional on the covariance parameter matrix, $\boldsymbol{\Sigma}$.

We implement a Bayesian approach for parameter inferences, using a Gibbs sampling algorithm. The algorithm was programmed using the R interface, *rbugs*, which accesses the software 'OpenBUGS' [21]. We implement the Gibbs sampler using weakly informative priors for all parameters. Fixed effects are normal with large variance components. Error variances and random effect covariances are distributed as inverse gamma and inverse Wishart distributions, respectively [2]. Code used in the analyses is available from the first author upon request.

3.2 Model selection criteria

Spiegelhalter *et al*. [28] proposed the Deviance Information Criterion (DIC) for model selection. The DIC is a generalization of the Akaike Information Criterion (AIC) [1] for hierarchical models based on the *deviance* of the posterior distribution. For parameter vector θ and observed data vector, $\boldsymbol{y}$, let $D(\theta)$ be the deviance, $D(\theta) = -2 \log f(\boldsymbol{y}|\theta) + 2 \log h(\boldsymbol{y})$, where $f(\boldsymbol{y}|\theta)$ is the likelihood function and $h(\boldsymbol{y})$ is a standardizing function of the data alone [2]. We define $D(\theta)$ to be $D(\boldsymbol{\theta}) = -2\sum_{i=1}^{n} \log f(\mathbf{y}_i, T_i^*, \delta_i;\boldsymbol{\theta})$ in our joint model, $\overline{D(\boldsymbol{\theta})} = E_{\theta|y}[D(\boldsymbol{\theta})]$ to denote the posterior expected deviance and $\bar{\boldsymbol{\theta}} = E_{\theta|y}[\boldsymbol{\theta}]$ to denote the mean of the posterior distribution of the parameters. The effective number of parameters, p_D, which can capture the complexity of a model is $p_D = \overline{D(\boldsymbol{\theta})} - D(\bar{\boldsymbol{\theta}})$. Then, the DIC is defined as $DIC = D(\bar{\boldsymbol{\theta}}) + 2p_D = \overline{D(\boldsymbol{\theta})} + p_D$. Smaller values of the DIC indicate better fitting models. As with the AIC, differences of the DIC between models are a tool used for model selection. Differences of 3 to 5 are considered meaningful. In this paper, we use the DIC for model selection.

4. Dynamic prediction of event probabilities

We compute dynamic probabilities of an event occurring within a fixed window as proposed by van Houwelingen and Putter [30]. Specifically, we considered the conditional survival function of time $v \geq t$ given survival to time t:

$$S(v|t)=P(T>v|T \geq t, \boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta}, v \geq t)=\frac{S(v|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta}, v \geq t)}{S(t|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta})}. \quad (5)$$

Equation (5) can also be expressed using a hazard function as

$$\begin{aligned} S(v|t)&=\exp[-\{\Lambda(v|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta})-\Lambda(t|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta})\}] \\ &=\exp\left(-\int_t^v \lambda(s|\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{b};\boldsymbol{\theta})ds\right) \end{aligned} \quad (6)$$

where $\Lambda(\cdot)$ denotes the cumulative hazard function and $\lambda(\cdot)$ denotes the instantaneous hazard function. Equation (6) implies that only the hazard on the interval $[t, v]$ is necessary to predict the probability of event up to time v for a subject at risk just before time t. Let $v = t + w$ be a fixed window of width w. We can relate the survival function to the cumulative distribution function as follows:

$$F_w(t)=1-S(t+w|t).$$

$F_w(t)$ is evaluated at all-time points t where the estimates change value. The variance of this function is based on the Nelson-Aalen estimate of that cumulative hazard that is given as

$$se^2[\ln\{\hat{F}_w(t)\}]=\sum_{\substack{t \leq t_i \leq t+w \\ t_i \in \mathfrak{D}}} \frac{1}{R(t_i)^2}$$

where $\mathfrak{D}$ denotes the set of event times and $R(t)$ denotes the size of the risk set, i.e., the number of subjects with no event and still being followed just before time t.

5. Simulation Studies

We conducted simulation studies to investigate the performance of the proposed joint models with the longitudinal continuous and binary outcomes and the time-to-event outcome. The simulated longitudinal data consisted of a quantitative outcome and a dichotomous outcome with seven repeated measurements at fixed times 0, 0.5, 1, 1.5, 2, 2.5, and 3 years. We considered three sample sizes, N=100, N = 200 and N = 500 subjects. Both longitudinal outcomes depended on the same fixed covariates: a continuous variable (x_1) sampled from the normal distribution, $x_1 \sim N(0, 5^2)$, a dichotomous variable (x_2) sampled from the Bernoulli distribution with probability 0.5, $x_2 \sim Bernoulli(0.5)$, and time (t), $t = \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$, were included as fixed covariates. Subject-specific random intercepts and slopes were assumed. For the longitudinal continuous outcome, the measurement error

term was normally distributed with mean zero and variance, $\sigma^2_{\varepsilon_1}=1$. The continuous longitudinal outcome was generated from the model

$$y_1(t)=\beta_{11}+\beta_{12}x_1+\beta_{13}x_2+\beta_{14}t+b_{10}+b_{11}t+\varepsilon_1(t)$$

The binary longitudinal outcome was generated from the model

$$\text{logit}(\Pr(y_2(t)=1))=\beta_{21}+\beta_{22}x_1+\beta_{23}x_2+\beta_{24}t+\alpha_1 b_{10}+\alpha_2 b_{11}t$$

and the random effects were assumed to have distribution

$$\begin{pmatrix} b_{10} \\ b_{11} \end{pmatrix} \sim \text{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma^2_{b_{10}} & \rho_{12}\sigma_{b_{10}}\sigma_{b_{11}} \\ & \sigma^2_{b_{11}} \end{bmatrix}\right).$$

The time-to-event outcome was generated from an exponential distribution, $T \sim exp(\mu_3)$, where $\mu_3 = \exp(\beta_{31} + \beta_{32}x_2 + \alpha_3 b_{10} + \alpha_4 b_{11} + b_3)$ and $b_3 \sim N\left(0, \sigma^2_{b_3}\right)$ with $\sigma^2_{b_3}=0.25$. A non-informative censoring time C was generated from a uniform distribution on [0.2, 2] which resulted in roughly 35% censoring on average. For each simulation study, 200 replications were performed. In each analysis, a total of 15,000 MCMC iterations were used, discarding the first 5,000 iterations as a burn-in.

Table 1 shows the results of the simulation studies, including true parameter values, bias (defined as the true parameter minus the mean estimated parameter), standard errors of the parameter estimates (SE), mean squared error (MSE), and the coverage probability of the estimated 95% credibility intervals (CP). With a few exceptions, most parameters in the joint models show acceptably low levels of bias and good coverage probabilities. For N=100, the biases of variance and covariance of random effects are relatively higher than for larger sample sizes, and it might create very low coverage probabilities. For N=200, CP averages close to 95% for all parameters. Most parameters have very small biases, with the exception of somewhat higher biases in the α parameters. However, biases in α are fairly small as a proportion of the size of the parameter, and decrease with larger sample sizes suggesting that a bias of the links between the longitudinal and time-to-event processes are strongly related to the number of events in the survival outcome. In general, as expected, larger sample sizes (N=500) show better results as indicated by the smaller bias, SE, MSE, and less variable CP that average very close to the nominal 95% level.

6. Application to mortality in idiopathic pulmonary fibrosis outcomes study

6.1 Data description and Preliminary Analyses

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive lung disease characterized by fibrosis of unknown etiology [24]. Our primary concern is prediction of failure times, defined as time to either lung transplant or death, utilizing longitudinally measured forced

vital capacity (FVC), a continuous measure of pulmonary function. Individual profiles for the pulmonary function assessed by FVC are shown in Figure 1, with subjects randomly apportioned to six subplots in order to more clearly depict these trajectories. An additional disease progression indicator, defined as a decline of 5% or more in FVC from the baseline FVC measurement, is also used as a (binary) longitudinal response. Baseline covariates also considered for inclusion in the models were: age centered at time of diagnosis, baseline FVC, smoking, gender, and time in years.

Of the 125 patients considered in the study, 64 (51.2%) patients died or had lung transplants. Patient characteristics are summarized in Table 3. This dataset was comprised of irregular follow-up times across patients, but generally patients were followed every 3–4 months according to standard of care. The time origin for each patient was the baseline assessment of pulmonary function, performed upon evaluation at the Simmons Center for Interstitial Lung Disease (ILD). The FVC of each patient was repeatedly measured about 11 times on average, up to 43 times. Follow-up information for the study was reported through December 31, 2011. The average follow-up time was 3.7 years among all patients and 4.1 years among patients who did not die nor received a lung transplant. All IPF patients were evaluated at the University of Pittsburgh Medical Center and clinical data were obtained from the Simmons Center for ILD.

Fitted survival curves of the Weibull model and exponential model are shown in Figure 2. The survival curves of the Weibull model and exponential both show a good fit to the marginal survival function. Because the Weibull and exponential models are almost identical we used the more parsimonious exponential model. The estimated median transplant-free survival time from the initial visit date was 4.4 years.

6.2 Joint model

As before, let y_{1ij} be the continuous longitudinal outcome indicating FVC measurements, y_{2ij} be the binary longitudinal outcome indicating disease progression indicator, and T_i be the transplant-free survival time. More specifically, the model is given by

$$\begin{gathered}
y_{1ij}=\mu_{1i}(t_{ij})+U_{1i}(t_{ij})+\varepsilon_{1ij} \\
=\beta_{11}+\beta_{12}\text{BaselineFVC}_i+\beta_{13}\text{Time}_{ij}+b_{10i}+b_{11i}\text{Time}_{ij}+\varepsilon_1(t), \\
\begin{pmatrix} b_{10i} \\ b_{11i} \end{pmatrix} \sim \text{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma^2_{b_{10}} & \rho_{b_1}\sigma_{b_{10}}\sigma_{b_{11}} \\ \rho_{b_1}\sigma_{b_{10}}\sigma_{b_{11}} & \sigma^2_{b_{11}} \end{bmatrix}\right), \quad \varepsilon_{1ij}\sim N\left(0,\, \sigma^2_{\varepsilon_1}\right) \\
\text{logit}\left(\Pr(y_{2ij}=1)\right)=\mu_{2i}(t_{ij})+U_{2i}(t_{ij}) \\
=\beta_{21}+\beta_{22}\text{Age}_i+\beta_{23}\text{Time}_{ij}+\alpha_1 b_{10i}+\alpha_2 b_{11i}\text{Time}_{ij}, \\
T_i\sim\text{Exponential}(\mu_3), \\
\log(\mu_{3i})=\mathbf{x}_{3i}^T\boldsymbol{\beta}_3+U_{3i}=\beta_{31}+\beta_{32}\text{Gender}_i+\beta_{33}\text{Smoking}_i+\beta_{34}\text{BaselineFVC}_i+\alpha_3 b_{10i}+\alpha_4 b_{11i}+b_{3i}.
\end{gathered}$$

Under the shared random effects joint model, the random effects terms of the three submodels are defined as follows:

- $U_{1i}(t_{ij}) = b_{10i} + b_{11i}t_{ij}$
- $U_{2i}(t_{ij}) = \alpha_1 b_{10i} + \alpha_2 b_{11i}t_{ij}$

- $U_{3i} = \alpha_3 b_{10i} + \alpha_4 b_{11i} + b_{3i}$

The association between longitudinal responses and time-to-event outcome is explained by the parameter $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. For the continuous longitudinal response, we included the covariates of baseline FVC and time in years. For the binary longitudinal response, age and time in years were found to be significant covariates. Gender, smoking, and baseline FVC were significant in the model for transplant-free survival.

Non-informative priors were used for all parameters. For the coefficients of the fixed effects $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, and $\boldsymbol{\beta}_3$, we assume a multivariate normal distribution with mean zero and variance-covariance matrices $100\boldsymbol{I}_3$, $100\boldsymbol{I}_3$, and $100\boldsymbol{I}_4$, respectively, where $\boldsymbol{I}_k$ indicates an identity matrix with dimension k. We assumed that the association coefficient, $\boldsymbol{\alpha}$, had a normal distribution with mean zero and variance equal to 100. A total of 15,000 MCMC iterations were used discarding the first 5,000 iterations for the burn-in period.

6.3 Results

We considered all possible combinations of random effects resulting in 37 assessed models. Table 3 summarizes DICs for each submodel, the posterior expected deviance, $\bar{D}$, the effective number of parameters, p_D, and total DIC scores (DIC_{total}) for each joint model. Note that DIC_{y1}, DIC_{y2}, and DIC_T denote DIC from the longitudinal continuous outcome, the longitudinal binary outcome, and the time-to-event outcome submodels, respectively. These scores show relative contributions of each component to the overall model DIC. DIC_{total}, was used to choose the “best” model.

Joint models with both subject-specific random intercepts and slopes in both the longitudinal continuous outcome submodel and binary outcome submodel (Models 29–37) have the smallest DICs, and hence the smallest DIC_{total} compared with other joint models with random effect structures. In the time-to-event submodel, adding a frailty term leads to only a very slight decrease in DIC_T. Model 32 with the smallest DIC_{total} was selected as our best model. Under Model 32, the longitudinal binary outcome shares the subject-specific random intercept and slope of the longitudinal continuous outcome and the time-to-event outcome is related to the subject-specific random slope of the longitudinal continuous outcome.

In Table 4, the posterior estimates for Model 32 parameters along with their standard errors and 95% credibility intervals are presented. For longitudinal continuous FVC, both the baseline FVC and time are statistically significant. Hence, patients with high values of baseline FVC have high FVC measurements and FVC measurements tend to decrease as time progresses. For the longitudinal binary outcome, there is a significant time effect. For the time-to-event outcome, male smokers with lower baseline FVC have a higher risk of death or transplant. The association parameter, α_4 indicates that there is a nearly significant negative association between the subject-specific random slopes of FVC measurements and the hazard of the transplant-free survival outcome.

Figure 3 shows the predicted four-year-ahead failure probabilities for the IPF data set along with 95% confidence intervals, obtained from the best-fitting Model 32. In general, the probability of event within the next four years decreases slowly, with a deceleration

apparent if the patient survives longer than two years, so that at baseline subjects have ~45% chance of failure in four years, whereas if a patient does not experience failure by year five, the probability of a failure in the next four years decreases to ~35%. As time progresses, confidence intervals become wider because of the sparsity of the data.

7. Discussion

In this paper, we have developed a joint model for multivariate longitudinal continuous and binary responses and a time-to-event outcome, applying this model to the prediction of failure times from idiopathic pulmonary fibrosis. This general approach was introduced in Rizopoulos and Ghosh [26] who used natural splines to model three mixed longitudinal processes (two continuous and one binary) as predictors of a time-to-event outcome. In our case, we used parametric models for the longitudinal processes as the number of longitudinal observations per individual was relatively small and hence parametric random effect models were more tractable. Simulation studies indicate good performance for our models.

Our joint model was developed utilizing a unified Bayesian framework, including model fitting, parameter inferences, computation of dynamic prediction probabilities with posterior credible intervals, and a DIC model selection criterion that parses out the contributions of the joint sub-models to the overall fit. Results from the best fitting model indicate that the two longitudinal responses jointly contribute nearly significantly to prediction of failure times in IPF as indicated by the value of the shared parameter, α_4, and its 95% credibility interval (Table 4). Our approach for choosing prior distributions followed that of Guo and Carlin [14], which is quite general and easier to implement than that used by Rizopoulos and Ghosh [26].

In our present investigation, we utilized shared random effects to characterize the relationship between the longitudinal continuous and binary outcomes. A disadvantage of the shared random effects joint model is that the structure of the random effects for both the longitudinal outcomes is limited. Nevertheless, shared parameter random effects models do gain efficiency by using a smaller number of random effects parameters compared to other joint model formulations.

A practical purpose our investigation was to illustrate the utility of joint models for computing dynamic predictions of the probabilities of events occurring within a fixed window of time. Given a subject is at risk just before time *t*, the probability of an event occurring within the next fixed window of time is predicted using the survival function obtained from our proposed joint models. By predicting the probability of event within a fixed window of time we can easily visualize conditional probabilities of events. Plots such as these can be an effective means for presenting dynamic risk of failure times to clinician researchers. Finally, we have implemented our shared random effects joint models in an R interface using OpenBugs, available from the first author upon request.

Finally, in related work, we have implemented a *correlated random effects* joint model [5] which we plan to investigate further. As part of this future investigation, we wish to extend models proposed by Lee and Nelder [17,18] that expand the correlated random effects

models by introducing a general method called maximum *h*-likelihood estimation. In the correlated random effects model, the relationship between the continuous and binary outcomes can be investigated through correlations among the random effects $\boldsymbol{b}_{1i}$ and $\boldsymbol{b}_{2i}$.

Acknowledgments

The authors would like to thank Drs. Naftali Kaminski and Charles F. Reynolds III for providing data which helped motivate the methods demonstrated in this manuscript. This work was partially supported by NCI grant numbers U10-CA069974 and U10-CA069651 (Anderson) and NIMH grant number P30-MH-090333 (Anderson and Thompson).

References

1. Akaike, H. Information theory and an extension of the maximum likelihood principle. In: Petrov, BN.; Csáki, F., editors. Proc 2nd Int Symp Information Theory. Budapest: Akadémiai Kiadó; 1973. p. 267-281.
2. Carlin, BP.; Louis, TA. Bayesian Methods for Data Analysis. 3. Chapman and Hall/CRC Press; Boca Raton, FL: 2009.
3. Catalano PJ, Ryan LM. Bivariate latent variable models for clustered discrete and continuous outcomes. J Amer Statist Assoc. 1992; 87:651–658.
4. Chi YY, Ibrahim JG. Joint models for multivariate longitudinal and multivariate survival data. Biometrics. 2006; 62:432–445. [PubMed: 16918907]
5. Choi, J-I. PhD dissertation. University of Pittsburgh; 2012. Prediction in the Joint Modeling of Mixed Types of Multivariate Longitudinal Outcomes and a Time-to-Event Outcome.
6. Collett, D. Modelling Survival Data in Medical Research. Chapman & Hall; London: 1994.
7. Faucett CJ, Thomas DC. Simultaneously modeling censored survival data and repeated measured covariates: A Gibbs sampling approach. Stat Med. 1996; 15:1663–1685. [PubMed: 8858789]
8. Fieuws S, Verbeke G. Joint modelling of multivariate longitudinal profiles: pitfalls of the random-effects approach. Stat Med. 2004; 23:3093–3104. [PubMed: 15449333]
9. Fieuws S, Verbeke G, Maes B, Vanrenterghem Y. Predicting renal graft failure using multivariate longitudinal profiles. Biostatistics. 2008; 9:419–431. [PubMed: 18056686]
10. Fitzmaurice, GM.; Davidian, M.; Verbeke, G.; Molenberghs, G. Longitudinal Data Analysis. Chapman and Hall/CRC Press; Boca Raton, FL: 2009.
11. Fitzmaurice GM, Laird NM. Regression models for a bivariate discrete and continuous outcome with clustering. J Amer Statist Assoc. 1995; 90:845–852.
12. Fitzmaurice, GM.; Laird, NM.; Ware, J. Applied Longitudinal Data. Wiley; Hoboken, New Jersey: 2004.
13. Gueorguieva RV, Sanacora G. Joint analysis of repeatedly observed continuous and ordinal measures of disease severity. Stat Med. 2006; 25:1307–1322. [PubMed: 16217846]
14. Guo X, Carlin BP. Separate and joint modeling of longitudinal and event time data using standard computer packages. Amer Statist. 2004; 58:16–24.
15. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. Biostatistics. 2000; 1:465–480. [PubMed: 12933568]
16. Hogan JW, Laird NM. Mixture models for the joint distribution of repeated measures and event times. Stat Med. 1997; 16:239–257. [PubMed: 9004395]
17. Lee Y, Nelder JA. Hierarchical generalized linear models. Journal of the Royal Statistical Society Series B (Methodological). 1996:619–678.
18. Lee, Y.; Nelder, JA.; Pawitan, Y. Generalized linear models with random effects: unified analysis via H-likelihood. CRC Press; 2006.
19. Little RJA. Pattern-mixture models for multivariate incomplete data. J Amer Statist Assoc. 1993; 88:125–134.

20. Liu X, Daniels MJ, Marcus B. Joint models for the association of longitudinal binary and continuous processes with application to a smoking cessation trial. J Amer Statist Assoc. 2010; 104:429–438.
21. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions (with discussion). Stat Med. 2009; 28:3049–3082. [PubMed: 19630097]
22. Molenberghs, G.; Verbeke, G. Models for Discrete Longitudinal Data. Springer; New York: 2005.
23. Molenberghs G, Verbeke G, Demetrio CGB, Vieira AM. A family of generalized linear models for repeated measures with normal and conjugate random effects. Statist Sci. 2010; 25:325–347.
24. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO, Klesen M, Zhang Y, Gibson KF. Peripheral Blood Proteins Predict Mortality in Idiopathic Pulmonary Fibrosis. American Journal of Respiratory and Critical Care Medicine. 2012; 185:67–76. [PubMed: 22016448]
25. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. Biometrics. 2011; 67:819–829. [PubMed: 21306352]
26. Rizopoulos D, Ghosh P. A Bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event. Stat Med. 2011; 30:1366–1380. [PubMed: 21337596]
27. Song X, Davidian M, Tsiatis AA. A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. Biometrics. 2002; 58:742–753. [PubMed: 12495128]
28. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian Measures of Model Complexity and Fit. J R Stat Soc Ser B Stat Methodol. 2002; 64:583–639.
29. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: An overview. Statist Sinica. 2004; 14:809–834.
30. van Houwelingen, HC.; Putter, H. Dynamic Prediction in Clinical Survival Analysis. Chapman and Hall/CRC Press; Boca Raton, FL: 2011.

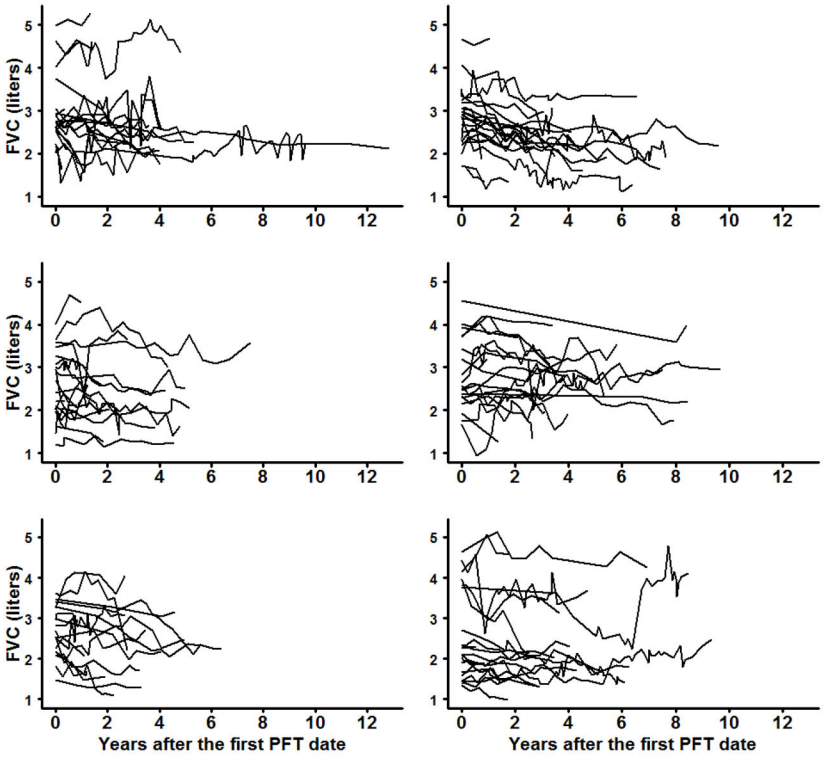

Figure 1.
Subject-specific trajectories of FVC measurements. (PFT: pulmonary function tests)

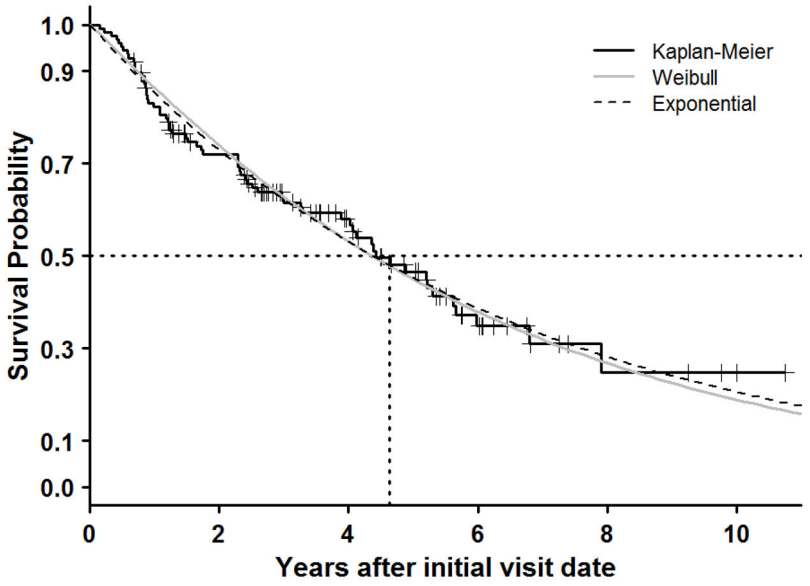

Figure 2.

The Kaplan-Meier plot of transplant-free survival (*black solid curve*); Weibull survival curve (*grey solid curve*) and exponential survival curve (*black broken curve*).

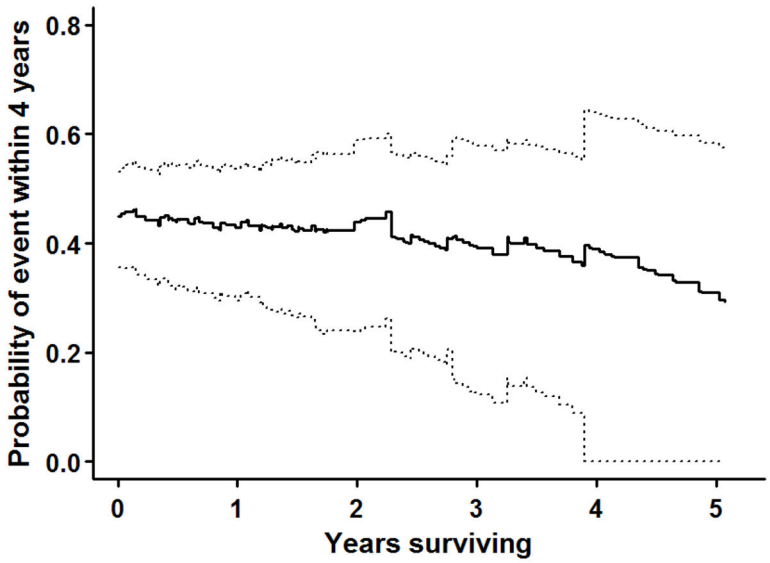

Figure 3.

Probability of transplant or death within the next four years using shared random effects joint model 32: Point estimates (*solid curve*) and 95 % confidence intervals (*dotted curves*).

Table 1

Simulation results

		N = 100				N = 200				N = 500			
Parameter	***True***	***Bias***	***SE***	***MSE***	***CP***	***Bias***	***SE***	***MSE***	***CP***	***Bias***	***SE***	***MSE***	***CP***
β_{11}	5.0	−0.002	0.104	0.024	0.930	−0.015	0.078	0.014	0.915	0.003	0.051	0.005	0.935
β_{12}	0.8	−0.005	0.138	0.047	0.920	0.008	0.107	0.029	0.880	−0.010	0.071	0.010	0.955
β_{13}	−0.2	0.000	0.012	0.000	0.925	0.001	0.009	0.000	0.925	0.000	0.006	0.000	0.950
β_{14}	−0.2	−0.004	0.056	0.007	0.930	−0.005	0.044	0.004	0.910	−0.004	0.028	0.002	0.945
β_{21}	0.3	0.018	0.223	0.109	0.940	−0.009	0.164	0.054	0.945	−0.003	0.103	0.022	0.960
β_{22}	0.3	−0.016	0.266	0.163	0.920	0.013	0.198	0.084	0.930	0.001	0.125	0.029	0.965
β_{23}	−0.2	−0.005	0.028	0.002	0.965	−0.001	0.020	0.001	0.970	−0.001	0.012	0.000	0.940
β_{24}	−0.25	−0.013	0.114	0.028	0.930	−0.015	0.083	0.015	0.920	−0.004	0.052	0.006	0.925
β_{31}	0.2	−0.023	0.345	0.266	0.930	−0.074	0.266	0.156	0.940	−0.021	0.171	0.059	0.950
β_{32}	1.0	0.003	0.419	0.398	0.945	0.031	0.307	0.216	0.915	−0.002	0.194	0.072	0.985
a_1	1.6	0.224	0.536	0.632	0.945	0.280	0.331	0.296	0.890	0.147	0.198	0.098	0.870
a_2	1.2	0.154	0.244	0.147	0.935	0.040	0.161	0.055	0.935	−0.012	0.100	0.022	0.930
a_3	2.0	0.181	0.891	1.308	0.975	0.164	0.585	0.595	0.980	0.052	0.344	0.211	0.995
a_4	3.5	0.369	0.647	0.836	0.975	0.153	0.457	0.379	0.985	0.087	0.291	0.158	0.960
$\sigma^2_{\varepsilon_1}$	1.0	0.029	0.061	0.008	0.925	0.023	0.043	0.004	0.940	0.008	0.027	0.001	0.975
$\sigma^2_{b_3}$	0.25	0.005	0.058	0.003	1.000	0.001	0.055	0.003	1.000	−0.011	0.049	0.003	1.000
$\sigma^2_{b_{10}}$	0.2	−0.073	0.020	0.006	0.010	−0.024	0.026	0.002	0.865	0.019	0.026	0.001	0.955
$\sigma^2_{b_{11}}$	0.25	−0.081	0.022	0.007	0.080	−0.016	0.025	0.001	0.895	0.013	0.021	0.001	0.950
ρ_{12}	0.5	−0.220	0.088	0.060	0.165	−0.027	0.067	0.007	0.995	0.043	0.052	0.006	0.945

Table 2

Patient characteristics of Idiopathic Pulmonary Fibrosis (IPF) outcome data

	N(=125)	%
Gender		
Male	84	67.2
Female	41	32.8
Race		
White	122	97.6
Black	1	0.8
American Indian	1	0.8
Oriental	1	0.8
Smoking		
Ever	88	70.4
Never	37	29.6
Diagnosis made		
Clinically	63	50.4
Historically	62	49.6
Transplant		
Yes	23	18.4
No	102	81.6
	Mean	**SD**
Age (year)		
Overall	65.2	9.34
Male	65.7	9.00
Female	64.2	10.04
Follow-up (year)		
All patients	3.7	2.32
Alive and not transplanted	4.1	2.35
Baseline PFTs		
FVC	2.7	0.85
FEV1	2.2	0.65
DLCO	12.7	4.36

Definition of abbreviations: PFT = pulmonary function tests; FVC = forced vital capacity; FEV1 = forced expiratory volume in 1 second; DLCO = diffusing capacity of carbon monoxide.

Table 3

Bayesian Model Selection of shared random effects joint model

Model	$U_1(t)$	$U_2(t)$	U_3	DIC_{y1}	DIC_{y2}	DIC_T	$\bar{D}$	p_D	DIC_{total}
	no random effects								
1	0	0	0	1795	1715	356.2	3856	11.02	3867
2	0	0	b_3	1795	1715	354.3	3842	23.11	3865
	random intercepts								
3	b_{10}	0	0	763.8	1715	356.1	2719	115.7	2835
4	b_{10}	0	b_3	763.9	1715	354.5	2706	127.9	2833
5	b_{10}	0	a_3b_{10}	763.6	1715	357.3	2720	116.3	2836
6	b_{10}	0	$a_3b_{10} + b_3$	763.6	1715	355.2	2706	128.2	2834
7	b_{10}	a_1b_{10}	0	719.5	1189	356.2	2143	122.5	2265
8	b_{10}	a_1b_{10}	b_3	718.8	1189	354.5	2128	134.3	2263
9	b_{10}	a_1b_{10}	a_3b_{10}	719.4	1189	357.8	2143	122.9	2266
10	b_{10}	a_1b_{10}	$a_3b_{10} + b_3$	718.8	1189	355.3	2128	134.7	2263
	random intercepts and random slopes								
11	$b_{10} + b_{11}t$	0	0	241.8	1715	356.2	2113	200.4	2313
12	$b_{10} + b_{11}t$	0	b_3	241.9	1715	354.2	2099	212.5	2311
13	$b_{10} + b_{11}t$	0	a_3b_{10}	241.7	1715	355.2	2111	200.8	2312
14	$b_{10} + b_{11}t$	0	a_4b_{11}	241.9	1715	350.4	2104	203.1	2307
15	$b_{10} + b_{11}t$	0	$a_3b_{10} + b_3$	241.9	1715	353.0	2098	212.2	2310
16	$b_{10} + b_{11}t$	0	$a_4b_{11} + b_3$	241.0	1715	351.5	2095	212.1	2307
17	$b_{10} + b_{11}t$	0	$a_3b_{10} + a_4b_{11}$	242.3	1715	351.1	2106	202.1	2308
18	$b_{10} + b_{11}t$	0	$a_3(b_{10} + b_{11})$	238.9	1715	352.7	2106	200.7	2307
19	$b_{10} + b_{11}t$	0	$a_3b_{10} + a_4b_{11} + b_3$	238.9	1715	349.7	2092	211.4	2304
20	$b_{10} + b_{11}t$	a_1b_{10}	0	374.3	1227	355.9	1742	215.1	1957
21	$b_{10} + b_{11}t$	a_1b_{10}	b_3	373.7	1227	354.6	1728	227.3	1955
22	$b_{10} + b_{11}t$	a_1b_{10}	a_3b_{10}	374.4	1226	357.5	1743	215.8	1958
23	$b_{10} + b_{11}t$	a_1b_{10}	a_4b_{11}	372.6	1226	346.6	1727	217.7	1945

Model	$U_1(t)$	$U_2(t)$	U_3	DIC_{y1}	DIC_{y2}	DIC_T	$\bar{D}$	p_D	DIC_{total}
24	$b_{10} + b_{11}t$	a_1b_{10}	$a_3b_{10} + b_3$	375.0	1228	355.1	1729	228.6	1958
25	$b_{10} + b_{11}t$	a_1b_{10}	$a_4b_{11} + b_3$	372.6	1227	348.0	1719	228.6	1947
26	$b_{10} + b_{11}t$	a_1b_{10}	$a_3b_{10} + a_4b_{11}$	372.8	1225	347.0	1726	218.9	1945
27	$b_{10} + b_{11}t$	a_1b_{10}	$a_3(b_{10} + b_{11})$	373.8	1227	356.2	1741	216.2	1957
28	$b_{10} + b_{11}t$	a_1b_{10}	$a_3b_{10} + a_4b_{11} + b_3$	373.6	1226	348.9	1719	229.4	1949
29	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	0	167.2	923.2	356.1	1235	211.6	1447
30	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	b_3	167.0	923.2	354.4	1221	223.9	1445
31	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	a_3b_{10}	168.6	922.4	356.3	1235	211.9	1447
32	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	a_4b_{11}	164.1	923.0	352.9	1229	210.6	1440
33	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	$a_3b_{10} + b_3$	166.0	923.1	353.4	1218	224.0	1442
34	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	$a_4b_{11} + b_3$	167.2	924.0	351.8	1219	224.3	1443
35	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	$a_3b_{10} + a_4b_{11}$	166.2	923.8	352.8	1230	212.7	1443
36	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	$a_3(b_{10} + b_{11})$	165.3	922.5	354.8	1232	211.0	1443
37	$b_{10} + b_{11}t$	$a_1b_{10} + a_2b_{11}t$	$a_3b_{10} + a_4b_{11} + b_3$	167.6	924.3	351.0	1218	224.5	1443

Table 4

Posterior estimates for Model 32

Parameter	Posterior Mean	Std. Error	95% CI
Longitudinal continuous submodel			
Intercept (β_{11})	0.167	0.048	(0.072, 0.263)
BaselineFVC (β_{12})	0.933	0.014	(0.906, 0.961)
Time (β_{13})	−0.094	0.016	(−0.122, −0.061)
$\sigma^2_{b_{10}}$	0.081	0.008	(0.066, 0.098)
$\sigma^2_{b_{11}}$	0.068	0.007	(0.056, 0.084)
ρ_{b_1}	−0.068	0.069	(−0.200, 0.067)
σ^2_{ε}	0.060	0.002	(0.055, 0.064)
Longitudinal binary submodel			
Intercept (β_{21})	−1.424	0.325	(−2.043, −0.796)
Age (β_{22})	−0.015	0.014	(−0.044, 0.012)
Time (β_{23})	1.019	0.180	(0.6513, 1.341)
α_1	−10.67	0.835	(−12.38, −9.145)
α_2	−10.31	0.727	(−11.82, −8.961)
Time-to-event submodel			
Intercept (β_{31})	−1.967	0.530	(−3.003, −0.939)
Male (β_{32})	0.903	0.360	(0.214, 1.648)
Smoking (β_{33})	0.719	0.341	(0.087, 1.422)
BaselineFVC (β_{34})	−0.405	0.203	(−0.807, −0.013)
α_4	−1.613	0.810	(−3.201, −0.025)