

# Bayesian inference for nonlinear mixed-effects location scale and interval-censoring cure-survival models: An application to pregnancy miscarriage

Statistical Methods in Medical Research

2025, Vol. 34(8) 1525–1533

© The Author(s) 2025



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/09622802251345485

[journals.sagepub.com/home/smm](https://journals.sagepub.com/home/smm)

Danilo Alvares<sup>1</sup> , Cristian Meza<sup>2</sup> and Rolando De la Cruz<sup>3,4</sup> 

## Abstract

Motivated by a pregnancy miscarriage study, we propose a Bayesian joint model for longitudinal and time-to-event outcomes that takes into account different complexities of the problem. In particular, the longitudinal process is modeled by means of a nonlinear specification with subject-specific error variance. In addition, the exact time of fetal death is unknown, and a subgroup of women is not susceptible to miscarriage. Hence, we model the survival process via a mixture cure model for interval-censored data. Finally, both processes are linked through the subject-specific longitudinal mean and variance. A simulation study is conducted in order to validate our joint model. In the real application, we use individual weighted and Cox-Snell residuals to assess the goodness-of-fit of our proposal versus a joint model that shares only the subject-specific longitudinal mean (standard approach). In addition, the leave-one-out cross-validation criterion is applied to compare the predictive ability of both models.

## Keywords

Joint models, longitudinal data, mixed-effects location scale, three-parameter logistic model, time-to-event

## 1 Introduction

In Obstetrics research, a recurring interest is to study longitudinal beta-human chorionic gonadotropin ( $\beta$ -HCG) hormone measurements from women during the first quarter of their pregnancies and the pregnancy outcome from some women who had complications leading to miscarriage.<sup>1</sup> During the early stages of pregnancy, it is important to consider how the fluctuation in hormone concentration happens within such a framework, since it might alter the pregnancy's outcome.

This problem was first modeled by Marshall and Barón,<sup>2</sup> where they proposed a nonlinear mixed-effects model using a parametric logistic function to model hormone concentration over time using maximum likelihood estimates, and De la Cruz-Mesía and Quintana<sup>3</sup> provided a Bayesian approach. To model this biomarker and a pregnancy outcome together, De la Cruz et al.<sup>4,5</sup> explain the relationship between a binary response (pregnancy outcome) and the characteristics of longitudinal measurements (hormone levels). So, the joint model is made up of a logistic regression that has individual-specific random effects from a nonlinear mixed-effects model as variables. In De la Cruz et al.,<sup>4</sup> the authors compared a number of estimation techniques, including the Laplacian approximation, the naïve two-stage method, best linear unbiased

<sup>1</sup>MRC Biostatistics Unit, University of Cambridge, Cambridge, UK

<sup>2</sup>INGEMAT-CIMFAV, Faculty of Engineering, Universidad de Valparaíso, Valparaíso, Chile

<sup>3</sup>Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Penalolén, Chile

<sup>4</sup>Data Observatory Foundation, ANID Technology Center, Chile

### Corresponding author:

Danilo Alvares, MRC Biostatistics Unit, University of Cambridge, Forvie Site, Robinson Way, Cambridge, CB20SR, UK.

Email: [danilo.alvares@mrc-bsu.cam.ac.uk](mailto:danilo.alvares@mrc-bsu.cam.ac.uk)

prediction, and Gaussian and adaptive Gaussian quadratures. In De la Cruz et al.,<sup>5</sup> the authors proposed a Bayesian inference based on a Markov chain Monte Carlo sampler and introduced autocorrelated errors into the joint model.

Clinicians have a critical interest in being able to evaluate the relationship between longitudinally recorded  $\beta$ -HCG and time to early miscarriage. So, the use of subject-specific random effects from a mixed-effects model for longitudinal  $\beta$ -HCG data as predictors in a survival model is a typical joint modeling strategy proposed to accomplish this goal. However, it frequently occurs while examining time-to-event data that a portion of participants will never experience the relevant event. From a modeling perspective, the so-called cure models<sup>6</sup> incorporate such a characteristic, in which participants are believed to have been cured, and these event periods are considered limitless. Hence, our proposal includes a mixture cure specification<sup>7</sup> to model the time until fetal death, where such times are interval-censored. In addition, a nonlinear mixed-effects location scale (MELS) model<sup>8</sup> is proposed to capture the trajectories of  $\beta$ -HCG hormone and share the within-subject variability.

The remainder of the paper is organized as follows. In Section 2, we present a Chilean pregnancy miscarriage dataset, which was the motivation for the modeling developed in this paper. In Section 3, we gradually introduce the proposed model formulation, as well as its likelihood function and prior distributions. In Section 4, we evaluate the performance of our proposal with a simulation study. In Section 5, we discuss the results of two specifications of shared elements between the longitudinal and survival submodels. Finally, in Section 6, we conclude with a few general remarks. The models implemented in this paper were written in Stan<sup>9</sup> and are available at [www.github.com/danioloavres/BJM-MELS-Cure](http://www.github.com/danioloavres/BJM-MELS-Cure).

## 2 Pregnancy miscarriage data

Our motivation comes from a clinical trial study in a Chilean private assisted reproduction center. The data consist of longitudinal  $\beta$ -HCG hormone measurements (in a  $\log_{10}$  scale, which from now on will be denoted as  $\log(\beta\text{-HCG})$ ) from 173 young women during the first quarter of their pregnancies. This hormone is produced by the placenta during pregnancy. Typically,  $\beta$ -HCG levels increase steadily until the end of the first trimester (10 weeks of pregnancy), then decline as the pregnancy progresses.<sup>10</sup>

In our study data, 49 women had complications that led to a miscarriage. From now on, the term *abnormal* will be used to refer to this women's group. In contrast, 124 women had regular pregnancies and formed the *normal* group. Unfortunately,  $\beta$ -HCG levels during the first weeks of pregnancy are recorded infrequently and not always at the same stage of pregnancy for every woman.<sup>11,12</sup> Figure 1 shows the  $\beta$ -HCG profiles over time for both groups.

Figure 1 reveals that there is a notable difference between the longitudinal profiles of each group. In particular, the normal group has an increasing, nonlinear, and homogeneous evolution, while the abnormal group trajectories do not follow a clear pattern, but they have lower  $\beta$ -HCG hormone levels than the normal group and much more variability. These preliminary visual analyzes suggest that the distribution of longitudinal measurements depends on the pregnancy group. Hence, we can assume that the shared characteristics of  $\beta$ -HCG trajectories potentially help to explain the time until miscarriage. Furthermore, as previously pointed out, we can also see that both groups have irregular frequency and number of measurements. Table 1 shows the frequency of longitudinal measurements by pregnancy group.

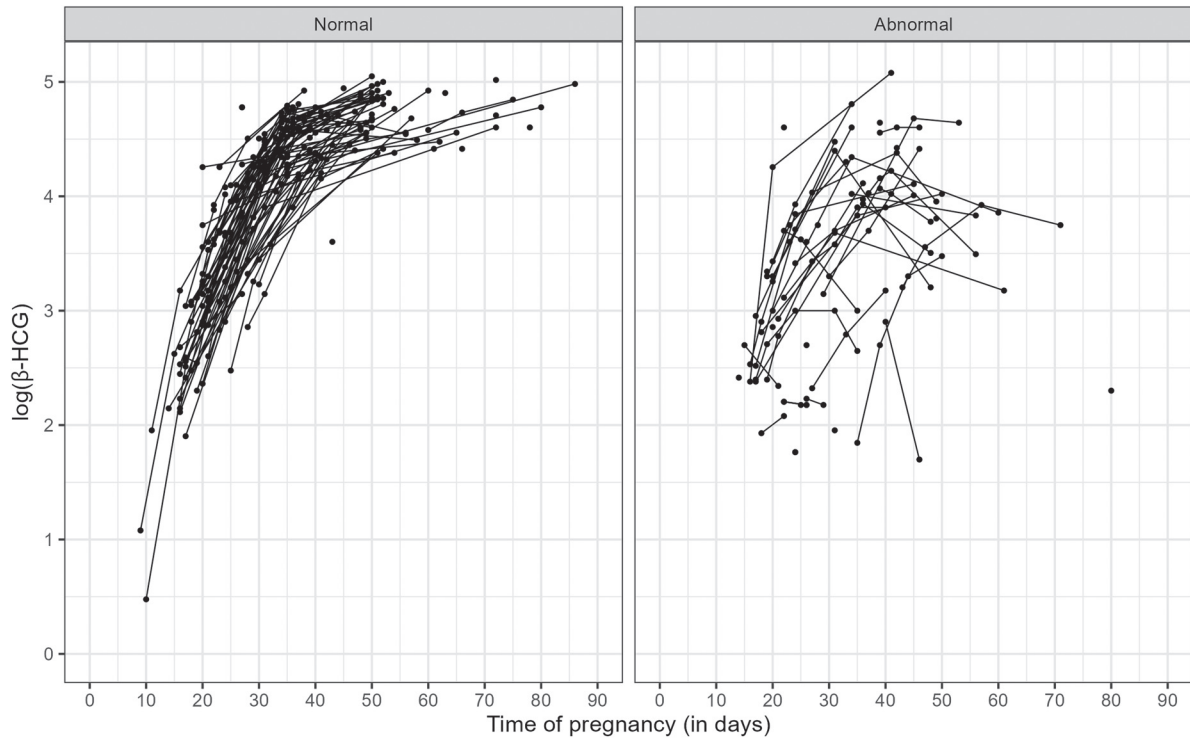
The main objective of this study is to analyze the association of the  $\beta$ -HCG levels with the time until fetal death. Here, it is important to note that only the abnormal group experiences the event of interest (fetal death). Another relevant characteristic of the problem is that the exact time of fetal death is unknown. However, miscarriage symptoms and emergency medical follow-up usually occur within a period of 10 days (regular time for miscarriage detection through medical visit and/or clinical examination in Chile) after the last measurement of the  $\beta$ -HCG hormone.<sup>13</sup> Assuming such a time window, Figure 2 shows the time interval in which fetal death occurred for the 49 women in the abnormal group. We can observe that most women had an early pregnancy loss between the 20th and 60th day of pregnancy.

The data of this study present a severe limitation which is the absence of (baseline) covariates. So, conclusions are limited to the knowledge that can be extracted from longitudinal measurements and survival times.

## 3 The Bayesian joint model

Conceptually, a joint model connects two or more processes through shared terms.<sup>14,15</sup> Here, such processes are described by a longitudinal submodel for the  $\beta$ -HCG hormone (endogenous time-varying covariate) and a survival submodel for the time until fetal death. Each element of our joint modeling proposal is introduced in the following.

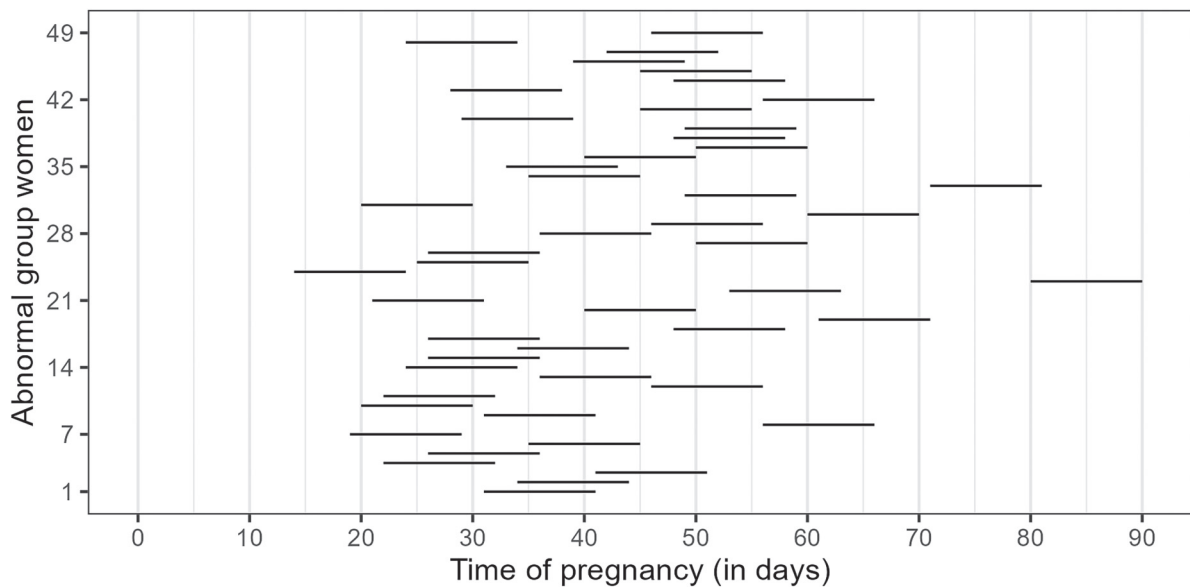
Let  $y_i(t)$  be the  $\log(\beta\text{-HCG})$  measurement associated with the  $i$ th woman,  $i = 1, \dots, N$ , measured at time  $t$ . It is worth noting that  $\log(\beta\text{-HCG})$  is always positive once  $\beta$ -HCG for pregnant women is greater than 1. Define the conditional distribution of  $y_i(t)$  given  $\theta$  (parameters),  $b_i$  (random effects), and  $\sigma$  (error standard deviation) as a generic additive error



**Figure 1.** Longitudinal measurements of  $\log(\beta\text{-HCG})$  by pregnancy group.  $\beta\text{-HCG}$ : beta human chorionic gonadotropin.

**Table 1.** Frequency of the number of individual longitudinal measurements by pregnancy group.

Pregnancy group	Number of longitudinal measurements					
	1	2	3	4	5	6
Normal	35	44	42	3	0	0
Abnormal	17	9	16	5	1	1



**Figure 2.** Time interval (10-day range from the last measurement of the  $\beta\text{-HCG}$  hormone) in which fetal death occurred for abnormal group women.  $\beta\text{-HCG}$ : beta human chorionic gonadotropin.

model:

$$y_i(t | \boldsymbol{\theta}, \mathbf{b}_i, \sigma) = \mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i) + \epsilon_i(t | \sigma) \tag{1}$$

where  $\mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i)$  represents the mean response at time  $t$  and  $\epsilon_i(t | \sigma)$  is a residual error. We assume that random effects,  $\mathbf{b}_i$ , given  $\boldsymbol{\Sigma}$ , follow a multivariate normal distribution with zero-mean vector and variance-covariance matrix  $\boldsymbol{\Sigma}$ . The residual errors are assumed to be conditionally independent and identically distributed as  $\epsilon_i(t | \sigma) \sim \text{Normal}(0, \sigma^2)$ .

The survival submodel aims to model the time until fetal death, which occurred within 10 days after the last measurement of the  $\beta$ -HCG hormone. Here, two important characteristics should be noted: (i) A part of the women (normal group) are not susceptible to the event of interest, and (ii) the exact time of fetal death is unknown, leading to interval-censored observations.

From a modeling perspective, (i) requires a mixture cure model, since some women have given birth and are therefore not susceptible to fetal loss.<sup>7</sup> Specifically, let  $Z$  be a binary random variable defined as 0 for a susceptible woman and 1 for an immune woman. So, the *incidence model* is given by  $P(Z_i = 1) = \eta = 1/(1 + \exp(-\nu))$ , where  $\eta$  represents the cure fraction. In addition, let  $T_i$  be the time until fetal death for the susceptible woman  $i$  (i.e.  $T_i$  conditional on  $Z_i = 0$ ), so the *latency model* is expressed through a proportional hazard specification:

$$h(t | \phi, \lambda, \alpha_1, \boldsymbol{\theta}, \mathbf{b}_i) = \phi t^{\phi-1} \exp \{ \lambda + \alpha_1 \mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i) \} \tag{2}$$

where  $\phi$  and  $\lambda$  are Weibull hazard shape and log-scale parameters, respectively. The term  $\mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i)$  has the role of connecting the longitudinal and survival submodels, while  $\alpha_1$  measures the strength of this association. We chose a Weibull baseline hazard because a similar study<sup>13</sup> using the same data corroborated that this specification is sufficiently adequate, but other alternatives could also be employed, such as piecewise and spline functions.<sup>16</sup>

### 3.1 Adding the three-parameter logistic specification

Typically, the longitudinal submodel (1) of a joint model is defined as a linear mixed-effects model with random intercept and slope.<sup>17-22</sup> However,  $\beta$ -HCG hormone trajectories clearly show nonlinear patterns (see Figure 1) that are not captured well with such a structure. To get around this issue, Marshall and Barón<sup>2</sup> successfully proposed a three-parameter logistic specification given by:

$$\mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i) = \frac{a_{i1}}{1 + \exp \left\{ -\frac{(t-a_{i2})}{a_{i3}} \right\}} \tag{3}$$

where  $a_{i1} = \exp\{\theta_1 + b_{i1}\}$ ,  $a_{i2} = \exp\{\theta_2 + b_{i2}\}$ , and  $a_{i3} = \exp\{\theta_3 + b_{i3}\}$ . The joint model (1)-(2) using the three-parameter logistic specification (3) will be called the *reference joint model*.

### 3.2 Adding the MELS specification

Figure 1 suggests that the variability of longitudinal trajectories may be a risk factor for the time until fetal death. In order to incorporate this characteristic into the modeling, we specify within-subject variances using a MELS model<sup>8</sup> as follows:

$$y_i(t | \boldsymbol{\theta}, \mathbf{b}_i) = \mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i) + \epsilon_i(t | \sigma_i) \tag{4}$$

where  $\sigma_i = \exp\{\theta_4\}$  for  $n_i = 1$  (number of longitudinal measurements) and  $\sigma_i = \exp\{\theta_4 + b_{i4}\}$  for  $n_i \geq 2$ . Hence, following the proposal of Barrett et al.,<sup>23</sup> we rewrite the hazard function (2) including  $\sigma_i^2$  as a second shared term:

$$h(t | \phi, \lambda, \alpha_1, \alpha_2, \boldsymbol{\theta}, \mathbf{b}_i) = \phi t^{\phi-1} \exp \{ \lambda + \alpha_1 \mu_i(t | \boldsymbol{\theta}, \mathbf{b}_i) + \alpha_2 \sigma_i^2 \} \tag{5}$$

### 3.3 Likelihood and priors

The likelihood function of the full parameter vector and random effects of the joint model (4)-(5) using the three-parameter logistic specification (3) is given by:

$$\mathcal{L}(\boldsymbol{\Phi}) = \prod_{i=1}^N \prod_{j=1}^{n_i} f(y_{ij} | \boldsymbol{\Phi}) f(\mathbf{b}_i | \boldsymbol{\Sigma}) \prod_{i \in \mathcal{I}} (1 - \eta) [S(t_{i,L} | \boldsymbol{\Phi}) - S(t_{i,R} | \boldsymbol{\Phi})] \prod_{i \in \mathcal{R}} [\eta + (1 - \eta) S(t_i | \boldsymbol{\Phi})] \tag{6}$$

**Table 2.** Bias and 95% CP for  $\theta$  and  $\alpha$  in Scenarios I (true model: (1)-(2)) and II (true model: (4)-(5)) considering 1–4 and 5–10 LMPI with  $N = 200$ .

Scenario	Parameter	1–4 LMPI		5–10 LMPI					
		Joint model (1)-(2)		Joint model (4)-(5)		Joint model (1)-(2)		Joint model (4)-(5)	
		Bias	95% CP	Bias	95% CP	Bias	95% CP	Bias	95% CP
I	$\theta_1$	−0.007	0.98	−0.008	0.97	−0.005	0.95	−0.003	0.95
	$\theta_2$	−0.020	0.97	−0.035	0.98	−0.015	0.94	−0.026	0.94
	$\theta_3$	−0.011	0.98	−0.001	0.96	−0.002	0.94	−0.001	0.95
	$\alpha_1$	0.046	0.92	0.032	0.95	0.002	0.96	−0.008	0.96
	$\theta_1$	−0.005	0.98	−0.003	0.96	−0.003	0.96	−0.002	0.95
	$\theta_2$	−0.018	0.94	−0.014	0.94	−0.011	0.95	−0.008	0.94
	$\theta_3$	−0.010	0.98	−0.008	0.97	−0.008	0.96	−0.005	0.96
	$\theta_4$	–	–	0.032	0.93	–	–	0.004	0.95
	$\alpha_1$	−0.331	0.83	−0.021	0.93	−0.210	0.89	−0.012	0.94
II	$\alpha_2$	–	–	−0.024	0.93	–	–	0.013	0.95

CP: coverage probability; LMPI: longitudinal measurements per individual.

where  $y_{ij}$  is the value of  $\log(\beta\text{-HCG})$  for the  $i$ th woman at visit  $j$ ;  $(t_{i,L}, t_{i,R})$  is the time interval in which the miscarriage occurred for a woman  $i$  belonging to the abnormal group ( $\mathcal{I}$ );  $t_i$  is the right-censored observation (last longitudinal measurement time) of a woman  $i$  belonging to the normal group ( $\mathcal{R}$ );  $\Phi = (\theta, \nu, \phi, \lambda, \alpha_1, \alpha_2, \Sigma, b_1, \dots, b_n)$  denotes the full parameter vector and random effects;  $f(y_{ij} | \Phi)$  represents the conditional probability density function of  $y_{ij}$  given  $\Phi$  described in (4) with respective random effects density function denoted by  $f(b_i | \Sigma)$ ; and  $S(t | \Phi)$  is the survival function derived from (5).

We assume independent and proper prior distributions.<sup>24</sup> More specifically, all longitudinal and survival fixed effects,  $(\theta_1, \theta_2, \theta_3, \theta_4, \nu, \lambda, \alpha_1, \alpha_2)$ , follow a  $\text{Normal}(0, 10^2)$ ; the error variance,  $\sigma^2$ , and the Weibull shape parameter,  $\phi$ , follow a  $\text{half-Cauchy}(0, 1)$ ;<sup>25</sup> and the random effects variance-covariance matrix  $\Sigma$  follows an  $\text{inverse-Wishart}(\mathbf{I}_5, 4)$ ,<sup>26</sup> where  $\mathbf{I}_5$  represents a  $5 \times 5$  identity matrix. We previously investigated the sensitivity of our prior distributions compared to vaguer ones,  $\text{Normal}(0, 100^2)$  and  $\text{half-Cauchy}(0, 10)$ , and we concluded that our choice is weakly informative, since the results were equivalent, differing only in computational time.

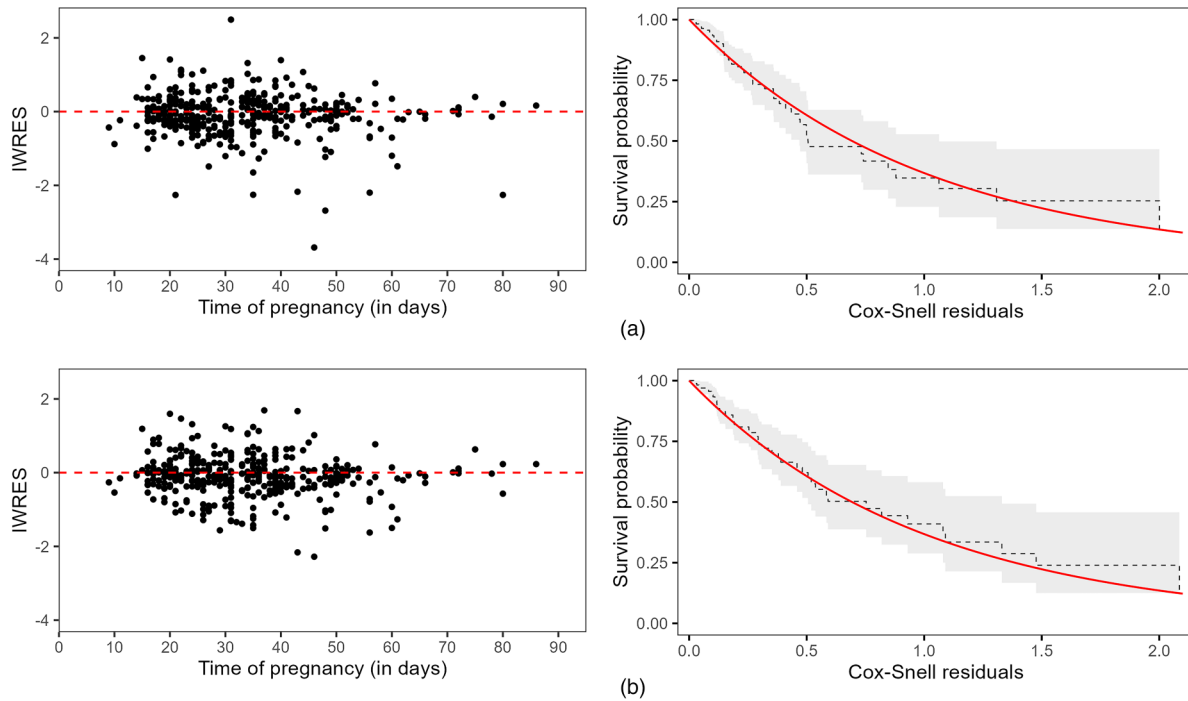
### 4 Simulation study

We conducted a simulation study to evaluate the performance of our proposal in estimating the parameters  $\theta$  and  $\alpha$ , compared to the reference joint model. We explored two scenarios: (I) simulated data from the joint model (1)-(2) (share  $\mu_i$ ) and (II) simulated data from the joint model (4)-(5) (share  $\mu_i$  and  $\sigma_i^2$ ). In both cases, we considered 1–4 and 5–10 longitudinal measurements per individual,  $N = 200$  (sample size), and 1000 repetitions. The specification of the parameters is based on the fit of each model with the pregnancy miscarriage data (see Section 2). Specifically, Scenario I:  $\theta_1 = 1.5, \theta_2 = 2.7, \theta_3 = 1.9, \sigma = 0.25, \lambda = -14.5, \phi = 4, \alpha_1 = -0.6$ , and  $\Sigma = \text{diag}(0.02, 0.08, 0.17)$ ; Scenario II:  $\theta_1 = 1.5, \theta_2 = 2.7, \theta_3 = 1.9, \theta_4 = -2, \lambda = -14.7, \phi = 3.9, \alpha_1 = -0.4, \alpha_2 = 3.1$ , and  $\Sigma = \text{diag}(0.02, 0.07, 0.1, 0.6)$ . Table 2 summarizes the results and terms of bias and 95% coverage probability.

The population parameters  $\theta$ 's are well estimated in all scenarios for both joint models. In Scenario I, where (1)-(2) is the true model, our proposal appropriately estimates the association parameter  $\alpha_1$  even with few longitudinal measurements per individual. In Scenario II, where (4)-(5) is the true model, our proposal is also suitable, as expected, while the reference joint model produces biased estimates of the parameter associated with the shared mean response.

### 5 Application

We implemented Bayesian joint models (1)-(2) (share  $\mu_i$ ) and (4)-(5) (share  $\mu_i$  and  $\sigma_i^2$ ) in `rstan`<sup>9</sup> R-package (version 2.32.7) and run each of them with three Markov chains and 6000 iterations. The first half of the posterior samples was discarded (warm-up period), and then we made the inference with the remaining ones (9000 posterior samples). Convergence and efficiency were checked through  $Rhat$  and effective sample size.<sup>24</sup> All models were run on a Dell laptop with 2.2 GHz Intel Core i7, 32 GB RAM, OS Windows.



**Figure 3.** First column: Individual weighted residuals (IWRES). Second column: Kaplan–Meier estimates of the Cox–Snell residuals (dashed black line) and its 95% confidence interval (gray shadow), where the solid red line represents the survival function of the unit exponential distribution. (a) Longitudinal and survival residuals from joint model (1)–(2) (share  $\mu_i$ ); (b) Longitudinal and survival residuals from joint model (4)–(5) (share  $\mu_i$  and  $\sigma_i^2$ ).

**Table 3.** Posterior summary for the parameters of interest from joint models (1)–(2) (share  $\mu_i$ ) and (4)–(5) (share  $\mu_i$  and  $\sigma_i^2$ ) using the three-parameter logistic specification (3).

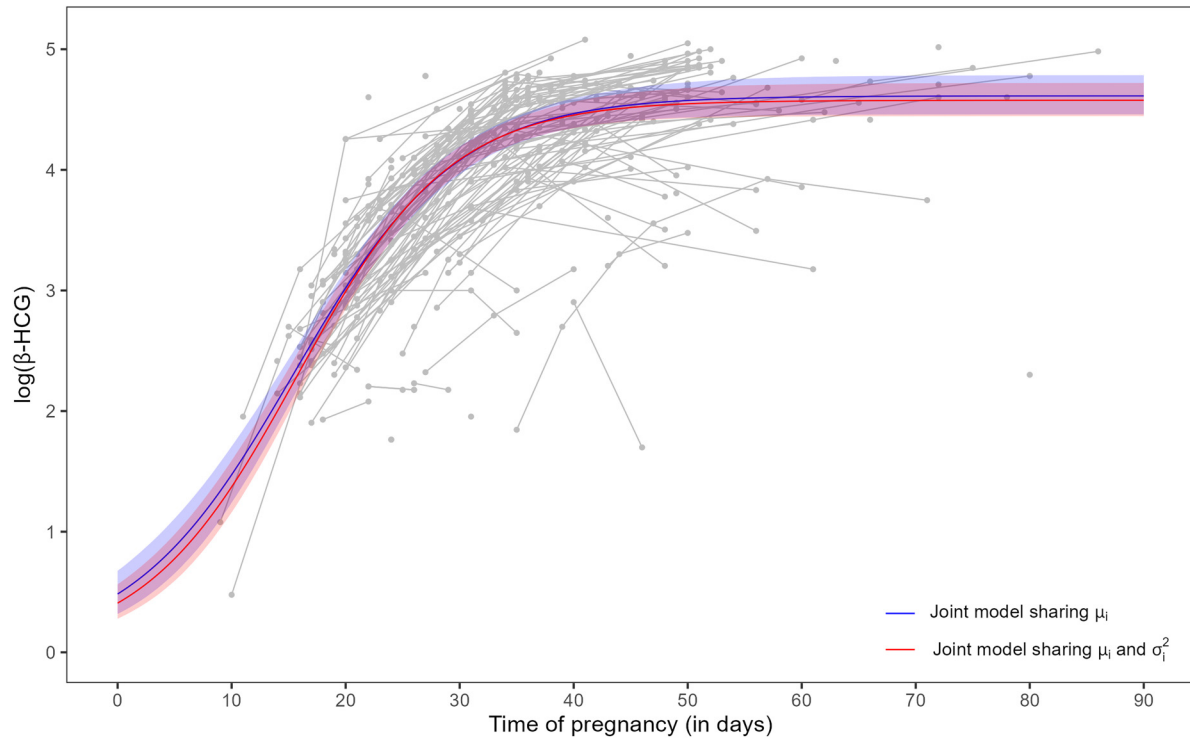
Parameter	Joint model (1)–(2)		Joint model (4)–(5)	
	Mean	95% CI	Mean	95% CI
$a_{10} = \exp\{\theta_1\}$	4.620	(4.469, 4.788)	4.577	(4.446, 4.721)
$a_{20} = \exp\{\theta_2\}$	15.438	(14.139, 16.621)	15.712	(14.589, 16.726)
$a_{30} = \exp\{\theta_3\}$	7.196	(6.182, 8.348)	6.747	(5.927, 7.679)
$\alpha_1$	−0.610	(−1.027, −0.190)	−0.440	(−0.839, −0.029)
$\alpha_2$	–	–	3.149	(0.010, 8.212)
<b>LOO-CV</b>	612		392	
<b>Time (in minutes)</b>	27		80	

LOO-CV: leave-one-out cross-validation.

We analyzed the goodness-of-fit through longitudinal and survival residuals.<sup>27</sup> Specifically, individual weighted residuals (IWRES) for longitudinal submodels and Cox-Snell residuals for survival submodels, considering interval-censored observations.<sup>28</sup> Figure 3 shows both residuals by joint model.

In both cases, IWRES did not suggest any model misspecification, but it is possible to observe less dispersion of residuals considering the MELS specification (Figure 3b). Additionally, the Kaplan–Meier estimates of the Cox–Snell residuals were close to the theoretical survival curves (unit exponential distribution), indicating a suitable fit for both models.

We used the leave-one-out cross-validation (LOO-CV)<sup>29</sup> to select the best joint model specification. This criterion is based on the out-of-sample prediction accuracy from a fitted Bayesian model using the log-likelihood evaluated at posterior simulations of the parameter values.<sup>30</sup> Interpretatively, a lower LOO-CV value indicates a better model fit. It is worth noting that LOO-CV compares models based on their predictive performance, so it can be used for different classes of models, including non-nested specifications. Table 3 shows a posterior summary for both joint models as well as their respective LOO-CV.



**Figure 4.** Posterior mean and 95% credible interval for the three-parameter logistic specification (3) from joint models (1)-(2) (share  $\mu_i$ , blue color) and (4)-(5) (share  $\mu_i$  and  $\sigma_i^2$ , red color) using population parameters  $a_{10}$ ,  $a_{20}$ , and  $a_{30}$  (see Table 3). Trajectories in gray color are the longitudinal measurements of  $\log(\beta\text{-HCG})$  of the 173 women in the study (both groups).  $\beta\text{-HCG}$ : beta-human chorionic gonadotropin.

In Table 3, we can see that the population parameters,  $a_{10}$ ,  $a_{20}$  and  $a_{30}$ , that describe the median trajectory of the three-parameter logistic specification (3), are slightly different. Figure 4 shows such a trajectory using each joint model. In fact, the difference between the curves is minimal over the range of measurements.

Still in Table 3, the association parameter  $\alpha_1$ , which connects the longitudinal process mean ( $\mu_i$ ) to the survival submodel, is negative for both joint models. This means that higher  $\beta\text{-HCG}$  hormone levels have a protective effect in terms of time until fetal death. This result corroborates the previous visual inspection (see Figure 1) that the abnormal group (women susceptible to miscarriage) presents longitudinal trajectories with lower  $\beta\text{-HCG}$  than the normal group. The joint model (4)-(5) also shows a strong and positive association ( $\alpha_2$ ) between the within-subject variance  $\sigma_i^2$  and the time until fetal death. This is interpreted as greater intra-woman  $\beta\text{-HCG}$  variability leads to a higher risk of miscarriage. This result is also consistent with the observed pattern of  $\beta\text{-HCG}$  hormone levels in the abnormal group. In terms of model selection, the LOO-CV criterion indicates a better fit using the joint model that shares  $\mu_i$  and  $\sigma_i^2$ , but this model takes more than twice as long as the reference joint model.

## 6 Discussion

In this paper, we have proposed a Bayesian joint model based on a nonlinear MELS submodel for longitudinal data and a mixture cure submodel for interval-censored survival data. In addition, such submodels have shared terms described through the subject-specific longitudinal mean and variance.

We have compared our proposal with a reference joint model that shares only the subject-specific longitudinal mean. In the simulation study, our proposal has performed equivalently or better than the competing model. In the application, both approaches have shown suitable goodness-of-fit in terms of longitudinal and survival residuals (see Figure 3). However, the inclusion of within-subject variance as a shared term contributed to a better understanding of the  $\beta\text{-HCG}$  hormone pattern of women who had a miscarriage. Specifically, we have argued that increasing such variance leads to higher risks of fetal loss (see Table 3). Still, we highlight that this conclusion should be taken with extreme caution, as our study does not include baseline variables (not available) that could potentially be relevant risk factors.

Both joint models presented high computational times (27 and 80 minutes) given that our Chilean pregnancy miscarriage study has a relatively small sample size (173 women with few longitudinal measurements). These times may be reduced using two-stage strategies that preferentially correct the estimation bias.<sup>31</sup>

In conclusion, we hope this paper inspires other authors to consider all complex elements of real data in their joint modeling. In particular, we encourage researchers to adapt our codes for other problems, as well as to implement our joint model proposal in other statistical Bayesian model tools, such as JAGS<sup>32</sup> and INLA.<sup>33</sup>

### Acknowledgments

The authors thank Guillermo Marshall for facilitating the pregnancy miscarriage data, as well as the Editor, the Associated Editor, and three referees for their multiple, detailed, and constructive reports.


### Declaration of conflicting interests


The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: D.A. was supported by the Medical Research Council grant MC\_UU\_00002/5. C.M. was supported by the ANID/FONDECYT grant 1190801. R.D.L.C. was supported by grants ANID/FONDECYT 1181662, ANID/PIA/ANILLOS ACT210096 & Data Observatory Foundation, and ANID Technology Center DO210001.

### ORCID iDs

Danilo Alvares  <https://orcid.org/0000-0003-3764-0397>

Rolando De la Cruz  <https://orcid.org/0000-0002-1415-1809>

### References

1. Confino E, Demir RH, Friberg J, et al. The predictive value of hCG beta subunit levels in pregnancies achieved by in vitro fertilization and embryo transfer: An international collaborative study. *Fertil Steril* 1986; **45**: 526–531.
2. Marshall G and Barón AE. Linear discriminant models for unbalanced longitudinal data. *Stat Med* 2000; **19**: 1969–1981.
3. De la Cruz-Mesía R and Quintana FA. A model-based approach to Bayesian classification with applications to predicting pregnancy outcomes from longitudinal beta-hCG profiles. *Biostatistics* 2007; **8**: 228–238.
4. De la Cruz R, Marshall G and Quintana FA. Logistic regression when covariates are random effects from a non-linear mixed model. *Biometrical J* 2011; **53**: 735–749.
5. De la Cruz R, Meza C, Arribas-Gil A, et al. Bayesian regression analysis of data with random effects covariates from nonlinear longitudinal measurements. *J Multivar Anal* 2016; **143**: 94–106.
6. Peng Y and Yu B. *Cure models: Methods, applications, and implementation*. 1st ed. New York, NY, USA: Chapman & Hall/CRC, 2021.
7. Berkson J and Gage RP. Survival curve for cancer patients following treatment. *J Am Stat Assoc* 1952; **47**: 501–515.
8. Hedeker D, Mermelstein RJ and Demirtas H. An application of a mixed-effects location scale model for analysis of ecological momentary assessment (EMA) data. *Biometrics* 2008; **64**: 627–634.
9. Stan Development Team. *RStan: The R interface to Stan*. R package version 2.32.7, New York, http://mc-stan.org/, 2025.
10. Nwabuobi C, Arlier S, Schatz F, et al. HCG: Biological functions and clinical applications. *Int J Mol Sci* 2017; **18**: 1–15.
11. De la Cruz-Mesía R, Fuentes C, Meza C, et al. Predicting pregnancy outcomes using longitudinal information: A penalized splines mixed-effects model approach. *Stat Med* 2017; **36**: 2120–2134.
12. Gaskins JT, Fuentes C and De la Cruz R. A Bayesian nonparametric model for classification of longitudinal profiles. *Biostatistics* 2023; **24**: 209–225.
13. De la Cruz R, Lavielle M, Meza C, et al. A joint analysis proposal of nonlinear longitudinal and time-to-event right-, interval-censored data for modeling pregnancy miscarriage. *Comput Biol Med* 2024; **182**: 1–17.
14. Rizopoulos D. *Joint models for longitudinal and time-to-event data: With applications in R*. 1st ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2012.
15. Elashoff R, Li G and Li N. *Joint modeling of longitudinal and time-to-event data*. 1st ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2016.
16. Lázaro E, Armero C and Alvares D. Bayesian regularization for flexible baseline hazard functions in Cox survival models. *Biometrical J* 2021; **63**: 7–26.
17. Crowther MJ, Abrams KR and Lambert PC. Joint modeling of longitudinal and survival data. *Stata J* 2013; **13**: 165–184.
18. Proust-Lima C, Séne M, Taylor JMG, et al. Joint latent class models for longitudinal and time-to-event data: A review. *Stat Methods Med Res* 2014; **23**: 74–90.
19. Furgal AKC, Sen A and Taylor JMG. Review and comparison of computational approaches for joint longitudinal and time-to-event models. *Int Stat Rev* 2019; **87**: 393–418.

20. Papageorgiou G, Mauff K, Tomer A et al. An overview of joint modeling of time-to-event and longitudinal outcomes. *Annu Rev Stat Appl* 2019; **6**: 223–240.
21. Alsefri M, Sudell M, García-Fiñana M, et al. Bayesian joint modelling of longitudinal and time to event data: A methodological review. *BMC Med Res Methodol* 2020; **20**: 1–17.
22. Alvares D and Rubio FJ. A tractable Bayesian joint model for longitudinal and survival data. *Stat Med* 2021; **40**: 4213–4229.
23. Barrett JK, Huille R, Parker R, et al. Estimating the association between blood pressure variability and cardiovascular disease: An application using the ARIC study. *Stat Med* 2018; **38**: 1855–1868.
24. Gelman A, Carlin JB, Stern HS, et al. *Bayesian data analysis*. 3rd ed. Boca Raton, FL, USA: Chapman & Hall/CRC, 2013.
25. Rubio F and Steel M. Flexible linear mixed models with improper priors for longitudinal and survival data. *Electron J Stat* 2018; **12**: 572–598.
26. Schuurman NK, Grasman RPPP and Hamaker EL. A comparison of inverse-Wishart prior specifications for covariance matrices in multilevel autoregressive models. *Multivariate Behav Res* 2016; **51**: 185–206.
27. Desmée S, Mentré F, Veyrat-Follet C, et al. Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients. *Biometrics* 2017; **73**: 305–312.
28. Farrington CP. Residuals for proportional Hazards models with interval-censored survival data. *Biometrics* 2000; **56**: 473–482.
29. Vehtari A, Mononen T, Tolvanen V, et al. Bayesian leave-one-out cross-validation approximations for Gaussian latent variable models. *J Mach Learn Res* 2016; **17**: 1–38.
30. Vehtari A, Gelman A and Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* 2017; **27**: 1413–1432.
31. Alvares D and Leiva-Yamaguchi V. A two-stage approach for Bayesian joint models: Reducing complexity while maintaining accuracy. *Stat Comput* 2023; **33**: 1–11.
32. Alvares D, Lázaro E, Gómez-Rubio V, et al. Bayesian survival analysis with BUGS. *Stat Med* 2021; **40**: 2765–3020.
33. Alvares D, van Niekerk J, Krainski ET, et al. Bayesian survival analysis with INLA. *Stat Med* 2024; **43**: 3975–4010.