



Published in final edited form as:

Stat Biosci. 2019 July ; 11(2): 449–464. doi:10.1007/s12561-019-09252-4.

A Joint Model Approach for Longitudinal Data with no Time-Zero and Time-To-Event with Competing Risks

Sungduk Kim[†], Olive D. Buhule[‡], Paul S. Albert^{†,*}

[†]Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland, U.S.A.

[‡]Department of Biostatistics and Medical Informatics School of Medicine and Public Health, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.

Summary

The joint modeling of longitudinal and time-to-event data is an active area in biostatistics research. The focus of this article is on developing a modeling framework for these joint models when the longitudinal and time-to-event data do not have a meaningful time-zero. The motivating example is the study of a longitudinal assessment of station during child labor and its relationship to time-to-delivery. A good predictor of delivery type and timing would help obstetricians better manage the end of pregnancy and better facilitate delivery. One measure of labor progression is station, a measure of the position of the fetus' head in relation to the pelvis of the pregnant women, may be a good marker for delivery time and type. However, women enter the hospital, where their station is closely monitored, at arbitrary points in their labor process, resulting in no clear time zero. In addition, since delivery may be of various types, the competing risks due to type need to be accounted for. We develop a joint model of longitudinal and time-to-event data for this situation. The model is formulated through shared random effects between the survival and longitudinal processes, and parameter estimation is conducted through a Bayesian approach. The model is illustrated with longitudinal data on station where the relationship between station and event-time is studied and the model is used to assess the ability of longitudinal measures of station to predict the type and timing of pregnancy. We illustrate the methodology with longitudinal data taken during labor.

Keywords

Competing risk; longitudinal data; prediction; random effects; station; time-to-event

1. Introduction

Understanding the relationship between longitudinal biomarkers and time-to-event has been an important area for methodological research (Rizopoulos (2012), for example). These methods have been used for both characterizing the relationship between features of the longitudinal biomarker and the event-time distribution as well as to develop dynamic

*Paul Albert is the corresponding author.
Sungduk Kim and Olive Buhule are co-first authors.

predictors of clinical outcome from the repeated biomarker. This article proposes new statistical methodology for the situation where the longitudinal and survival processes have no clearly defined time zero and where there are competing risks for the time-to-event outcome. We present a shared random parameter model where the longitudinal process is linked to the event process through a series of random effects. Competing risks are incorporated by separately modeling the type and timing of the clinical event given these shared random effects.

The application of interest is on understanding the relationship between a longitudinal assessment of labor progression and the type and timing of delivery in women at the end of pregnancy. Specifically, researchers are interested in assessing the relationship between station, defined as the position of the fetuses in their mother's pelvis, and eventual delivery type and timing. Studying this relationship would provide useful information to the obstetrician about monitoring pregnant women for station during labor. The station in labor describes the descent of the fetus into the pelvis and is often measured on a -5 to $+5$ scale in centimeters, where the reference point of zero refers to the point where the head crosses the midpelvis (halfway out). Since women start getting measured for station once they come to the hospital in labor, there is no clear time zero relative to their start of followup. Instead, using the point where the station value crosses zero is a natural reference point. There are different ways that a fetus can be delivered, and labor progression may influence each delivery type differently. To account for this type of competing risk, we model delivery type using a polychotomous logistic regression model and then the time-to-delivery given the delivery type. The three delivery types are spontaneous which requires no intervention, C-section, and those requiring vacuum assistance.

Figure 1 shows a series of individual profiles for women who eventually have a spontaneous delivery. The figure demonstrates that the monitoring of station begins at different points in labor. Further, the longitudinal profile appears different before and after a station value of zero.

McLain and Albert (2014) proposed a model for longitudinal data with a random change point and no-time zero that was motivated by repeat labor dilation data. The methodology was used to characterize the relationship between key demographic factors and labor progression as assessed by repeated measures of a woman's cervical dilation. Further, this modeling strategy was used to predict future dilation measurements as well as the time to full dilation for a pregnant women based on prior dilation measurements. The current article asks a different, and perhaps the more important question of whether a repeated measure of labor progression affects the timing and type of pregnancy outcome.

In this article, we propose a joint model for longitudinal assessments of station with a random landmark and time to delivery and mode of delivery as a competing risk. We consider a shared random parameter model that links together model components for the longitudinal station process (with no time-zero) and both the time and type of delivery. Specifically, each model component includes random effects that are shared between these three components, inducing realistic dependence between these three model components. To this end, we develop an efficient Bayesian computational method for fitting this joint model

via several modified collapsed Gibbs samplers (Chen et al. 2000; Liu 1994). In addition, we derive the deviance information criterion (DIC) for comparing several variations of the proposed joint models, which is a modified DIC measure when mixtures of distributions or latent variables are present.

The rest of this paper is organized as follows. In Section 2, we provide the methodological development of a flexible class of joint models for the longitudinal station measurements and the time to delivery as well as the model of delivery. This includes specifying a skewed generalized t -distribution for the time to delivery. In Section 3, we discuss the prediction of both the timing and type of delivery from repeated station measurements taken at various time points that can be irregularly spaced across individuals. The prior and posterior distributions, the Markov chain Monte Carlo (MCMC) sampling algorithm, and the Bayesian model comparison criteria are presented in Sections 4.1, 4.2, and 4.3, respectively. In Section 5, we apply the methodology to longitudinal labor data in order to study the association between labor dynamics and delivery time. Finally, we conclude the article with some discussion for future research in Section 6.

2. Joint Models

Suppose that i denotes individual ($i = 1, \dots, I$) and j denotes time point ($j = 1, \dots, J_i$). We assume that each measurement is taken at repeated time points, which are potentially irregularly spaced times in the delivery. Further, we assume that there are I individuals in the study, each contributing J_i time points, where J_i denotes the number of repeated time points on the i th individual. Let $\mathbf{u}_i = (u_{i1}, \dots, u_{i,p_1})'$ and $\mathbf{z}_i = (z_{i1}, \dots, z_{i,p_2})'$ denote the vectors of fixed effect covariates, where \mathbf{z}_i may share common components with \mathbf{u}_i . Also let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_{p_1})'$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_{p_2})'$ denote the corresponding regression vectors for model components of the time to delivery and the mode of delivery, respectively. In addition, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_{p_3})'$ denotes the vector of regression parameters for the longitudinal measurements. Furthermore, we assume that t_{ij}^* denotes the j th time for the i th individual after arriving at the hospital, where $t_{i1}^* = 0$ (arriving time). Note that, since the application has three types of delivery (spontaneous, vacuum, and C-section), for ease of notation, the model is specified with three competing risks.

Let y_{ij} denote the j th longitudinal station measurement, D_i denotes the mode of delivery such as spontaneous, vacuum, or C-section, and S_i^* denotes the time to delivery from entry into the hospital for the i th individual, separately. Note that each individual enter the hospital with different station measurement so that there is no known time zero when labor starts. Furthermore, station measurements were taken on each woman from time of entry into the hospital to a few hours before giving birth, resulting in no clear time zero for use as a reference point for valid statistical inferences and predictions. Therefore, we re-scale the t_{ij}^* and S_i^* such that time zero is the time when a woman's station is zero (i.e. when a fetus is said to be engaged in the pelvis). We consider the joint models with shared random effects for the longitudinal measurements, time to delivery, and delivery type. First, we propose the following random effects regression model for the longitudinal measurements as

$$y_{ij} = \begin{cases} K + \beta_1 t_{ij} + \beta_2 t_{ij}^2 + b_{1i} t_{ij} + b_{2i} t_{ij}^2 + \epsilon_{ij}^y & \text{if } t_{ij} \geq 0 \\ K + \beta_3 t_{ij} + b_{3i} t_{ij} + \epsilon_{ij}^y & \text{if } t_{ij} < 0, \end{cases} \quad (2.1)$$

where K is the mean station when $t_{ij} = 0$, $t_{ij} = t_{ij}^* - \mu - b_{4i}$ is the re-scaled time for the j^{th} station measurement on the i^{th} woman with the mean time re-scaling factor μ and random effect b_{4i} allowing for heterogeneity, and ϵ_{ij}^y are random variables for the error distributions. A station value of zero, the point at which the head of the fetus crosses the mid-pelvis, is a natural reference choice; thus, $K = 0$. Furthermore, b_{1i} , b_{2i} , and b_{3i} are individual-level random effects, characterizing heterogeneity in the station process across women. Let $\epsilon_i^y = (\epsilon_{i1}^y, \dots, \epsilon_{iJ_i}^y)'$ and $\mathbf{b}_i = (b_{1i}, b_{2i}, b_{3i}, b_{4i})'$. In (2.1), we assume that (i) ϵ_i^y and \mathbf{b}_i are independent; (ii) ϵ_{ij}^y follows a normal distribution with mean 0 and variance σ_y^2 ; and (iii) the shared random effect \mathbf{b}_i follows a multivariate normal distribution mean 0 and 4×4 unstructured variance-covariance matrix Ω for $i = 1, \dots, I$ and $j = 1, \dots, J_i$. Furthermore, we assume quadratic trend after station zero and a linear form before and set $K = 0$. Let $\mathbf{y} = (y_{11}, \dots, y_{I,J_i})'$, $\mathbf{t}^* = (t_{11}^*, \dots, t_{I,J_i}^*)'$, $\mathbf{t} = (t_{11}, \dots, t_{I,J_i})'$, and $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_I)'$. Also let $D_{obs}^y = (\mathbf{y}, \mathbf{t}^*)$ denote the observed data. Given \mathbf{b} and D_{obs}^y , the likelihood function of (μ, β, σ_y^2) for longitudinal station measurements is given by

$$L_y(\mu, \beta, \sigma_y^2 | \mathbf{b}, D_{obs}^y) = \prod_{i=1}^I \prod_{j=1}^{J_i} N(y_{ij} | \mu_{y_{ij}}, \sigma_y^2), \quad (2.2)$$

where $\mu_{y_{ij}} = \beta_1 t_{ij} + \beta_2 t_{ij}^2 + b_{1i} t_{ij} + b_{2i} t_{ij}^2$ for $t_{ij} \geq 0$ and $\mu_{y_{ij}} = \beta_3 t_{ij} + b_{3i} t_{ij}$ for $t_{ij} < 0$. Note that $N(\cdot | a, b)$ denotes a normal distribution with mean a and variance b .

For time to delivery, we propose the following flexible regression model as

$$S_i = \mathbf{u}_i' \boldsymbol{\alpha} + b_{1i} \theta_1 + b_{2i} \theta_2 + b_{3i} \theta_3 + D_{i1} \theta_4 + D_{i2} \theta_5 + \epsilon_i^S, \quad (2.3)$$

where $S_i = S_i^* - \mu - b_{4i}$ is the re-scaled time to delivery for the i^{th} woman, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_5)$ are shared parameters accounting for correlation between \mathbf{y} , S_i , and $D_i = (D_{i1}, D_{i2})'$ which is dichotomized. We assume that ϵ_i^S and \mathbf{b}_i are independent, and ϵ_i^S are random variables for the flexible error distributions that will allow for long-tailed and skewed distributions. In this paper, we propose to use long-tailed and skewed distributions to accomplish this goal. We model

$$\epsilon_i^S = \delta(\xi_i - E(\xi_i)) + \eta_i, \quad (2.4)$$

where the first and second terms in (2.4) characterize skewness and long tails, respectively. Similar to Kim et al. (2008), we assume that (i) ξ_i and η_i are independent; (ii) $\xi_i \sim G_\xi$ are each independent, where G_ξ is the cumulative density function (cdf) of a skewed distribution defined on $R^+ = (0, \infty)$; and (iii) η_i follows a symmetric distribution. Also in (2.4), δ is a skewness parameter, where $\delta = 0$ reflects a symmetric distribution. Following Chen et al. (1999) and Kim et al. (2008), we assume that G_ξ has a known cdfs in order to ensure model identifiability. In this paper, we first specify several different distributions for G_ξ , and use

the DIC (Spiegelhalter et al. (2002)) to determine which G_{ξ} fits the data best. We consider the following distributions for G_{ξ} : (a) G_{ξ} is degenerated at 0, denoted by $\Delta_{\{0\}}$, yielding a symmetric distribution; (b) G_{ξ} is a standard exponential distribution (\mathcal{E}) with probability density function (pdf) $f_{\mathcal{E}}(\xi_i) = \exp(-\xi_i)$ if $\xi_i > 0$ and 0 otherwise; and (c) G_{ξ} is a half normal (\mathcal{HN}) with pdf $f_{\mathcal{HN}}(\xi_i) = \frac{2}{\sqrt{2\pi}} \exp(-\xi_i^2/2)$ if $\xi_i > 0$ and 0 otherwise. Thus given G_{ξ} , the model (2.4) yields a skewed distribution. Furthermore, we propose a skewed generalized t model for S_i , which is motivated by Kim et al. (2008) for a binary response data in a generalized linear model. Let $f_{gt, \nu_1, \nu_2}(\eta_i)$ denote the pdf of a generalized t -distribution (Abranowitz and Stegun (1972)) that is given by

$$f_{gt, \nu_1, \nu_2}(\eta_i) = \frac{1}{\sqrt{\pi\nu_2}} \frac{\Gamma\left(\frac{\nu_1+1}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)} \times \left(1 + \frac{\eta_i^2}{\nu_2}\right)^{-\frac{\nu_1+1}{2}}, \tag{2.5}$$

where ν_1 is a shape (or degrees of freedom) and ν_2 is a scale parameter. Let $\eta_i \sim f_{gt, \nu_1, \nu_2}$, then $E(\eta_i) = 0$ for $\nu_1 > 1$ and $\text{Var}(\eta_i) = \nu_2/(\nu_1 - 2)$ for $\nu_1 > 2$. To ensure identifiability, we assume ν_2 to be 1 (Kim et al. 2008). We note that based on (2.5), e_i^S in (2.4) leads to a skewed generalized t -distribution for S_i . The models defined in (2.3), (2.4), and (2.5) are general and flexible, which includes the normal and skewed t models as special cases. Figure 2 illustrates the skewed generalized t -distribution for positive, negative, and no skewness.

Let $\mathbf{S} = (S_1, \dots, S_I)'$, $\mathbf{S}^* = (S_1^*, \dots, S_I^*)'$, $\mathbf{U} = (\mathbf{u}'_1, \dots, \mathbf{u}'_I)'$, $\mathbf{D}_1 = (D_{11}, \dots, D_{I1})'$, and $\mathbf{D}_2 = (D_{12}, \dots, D_{I2})'$. Also let $\mathbf{D}_{obs}^D = (\mathbf{S}^*, \mathbf{U}, \mathbf{D}_1, \mathbf{D}_2)$ denote the observed data. Given \mathbf{b} and \mathbf{D}_{obs}^S , the likelihood function of $(\mu, \alpha, \theta, \delta, \nu_1, \sigma_S^2)$ for time to delivery is given by

$$L_S(\mu, \alpha, \theta, \delta, \nu_1, \sigma_S^2 | \mathbf{b}, \mathbf{D}_{obs}^S) = \prod_{i=1}^I \int f_{gt, \nu_1, \nu_2=1}(S_i | \mu_{S_i}, \sigma_S^2) dG_{\xi}(\xi_i), \tag{2.6}$$

where $\mu_{S_i} = \mathbf{u}'_i \alpha + b_{1i}\theta_1 + b_{2i}\theta_2 + b_{3i}\theta_3 + D_{i1}\theta_4 + D_{i2}\theta_5 + \delta(\xi_i - E(\xi_i))$. Note that $f_{gt, \nu_1, \nu_2=1}(\cdot | a, b)$ denotes a pdf of generalized t -distribution with mean a , variance b , ν_1 and $\nu_2 = 1$.

For the mode of delivery, we assume the multinomial logistic regression model as

$$P(D_i = l | \mathbf{z}, \mathbf{b}) = \frac{\exp(\mathbf{z}'_i \gamma_l + b_{1i}\phi_{l1} + b_{2i}\phi_{l2} + b_{3i}\phi_{l3})}{1 + \sum_{m=1}^{L-1} \exp(\mathbf{z}'_i \gamma_m + b_{1i}\phi_{m1} + b_{2i}\phi_{m2} + b_{3i}\phi_{m3})}, \tag{2.7}$$

where $\phi_l = (\phi_{l1}, \phi_{l2}, \phi_{l3})'$ are shared parameters accounting for correlation between \mathbf{y} , S_i , and $D_i, l = 1, \dots, L - 1$. Note that $\phi_L = 0$. Let $\mathbf{D} = (D_1, \dots, D_I)'$, $\mathbf{Z} = (\mathbf{z}'_1, \dots, \mathbf{z}'_I)'$, $\boldsymbol{\gamma} = (\gamma'_1, \dots, \gamma'_{L-1})'$, and $\boldsymbol{\phi} = (\phi'_1, \dots, \phi'_{L-1})'$. Also let $\mathbf{D}_{obs}^D = (\mathbf{D}, \mathbf{Z})$ denote observed data. Given \mathbf{b} and \mathbf{D}_{obs}^D , the likelihood function of $(\boldsymbol{\gamma}, \boldsymbol{\phi})$ for time to delivery is given by

$$L_D(\boldsymbol{\gamma}, \boldsymbol{\phi} | \mathbf{b}, \mathbf{D}_{obs}^D) = \prod_{i=1}^I \prod_{l=1}^{L-1} P(D_i = l | \mathbf{z}, \mathbf{b})^{1(D_i=l)}, \tag{2.8}$$

where $P(D_i = l | \mathbf{z}, \mathbf{b})$ is given in (2.7) and $1(a)$ is an indicator function, which is defined as $1(a) = 1$ if a is true and 0 otherwise.

The observed joint likelihood function of $(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi)$ is

$$\begin{aligned} & L(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, \Omega | D_{obs}^y, D_{obs}^S, D_{obs}^D) \\ &= \int L_y(\mu, \beta, \sigma_y^2 | \mathbf{b}, D_{obs}^y) \times L_S(\mu, \alpha, \theta, \delta, \nu_1, \sigma_S^2 | \mathbf{b}, D_{obs}^S) \times L_D(\gamma, \phi | \mathbf{b}, D_{obs}^D) \\ & \times \prod_{i=1} N(\mathbf{b}_i | 0, \Omega) d\mathbf{b}, \end{aligned} \quad (2.9)$$

where $L_y(\mu, \beta, \sigma_y^2 | \mathbf{b}, D_{obs}^y)$, $L_S(\mu, \alpha, \theta, \delta, \nu_1, \sigma_S^2 | \mathbf{b}, D_{obs}^S)$, and $L_D(\gamma, \phi | \mathbf{b}, D_{obs}^D)$ are given in (2.2), (2.6), and (2.8), respectively. Since it is not easy to work directly with the observed joint likelihood function of $(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi)$ in (2.9), we develop an efficient MCMC algorithm using the fact that the generalized t -distribution can be represented as a gamma mixture of normal distributions for ϵ_i^S as described in Section 4.2.

3. Predicting the time to delivery and delivery type

The joint model in (2.1), (2.3), and (2.7) relates the longitudinal station patterns to the probability of a delivery type and time to delivery through an individual's predicted station measurement, and can be used to develop a predictor of the timing and delivery type using individualized longitudinal assessments of station. To predict the time to delivery and mode of delivery at birth using the longitudinal station measurements, we let $\mathbf{y}^p = (y_{t_1}, \dots, y_{t_M})$ denote the longitudinal station measurement taken at time points t_1, \dots, t_M , where M is the number of repeated measurements in the predictor. Furthermore, let S^p and D^p denote the time to delivery and mode of delivery we wish to predict. Also let $\Theta = (\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi)$. Then the posterior predictive probability for D^p and posterior predictive distribution for S^p based on longitudinal measurements \mathbf{y}^p can be given by

$$\begin{aligned} P(D^p = l | \mathbf{y}^p) &= \int P(D^p = l | \mathbf{z}, \mathbf{b}^p) \pi(\Theta, \mathbf{b}^p | \mathbf{y}^p) d\mathbf{b}^p d\Theta, \\ \pi(S^p | \mathbf{y}^p) &= \int \int_{g^p, \nu_1, \nu_2 = 1} (S^p | \mu_{S^p}, \sigma_S^2) \pi(\Theta, \mathbf{b}^p | \mathbf{y}^p) d\mathbf{b}^p d\Theta, \end{aligned} \quad (3.1)$$

where $\mu_{S^p} = \mathbf{u}'\alpha + b_1^p\theta_1 + b_2^p\theta_2 + b_3^p\theta_3 + D_1^p\theta_4 + D_2^p\theta_5 + \delta(\xi - E(\xi))$, \mathbf{b}^p is a multivariate random effects ($\mathbf{b}^p \sim N(0, \Omega)$) and $\pi(\Theta, \mathbf{b}^p | \mathbf{y}^p)$ is the posterior distribution for Θ and \mathbf{b}^p based on \mathbf{y}^p . To obtain the posterior predictive probability for D^p and posterior predictive distribution for S^p in (3.1), we sample \mathbf{b}^p from joint posterior distribution based on test set data (with parameter estimates using training dataset). Specifically, to evaluate the prediction accuracy, we divide the data into a training and test set data. Note that mode of delivery is treated as a competing risk for predicting S^p .

4. Posterior inference

4.1. Prior and posterior distributions

We assume that $\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi,$ and Ω are independent *a priori*. The priors that we assume have as weak information as possible so that the data can dominate the posterior computation. Specifically, we assume that, $\mu \sim N(0, \sigma_\mu^2), \alpha \sim N_{p_1}(0, \sigma_\alpha^2 I), \gamma \sim N_{p_2}(0, \sigma_\gamma^2 I), \beta \sim N_{p_3}(0, \sigma_\beta^2 I), \theta \sim N_5(0, \sigma_\theta^2 I), \phi \sim N_3(0, \sigma_\phi^2 I), \delta \sim N(0, \sigma_\delta^2), \nu_1,$ and $\sim \text{Gamma}(a_1, b_1), \sigma_y^2 \sim \text{IG}(a_2, b_2), \sigma_s^2 \sim \text{IG}(a_3, b_3)$

$\Omega^{-1} \sim \text{Wishart}(d_0, d_0 V_0)$, where $\sigma_\mu^2, \sigma_\alpha^2, \sigma_\gamma^2, \sigma_\beta^2, \sigma_\theta^2, \sigma_\phi^2, \sigma_\delta^2, a_1, b_1, a_2, b_2, a_3, b_3, d_0,$ and $V_0,$ are the prespecified hyperparameters. Note that $\text{Gamma}(a, b)$ is a Gamma distribution with mean a/b , $\text{IG}(a, b)$ is an Inverse-gamma distribution with mean $b/(a - 1)$, and $\text{Wishart}(d, V)$ denotes a Wishart prior distribution with d degrees of freedom and mean dV . The hyperparameters of the prior were specified as $\sigma_\mu^2 = 100, \sigma_\alpha^2 = 100, \sigma_\gamma^2 = 100, \sigma_\beta^2 = 100, \sigma_\theta^2 = 100, \sigma_\phi^2 = 100, \sigma_\delta^2 = 100, a_1 = 1, b_1 = 0.1, a_2 = 2, b_2 = 1, a_3 = 2, b_3 = 1, d_0 = 4.1$

and $V_0 = 0.1I$ in the analysis. These choices of hyperparameters lead to noninformative priors. Let $D_{obs} = (D_{obs}^y, D_{obs}^S, D_{obs}^D)$. Based on the prior distributions specified above, the joint posterior distribution of $\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi,$ and Ω is thus given by

$$\begin{aligned} & \pi(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi, \Omega \mid D_{obs}) \\ & = L(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi, \Omega \mid D_{obs}) \\ & \times \pi(\mu)\pi(\beta)\pi(\sigma_y^2)\pi(\alpha)\pi(\theta)\pi(\delta)\pi(\nu_1)\pi(\sigma_s^2)\pi(\gamma)\pi(\phi)\pi(\Omega), \end{aligned} \tag{4.1}$$

where $L(\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi, \Omega \mid D_{obs})$ is defined in (2.9). We can generate a sample from this joint posterior distribution using Gibbs sampler and make appropriate inference of the various model parameters. A description of the MCMC algorithm is given in Section 4.2.

4.2. Computational Developments

The analytical evaluation of the posterior distribution of $\Theta = (\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_s^2, \gamma, \phi, \Omega)$ given in (4.1) is not available. However, we can develop an efficient MCMC sampling algorithm to sample from (4.1). For ease of presentation, we consider the standard exponential distribution (\mathcal{E}) for G_ξ only as the MCMC sampling algorithms for other choices of G_ξ are similar. Since it is difficult to work directly with the observed joint likelihood function of Θ in (2.9), we use the complete data joint likelihood function of Θ using the fact that the generalized t -distribution can be represented as a gamma mixture of normal distributions for $\epsilon_i^S: \eta_i \mid \lambda_i \sim N(0, 1/\lambda_i)$ and $\lambda_i \sim \text{Gamma}(\nu_i/2, 1/2)$, where λ_i is a mixing variable. Let $\xi = (\xi_1, \dots, \xi_t)'$ and $\lambda = (\lambda_1, \dots, \lambda_t)'$. Also let $D_c = (t, S, b, \xi, \lambda, D_{obs})$ denote complete data set. The joint posterior distribution of Θ based on the complete data joint likelihood function can be written as

$$\begin{aligned}
 &\pi(\Theta | D_i) \\
 &\propto \prod_{i=1}^I \prod_{j=1}^L (\sigma_y^2)^{-\frac{1}{2}} \exp \left(-\frac{1}{2\sigma_y^2} (y_{ij} - \mu_{y_{ij}})^2 \right) \\
 &\times \prod_{i=1}^I \left(\frac{\sigma_S^2}{\lambda_i} \right)^{-\frac{\nu_1}{2}} \exp \left(-\frac{\lambda_i}{2\sigma_S^2} (S_i - \mu_{S_i})^2 \right) \times \exp \left(-\xi_i \right) \times \frac{2^{-\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)} \lambda_i^{\frac{\nu_1}{2}-1} \exp \left(-\frac{1}{2} \lambda_i \right) \\
 &\times \prod_{i=1}^I \frac{\exp \left(\sum_{l=1}^{L-1} 1(D_i = l) (\mathbf{z}_i' \gamma_l + b_{1i} \phi_{l1} + b_{2i} \phi_{l2} + b_{3i} \phi_{l3}) \right)}{1 + \sum_{m=1}^{L-1} \exp \left(\mathbf{z}_i' \gamma_m + b_{1i} \phi_{m1} + b_{2i} \phi_{m2} + b_{3i} \phi_{m3} \right)} \\
 &\times \prod_{i=1}^I \left| \Omega \right|^{-\frac{1}{2}} \exp \left(-\frac{1}{2} \mathbf{b}_i' \Omega^{-1/2} \mathbf{b}_i \right) \\
 &\times \pi(\Theta),
 \end{aligned} \tag{4.2}$$

where $\mu_{y_{ij}}$, μ_{S_i} , and $\pi(\Theta)$ are given in (2.2), (2.6), and (4.1), respectively. From (4.2), the algorithm requires sampling the following parameters in turn from their respective full conditional distributions: (I) $[\xi_i | \mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, t, S, \mathbf{b}, \lambda, \Omega, D_{obs}]$; (II) $[\Omega | \mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, t, S, \mathbf{b}, \xi, \lambda, D_{obs}]$; (III) $[\gamma, \phi | \mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, t, S, \mathbf{b}, \xi, \lambda, \Omega, D_{obs}]$; and (IV) $[\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \mathbf{b}, \lambda | \gamma, \phi, t, S, \xi, \Omega, D_{obs}]$. Let $\mathbf{x}_i = (\mathbf{u}_i', b_{1i}, b_{2i}, b_{3i}, D_{i1}, D_{i2})'$, $X = (\mathbf{x}_1', \dots, \mathbf{x}_I')$, $\mathbf{w}_i = (\mathbf{z}_i', b_{1i}, b_{2i}, b_{3i})'$, $W = (\mathbf{w}_1, \dots, \mathbf{w}_I)'$, $\boldsymbol{\psi} = (\boldsymbol{\alpha}', \boldsymbol{\theta}')'$, $\boldsymbol{\varphi}_i = (\gamma_i', \boldsymbol{\phi}_i)'$ and $\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_{L-1})'$. For (I), observed that

$$\xi_i | \mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, t, S, \mathbf{b}, \lambda, \Omega, D_{obs} \sim N(A_{\xi_i}^{-1} B_{\xi_i}, A_{\xi_i}^{-1}) 1(\xi_i > 0), \tag{4.3}$$

where $B_{\xi_i} = \lambda_i(S_i - \mathbf{x}_i' \boldsymbol{\psi} + \delta E(\xi_i)) \delta / \sigma_S^2 - 1$ and $A_{\xi_i} = \lambda_i \delta^2 / \sigma_S^2$ for $i = 1, \dots, I$. For (II), we observe that

$$\begin{aligned}
 &\Omega^{-1} | \mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, t, S, \mathbf{b}, \xi, \lambda, D_{obs} \\
 &\sim \text{Wishart}_4 \left(d_0 + I, \left(d_0 V_0^{-1} + \sum_i \mathbf{b}_i \mathbf{b}_i' \right)^{-1} \right).
 \end{aligned} \tag{4.4}$$

For (III), the conditional posterior densities for $\boldsymbol{\varphi}$ do not have closed form. Therefore, we use the Metropolis-Hastings algorithm (Hastings (1970)) to sample $\boldsymbol{\varphi}$ from their conditional posterior distributions. For (IV), we apply the collapsed Gibbs technique of Liu (1994) via the following identity:

$$\begin{aligned}
 &[\mu, \beta, \sigma_y^2, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \mathbf{b}, \lambda | \gamma, \phi, t, S, \xi, \Omega, D_{obs}] \\
 &= [\sigma_y^2, \lambda | \mu, \beta, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \gamma, \phi, t, S, \mathbf{b}, \xi, \Omega, D_{obs}] \\
 &\times [\mu, \beta, \alpha, \theta, \delta, \nu_1, \sigma_S^2, \mathbf{b} | \gamma, \phi, t, S, \xi, \Omega, D_{obs}].
 \end{aligned}$$

That is, we sample $\mu, \beta, \alpha, \theta, \delta, \nu_1, \sigma_S^2$, and \mathbf{b} after collapsing out σ_y^2 and λ . Observe that

$$\begin{aligned} &\sigma_y^2 \mid \mu, \boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \delta, \nu_1, \sigma_s^2, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{t}, \boldsymbol{S}, \boldsymbol{b}, \boldsymbol{\xi}, \boldsymbol{\lambda}, \boldsymbol{\Omega}, D_{obs} \\ &\sim \text{IG} \left(a_2 + \frac{1}{2} \sum_i J_i, b_2 + \frac{1}{2} \sum_i \sum_j (y_{ij} - \mu_{y_{ij}})^2 \right), \end{aligned} \tag{4.5}$$

where $\mu_{y_{ij}}$ is given in (2.2). We also observe that

$$\begin{aligned} &\lambda_i \mid \mu, \boldsymbol{\beta}, \sigma_y^2, \boldsymbol{\alpha}, \boldsymbol{\theta}, \delta, \nu_1, \sigma_s^2, \boldsymbol{\gamma}, \boldsymbol{\phi}, \boldsymbol{t}, \boldsymbol{S}, \boldsymbol{b}, \boldsymbol{\xi}, \boldsymbol{\Omega}, D_{obs} \\ &\sim \text{Gamma} \left(\frac{1}{2}(\nu_1 + 1), \frac{1}{2} \left(1 + \frac{(S_i - \mu_{S_i})^2}{\sigma_s^2} \right) \right), \end{aligned} \tag{4.6}$$

where S_i and μ_{S_i} are given in (2.3). The conditional posterior densities for $\mu, \boldsymbol{\beta}, \boldsymbol{\psi}, \delta, \nu_1, \sigma_s^2$, and \boldsymbol{b} do not have closed form. Therefore, we use the Metropolis-Hastings algorithm (Hastings (1970)) to sample $\mu, \boldsymbol{\beta}, \boldsymbol{\psi}, \delta, \nu_1, \sigma_s^2$, and \boldsymbol{b} from their conditional posterior distributions.

4.3. Model Comparison

To assess the goodness of fit of the models for different choices of the distributions for ξ_i and η_i in (2.4), we use the DIC proposed by Celeux et al. (2006) without integrating out \boldsymbol{b} analytically in (2.1), (2.3), and (2.7). Although certain numerical integration or Monte Carlo methods may be used for evaluating those analytically intractable integrals, those methods are computationally intensive to carry out due to the large size of the data. Following the suggestions of Celeux et al. (2006), we use the DIC_4 , which is a modified DIC measure when mixtures of distributions or latent variables are present. The measure DIC_4 makes use of the complete-data likelihood in the presence of latent variables, which is given by

$$\begin{aligned} &L_c(\boldsymbol{\Theta} \mid \boldsymbol{b}, D_{obs}) \\ &= \prod_{i=1}^I \left[\prod_{j=1}^J f_n(y_{ij} \mid \mu_{y_{ij}}, \sigma_y^2) \right] \times \int f_{gt, \nu_1, \nu_2=1}(S_i \mid \mu_{S_i}, \sigma_s^2) dG_{\xi}(\xi_i) \\ &\times \prod_{i=1}^I \prod_{l=L}^L P(D_i = l \mid \boldsymbol{z}, \boldsymbol{b})^{1(D_i=l)} \times f_{mn}(\boldsymbol{b}_i \mid 0, \boldsymbol{\Omega}) \end{aligned} \tag{4.7}$$

where $f_{gt, \nu_1, \nu_2=1}(S_i \mid \mu_{S_i}, \sigma_s^2)$ and $P(D_i = l \mid \boldsymbol{z}, \boldsymbol{b})$ are defined in (2.5) and (2.7), $f_n(\cdot \mid a, b)$ and $f_{mn}(\cdot \mid 0, \boldsymbol{\Omega})$ denote a pdf of normal distribution with mean a and variance b , and multivariate normal distribution with mean 0 and variance-covariance matrix $\boldsymbol{\Omega}$, respectively. Then DIC_4 can be written as

$$DIC_4 = -4E_{\boldsymbol{\theta}, \boldsymbol{b}}[\log L_c(\boldsymbol{\Theta} \mid \boldsymbol{b}, D_{obs}) \mid D_{obs}] + 2E_{\boldsymbol{\theta}, \boldsymbol{b}}[\log L_c(E_{\boldsymbol{\theta}}(\boldsymbol{\Theta} \mid D_{obs}) \mid \boldsymbol{b}, D_{obs}) \mid D_{obs}], \tag{4.8}$$

where the integration for ξ_i is evaluated numerically. The smaller the DIC value, the better the model fits the data. The other properties of the DIC can be found in Spiegelhalter et al. (2002) and Celeux et al. (2006).

5. Analysis of the Labor Data

In this section, we used the proposed joint model in (2.1) - (2.7) to analyze data from the cohort study introduced in Section 1. The outcome variables were y_{ij}, S_i^* , and D_i , which

were defined as a station measurement at the j th time point, time to delivery, and the mode of delivery for the i th individual. $D_i = 1, 2,$ and 3 correspond to spontaneous, vacuum, and C-section, respectively. The time point t_{ij}^* is the j th time (hours) from time of entry into the hospital for the i th individual ($t_{i1}^* = 0$). We considered two covariates: Age (years) and the delivery type. Delivery type was dichotomized ($(D_1, D_2) = (0, 0)$ for spontaneous, $(1, 0)$ for vacuum, and $(0, 1)$ for C-section). In total, we had $I = 637$ and $2 \leq J_i \leq 15$, the range of the maximum number of repeated station measurements. In addition, 18.4%, 15.3%, and 66.3% of the women delivered through vacuum, C-section, and spontaneous, respectively. The medium delivery times since entry into the hospital were 6.27 hours for vacuum, 5.97 hours for C-section, and 4.6 hours for spontaneous, respectively. In addition, we divide the whole data into two data sets randomly with a 70% and 30% split of the sample into a training and test set, respectively (i.e., 446 in the training set and 191 in the test set). We used the training set data to develop the predictor, while test set data was used to validate the predictor with different accuracy measures. We fitted the proposed joint model in (2.1) - (2.7) and estimated all parameters using the training-set data.

To help the numerical stability and to improve convergence for the MCMC sampling algorithm, we standardized two covariates. The means and standard deviations were (28.7, 5.04) for age, (0.16, 0.37) for D_1 , and (0.15, 0.35) for D_2 , respectively. In all the Bayesian computations, we used 20,000 Gibbs samples, which were taken from every 5th iteration, after a burn-in of 6,000 iterations, to compute all the posterior estimates, including means (Estimates), SDs, 95% HPD intervals, as well as the DICs. The computer programs were written in FORTRAN 95 using IMSL subroutines with double precision accuracy. The convergence of the MCMC sampling algorithm for all the parameters was checked based on the recommendations of Cowles and Carlin (1996). All trace and autocorrelation plots showed good convergence and mixing of the MCMC sampling algorithm.

We are interested in investigating how the goodness of fit might be affected by the choices of G_{ξ} for joint model in (2.1) - (2.7) using the DIC discussed in Section 4.3. We considered the following distributions for η_i and G_{ξ} : (1) Normal and generalized t (GT) for η_i ; (2) $A_{(0)}$, \mathcal{E} , and $\mathcal{H}\mathcal{N}$ for G_{ξ} . Table 1 shows the DIC values for the six models under consideration, with the smallest value for DIC (3612.15) corresponding to skewed generalized t -distribution with the \mathcal{E} for G_{ξ} (SGTE), which fits training data the best among all models considered. This affirms the need for considering the skewed heavy tail distributions. In addition, the models with symmetric distributions for η_i had the larger DIC value, suggesting that these models fit data worse than the skewed models, demonstrating the importance of including skewness for the model of the time to delivery in (2.1).

The posterior estimates, including the posterior means, posterior standard deviations (SDs), and 95% highest posterior density (HPD) intervals of the parameters under the best model based on the training set are reported in Tables 2 and 3. We define a posterior estimate to be “statistically significant at a significance level of 0.05” if the corresponding 95% HPD interval does not contain 0. The results shown in Table 3 indicate that the fixed effects parameters β_1 , β_2 , and β_3 are all statistically significant with the slope before reaching a station value of zero being gradual as compared to the much larger increases after reaching

this reference time. The estimated mean scaling factor, $\hat{\mu}$, is -2.61 , suggesting that the typical fetus reaches a zero reference approximately 2.6 hours after entering the hospital in labor. Time to delivery (from a zero station) decreases with age (-0.176 hours per year) and this time for vacuum assisted and C-section deliveries is 0.152 longer and -0.296 shorter, respectively, relative to a spontaneous delivery. For a women at the average age of 28.7 years who has a spontaneous delivery, the median time to delivery relative to reaching a station value of zero is $\hat{\alpha}_0 = 7.8$ hours (this was computed based on age = 0 on the standardized scale and with the random effects and mode of delivery effects set to zero). The shared random effects b_{i1} , b_{i2} , and b_{i3} demonstrate that increases in the slopes before and after the zero station value are associated with a decreased time to delivery, with this association being substantially stronger for the slope after the zero value. Table 3 shows the relationship between longitudinal station measurements and the type of delivery. An increased subject-specific rate of change before and after the zero station value was associated with a decreased risk of a C-section compared with the reference category of a spontaneous delivery. There was no effect of the dynamics of labor progression on the risk of a vacuum assisted delivery as compared to this reference category.

We fitted the model with the training-set data and evaluated its performance on the test-set data. Figure 3 shows a plot of deciles of the predicted probabilities versus the corresponding sample proportions corresponding to those deciles for each delivery type. The plot shows that for spontaneous deliveries (the type that is not controlled by the obstetrician), the predictive model is well calibrated. Further the predictive probabilities range from a lower decile of 0.55 to an upper decile of 0.9, showing a large amount of potential discrimination. Figure 4 shows the predicted versus the observed delivery times for spontaneous deliveries. The figure demonstrates that for predicted values less than 2 hours, the model prediction is accurate (all subjects with predictive values less 2 hours had an observed delivery time of < 2 hours). However, the predicted delivery time is not as accurate when the predicted values are greater than 2 hours.

6. Discussion

This article proposed a joint model of longitudinal and time-to-event data when there is no natural time zero and when there are competing risks. For the labor example, station values are measured longitudinally and the time of study entry is not a natural reference point since expecting mothers come to the hospital at different points in their late pregnancy. The competing risk aspects need to be considered since delivery type is an important component of the endpoint and time-to-delivery may be influenced by the type of delivery. The model led to important scientific observations. First, the dynamics in labor (longitudinal station process) is associated with delivery time such that the time to delivery is shorter when the velocity of station is higher. This is particularly pronounced in the later stage of labor (after the station value crosses zero). The type of delivery is also associated with station velocity. Most notably, a C-section is much more likely when the velocity (before and after a station value of zero) is delayed relative to the typical profile. This later observation is consistent with clinical experience.

For the labor progression application, the reference time (time zero) used is the time from when the station value is zero. The proposed approach allows us to use all the subjects' data to estimate, including those who started the observation process before and after having reached time zero. In our application, we had such data. The model may be only weakly identifiable in situations where the observation process begins after a dilation value of zero. In this latter case, we recommend that the model not be used.

Although, the model is useful for understanding etiology through a close examination of the model parameters, the prediction of delivery times was not that good in some situations. Our prediction results demonstrated that prediction was accurate when the time to delivery was less than approximately 2 hours, while it was less accurate for longer range predictions. Although this is a bit disappointing, it is not surprising that a single longitudinal biomarker would be able to accurately predict delivery time among spontaneous pregnancies. The model can be extended to incorporate multiple longitudinal biomarkers in the future. Plots of the deciles of risk shows that there is a fairly large range in predictive probabilities and that the predictive model is well calibrated. Thus, the model may be very useful for predicting an individual woman's risk of a spontaneous delivery.

The proposed modeling framework addresses the issue of relating a longitudinal biomarker to survival when there is no clearly defined time zero. Although this was motivated by a study of labor progression where women come to the hospital at different times during their labor process, the methodology would apply to any study of the effect of a disease severity score on survival where subjects are enrolled at varying levels of initial disease severity. Another example comes from studying the relationship between disability and survival in multiple sclerosis. Disease severity in multiple sclerosis is often characterized using the expanded disability status scale (EDSS) which ranges between 0 and 10 in 0.5 increments. The proposed method may be used to assess the relationship between changes in EDSS and survival in a population where subjects enroll in the study at different levels of disability.

Acknowledgments

This research of Drs. Kim and Albert was supported by the intramural research program of National Cancer Institute.

References

- Abranowitz M and Stegun IA (1972). Handbook of mathematical functions with formulas, graphs, and mathematical tables. New York, Dover Publications, Inc.
- Celeux G, Forbes F, Robert CP, and D. M. Titterton D (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, 1:651–674.
- Chen M-H, Dey DK, and Shao Q-M (1999). A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, 94:1172–1186.
- Chen M-H, Shao OM, and Ibrahim JG (2000). *Monte Carlo Methods in Bayesian Computation*. New York, Springer-Verlag.
- Cowles C and Carlin BP (1996). Markov chain monte carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, 91:883–904.
- Hastings WK (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97–109.

- Kim S, Chen M-H, and Dey DK (2008). Flexible generalized t-link models for binary response data. *Biometrika*, 95:93–106.
- Liu JS (1994). The collapsed gibbs sampler in bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association*, 89:958–66.
- McLain AC and Albert PS (2014). Modeling longitudinal data with a random change point and no time-zero: applications to inference and prediction of the labor curve. *Biometrics*, 70:1052–1060. [PubMed: 25156417]
- Rizopoulos D (2012). *Joint models for longitudinal and time-to-event data: with applications in R*. Boca Raton: Chapman & Hall/CRC.
- Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society, B*, 64:583–639.

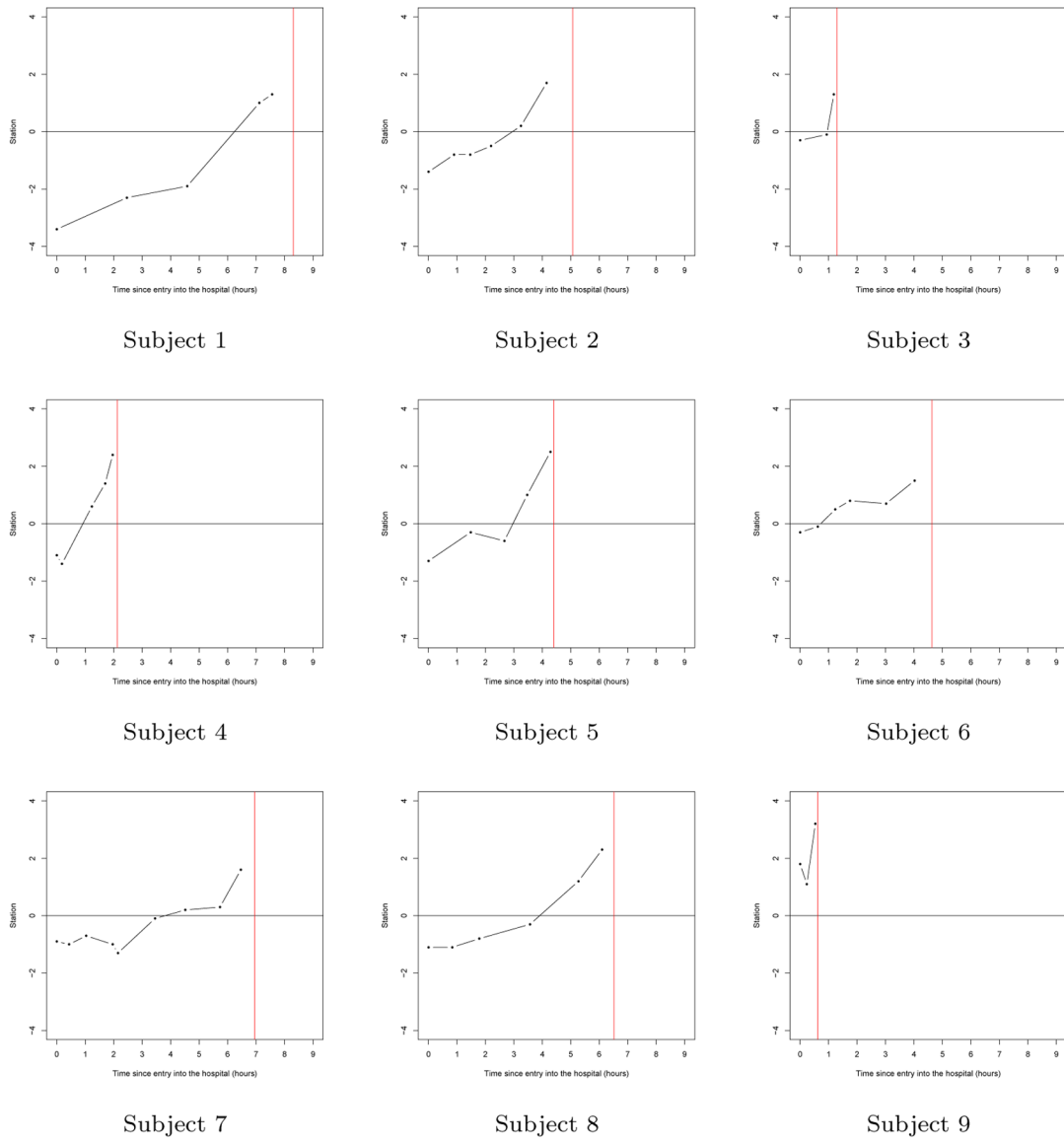


Figure 1: Representative station measurement trajectories resulting in a spontaneous delivery as a function of time from study entry (in hours). The red line represents the time of the spontaneous delivery.

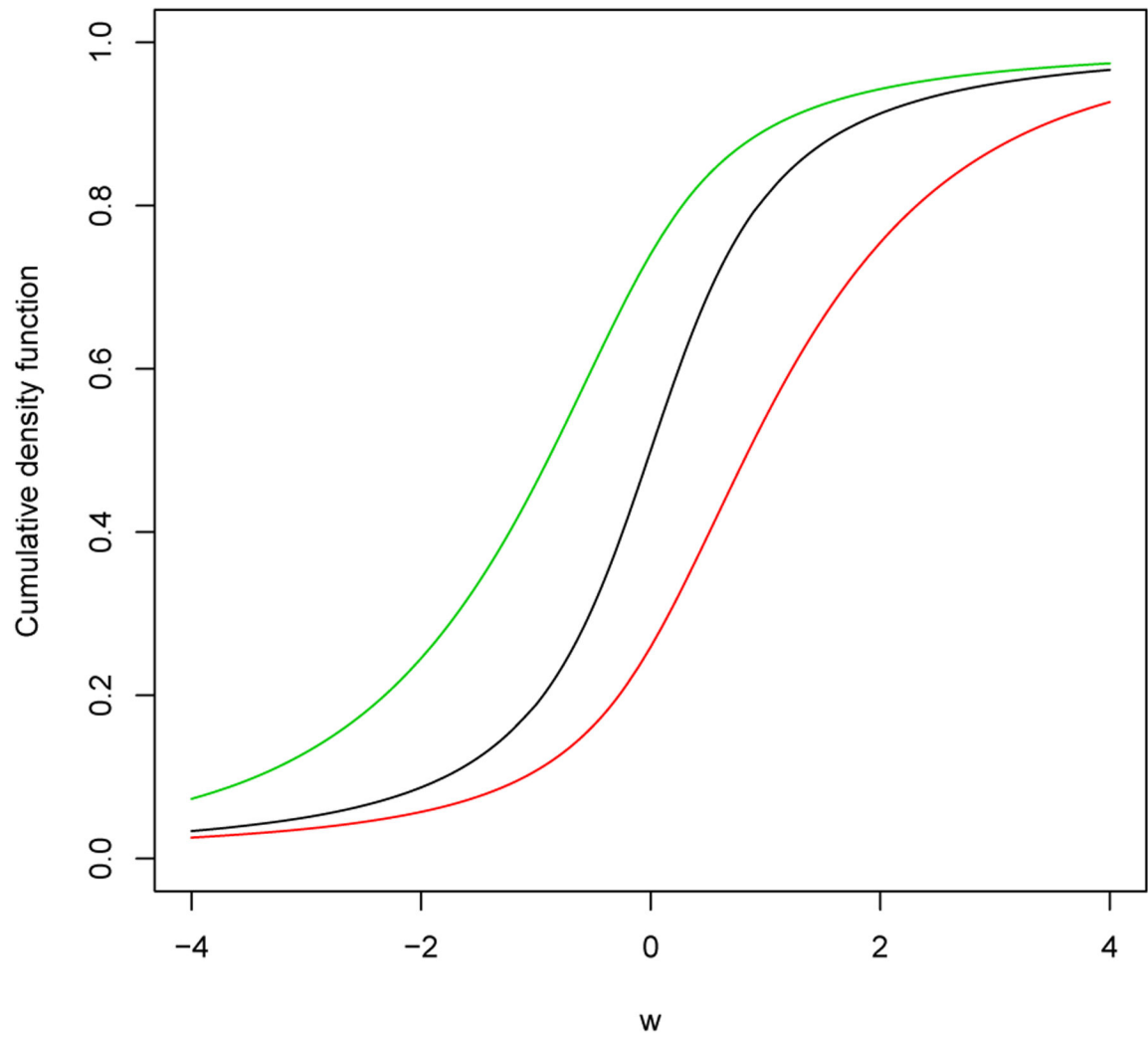


Figure 2: CDF plots of the skewed generalized t -distribution. Black line: symmetric, Red line: positive skewed, and Green line: negative skewed.

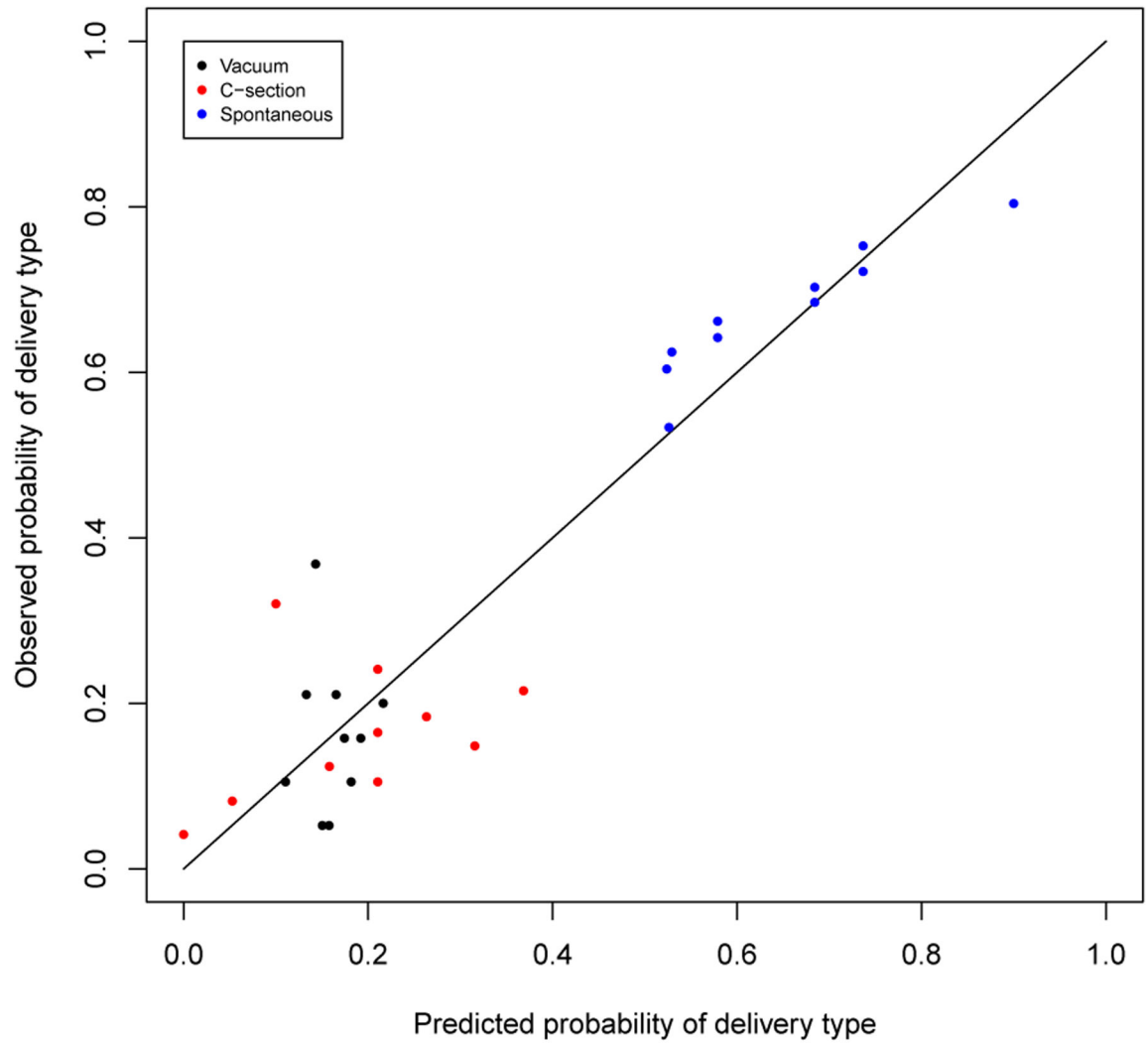


Figure 3: Decile plots of the predicted probability versus the corresponding observed proportions.

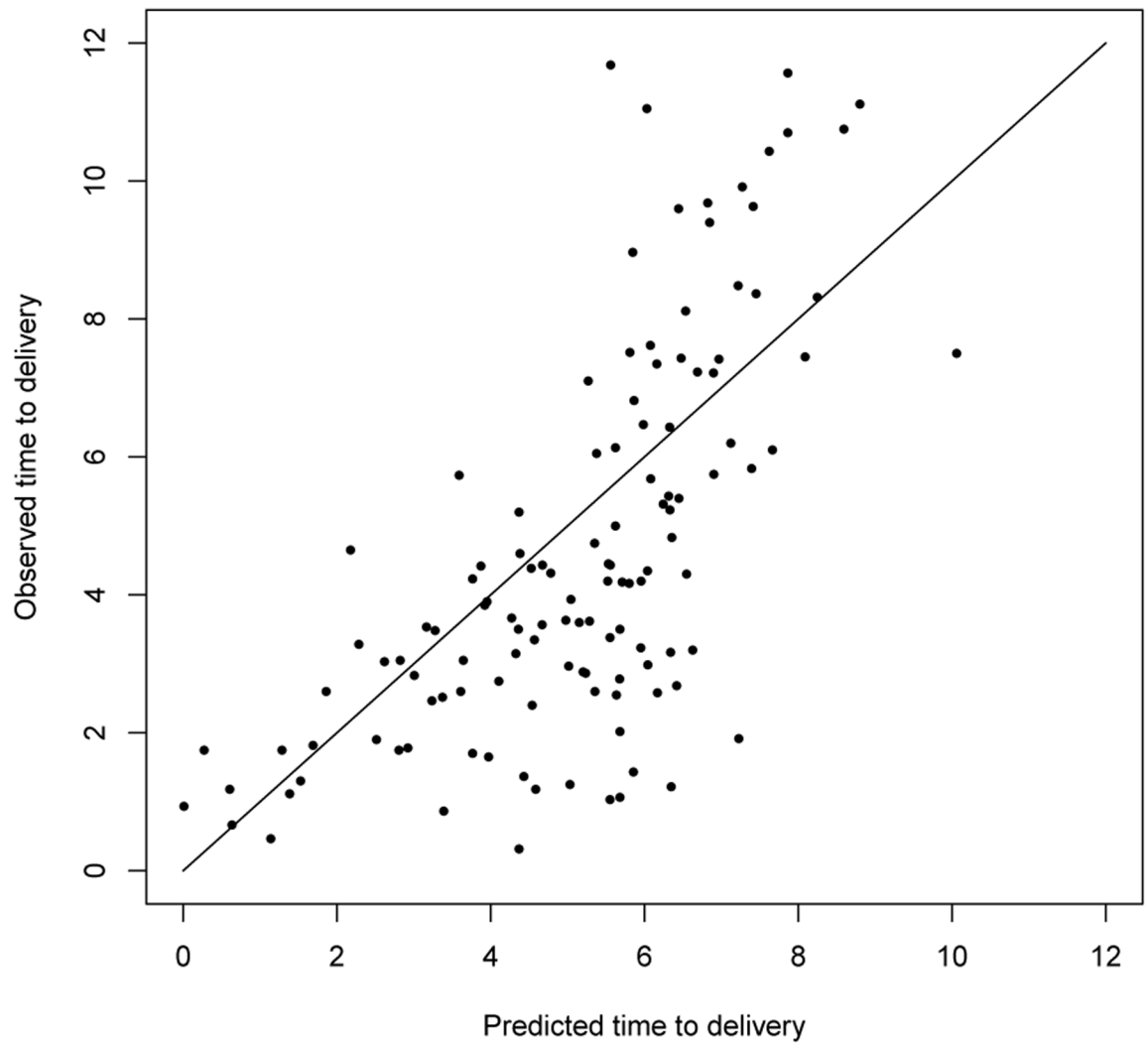


Figure 4:
Plots of predicted time to delivery from last observed station value

Table 1:

DIC Values

G_{ξ}	Normal	GT
$\Delta_{(0)}$	4278.64	4174.57
\mathcal{E}	4081.16	3612.15
$\mathcal{H}\mathcal{N}$	4140.03	3818.06

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2:

Posterior estimates under the best model

	Parameter	Posterior Mean	Posterior SD	95% HPD Interval
Longitudinal Measurements (y_{ij})	$\beta_1(t_{ij}): t_{ij} \geq 0$	-0.792	0.025	(-0.840, -0.743)
	$\beta_2(t_{ij}^2): t_{ij} \geq 0$	0.137	0.007	(0.125, 0.150)
	$\beta_3(t_{ij}): t_{ij} < 0$	1.687	0.294	(1.153, 2.355)
	σ_y^2	0.158	0.008	(0.143, 0.174)
	μ : mean scaling factor	-2.610	0.193	(-2.987, -2.253)
Time to Delivery (S_i)	α_0 (intercept)	7.830	0.216	(7.408, 8.257)
	α_1 (age)	-0.176	0.068	(-0.304, -0.040)
	$\theta_1(b_{i1})$	-5.367	0.736	(-6.834, -3.912)
	$\theta_2(b_{i2})$	-48.766	3.099	(-54.667, -42.536)
	$\theta_3(b_{i3})$	-0.091	0.616	(-1.337, 1.224)
	θ_4 (mode of delivery)	0.152	0.070	(0.010, 0.286)
	θ_5 (mode of delivery)	-0.296	0.114	(-0.538, -0.079)
	δ (skewness)	1.101	0.125	(0.863, 1.353)
	σ_t^2	0.664	0.517	(0.114, 1.643)
ν_1	13.857	9.866	(2.195, 34.350)	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3:

Posterior estimates under the best model

	Parameter	Posterior Mean	Posterior SD	95% HPD Interval
Mode of delivery (<i>D</i>)	γ_{10} (intercept) (vacuum)	-1.509	0.142	(-1.784, -1.230)
	γ_{11} (agle) (vacuum)	-0.133	0.138	(-0.401, 0.139)
	γ_{20} (intercept) (C-section)	-2.118	0.467	(-2.979, -1.468)
	γ_{21} (age) (C – section)	0.337	0.180	(-0.010, 0.692)
	$\phi_{11}(b_{i1})$ (vacuum)	0.310	0.813	(-1.298, 1.915)
	$\phi_{12}(b_{i2})$ (vacuum)	-4.148	3.159	(-10.223, 2.136)
	$\phi_{13}(b_{i3})$ (vacuum)	-0.155	0.574	(-1.309, 1.008)
	$\phi_{21}(b_{i1})$ (C – section)	-3.831	2.228	(-8.348, -0.018)
	$\phi_{22}(b_{i2})$ (C – section)	-16.161	4.405	(-24.730, -7.532)
	$\phi_{23}(b_{i3})$ (C – section)	-3.910	2.453	(-9.329, -0.800)
Variance of the random effects	Ω_{11}	0.109	0.016	(0.080, 0.141)
	Ω_{12}	-0.005	0.002	(-0.009, -0.002)
	Ω_{13}	-0.109	0.031	(-0.170, -0.053)
	Ω_{14}	0.538	0.103	(0.341, 0.738)
	Ω_{22}	0.004	0.001	(0.003, 0.005)
	Ω_{23}	-0.003	0.004	(-0.012, 0.004)
	Ω_{24}	0.027	0.016	(-0.003, 0.059)
	Ω_{33}	0.276	0.134	(0.068, 0.547)
	Ω_{34}	-1.210	0.339	(-1.893, -0.596)
	Ω_{44}	6.846	0.997	(5.072, 8.905)