


Article

A Two-Stage Approach for Bayesian Joint Models of Longitudinal and Survival Data: Correcting Bias with Informative Prior

Valeria Leiva-Yamaguchi [†] and Danilo Alvares ^{*,†} 

Department of Statistics, Pontificia Universidad Católica de Chile, Macul, Santiago 7820436, Chile; vjleiva@mat.uc.cl

* Correspondence: dalvares@mat.uc.cl

† These authors contributed equally to this work.

Abstract: Joint models of longitudinal and survival outcomes have gained much popularity in recent years, both in applications and in methodological development. This type of modelling is usually characterised by two submodels, one longitudinal (e.g., mixed-effects model) and one survival (e.g., Cox model), which are connected by some common term. Naturally, sharing information makes the inferential process highly time-consuming. In particular, the Bayesian framework requires even more time for Markov chains to reach stationarity. Hence, in order to reduce the modelling complexity while maintaining the accuracy of the estimates, we propose a two-stage strategy that first fits the longitudinal submodel and then plug the shared information into the survival submodel. Unlike a standard two-stage approach, we apply a correction by incorporating an individual and multiplicative fixed-effect with informative prior into the survival submodel. Based on simulation studies and sensitivity analyses, we empirically compare our proposal with joint specification and standard two-stage approaches. The results show that our methodology is very promising, since it reduces the estimation bias compared to the other two-stage method and requires less processing time than the joint specification approach.

Keywords: Bayesian inference; bias reduction; individual fixed-effect; Stan



Citation: Leiva-Yamaguchi, V.; Alvares, D. A Two-Stage Approach for Bayesian Joint Models of Longitudinal and Survival Data: Correcting Bias with Informative Prior. *Entropy* **2021**, *23*, 50. <https://doi.org/10.3390/e23010050>

Received: 30 November 2020

Accepted: 27 December 2020

Published: 31 December 2020

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Joint models of longitudinal and survival data have been an essential statistical tool in medical research [1,2]. This class of models became popular due to its ability to provide complete inference (longitudinal, survival, and association between both of them), reduce estimation bias, increase statistical efficiency, and conveniently make predictions of outcomes [3–5]. However, there ain't no such thing as a free lunch. The complexity of these models makes the computational process quite demanding and sometimes impractical.

In this paper, we focus on general contexts in which longitudinal measurements are observed strictly before the survival time [6]. This framework has been analysed in several applications, see References [7–9] for a review on joint models up to date, and it has at least two drawbacks: (i) identifiability problems due to the large number of parameters [7,10–13] and (ii) requirement for numerical integrations that can make the inferential process time-consuming [14–18].

Two-stage approaches alleviate both problems that arise with simultaneous inference for joint models [19,20]. Typically, the two-stage approach fits the longitudinal submodel first and then uses the estimated parameters to approximate the longitudinal trajectory, as an endogenous time-varying covariate, within the survival submodel. This strategy is usually simple to implement and allows us to use flexible models available in standard longitudinal and survival analyses packages (separately). In the current literature of joint models, there are different proposals for two-stage methods in both frequentist [20–24]

and Bayesian [25–27] approaches. These two-stage procedures speed up processing time by estimating two less complex submodels than the joint model. However, the main weakness of this methodology is that by ignoring the joint nature between both processes, the estimates of the survival regression parameters are often biased [22,28–30].

From a Bayesian perspective, we work around this problem by proposing a two-stage approach that, after fitting the longitudinal submodel, corrects bias through an individual and multiplicative fixed-effect with highly informative prior inserted in the survival submodel.

The paper is organised as follows—Section 2 introduces a general formulation of joint models. Sections 3 and 4 describe the standard joint and two-stage approaches. Section 5 presents our two-stage strategy. Section 6 validates and compares the performance of our proposal against the other standard approaches. Finally, Section 7 discusses the advantages, limitations and extensions of our methodology. Appendices A and B show sensitivity analyses and other simulated scenarios.

2. Bayesian Joint Model Formulation

We assume that there are n individuals with repeated measures and time to an event of interest individually associated. In particular, underlying characteristics from the longitudinal process, which models repeated measures, are shared with the time-to-event process [30].

2.1. Longitudinal Submodel

We use the well-known linear mixed-effects specification to model the repeated measures over time [31,32]. In this case, the response variable $y_i(t)$ of individual i at time t is given by:

$$y_i(t) = \mu_i(t) + \epsilon_i(t) = \mathbf{x}_{L,i}^\top(t)\boldsymbol{\beta} + \mathbf{z}_i^\top(t)\mathbf{b}_i + \epsilon_i(t), \quad (1)$$

$$\mathbf{b}_i \stackrel{\text{i.i.d.}}{\sim} \mathbf{N}(\mathbf{0}, \boldsymbol{\Sigma}) \quad \text{and} \quad \epsilon_i(t) \stackrel{\text{i.i.d.}}{\sim} \mathbf{N}(0, \sigma^2),$$

where the true unobserved value of the longitudinal outcome at time t , $\mu_i(t)$, is characterised by the linear combination between the covariate vectors, $\mathbf{x}_{L,i}(t)$ and $\mathbf{z}_i(t)$, and their respective fixed ($\boldsymbol{\beta}$) and random (\mathbf{b}_i) effect vectors; \mathbf{b}_i represents the vector of individual random effects with a $K \times K$ variance-covariance matrix $\boldsymbol{\Sigma}$, where K is the number of random effects; and $\epsilon_i(t)$ denotes the measurement error term with variance σ^2 .

2.2. Survival Submodel

The proportional hazards specification is widely used to model this type of problem [33]. Let T_i^* denote the event time for individual i , C_i the censoring time, $T_i = \min\{T_i^*, C_i\}$ the observed time, and $\delta_i = \mathbf{I}(T_i^* \leq C_i)$ the event indicator. So, the hazard function of the survival time T_i of individual i is expressed by:

$$h_i(t \mid \mathcal{M}_i(t)) = h_0(t) \exp\left\{\mathbf{x}_{S,i}^\top \boldsymbol{\gamma} + \alpha \mu_i(t)\right\}, \quad (2)$$

where $h_0(t)$ represents an arbitrary baseline hazard function at time t and $\mathbf{x}_{S,i}$ is a covariate vector with coefficients $\boldsymbol{\gamma}$. $\mathcal{M}_i(t) = \{\mu_i(l), 0 \leq l < t\}$ denotes the history of the longitudinal process up to t ; $\mu_i(t)$ is defined as in (1) and has the role of connecting both processes, while α measures the strength of this association. In order to simplify the notation, we will omit the term $\mathcal{M}_i(t)$ when specifying a hazard function.

2.3. Prior Distributions

To complete the Bayesian joint model formulation, we have to assign prior distributions to all parameters and hyperparameters. As a standard specification, we assume independent and diffuse prior, that is, proper distributions with a large variance [34]. More specifically, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and α follow Normal distributions with mean at zero and large variance; σ follows a weakly-informative half-Cauchy(0, 5) [35]; and $\boldsymbol{\Sigma}$ follows an inverse-Wishart(V, r), where

V is a $K \times K$ identity matrix, $r = K$ is the degrees-of-freedom parameter [36]. Once the baseline hazard function $h_0(t)$ is defined, diffuse priors are also specified for its parameters.

3. Joint Specification (JS) Approach

Let \mathbf{y} and \mathbf{s} be the longitudinal and survival data, respectively. The vector of all parameters and hyperparameters is specified by $\boldsymbol{\theta}$ and the random effects by \mathbf{b} . So, the full joint distribution of $(\mathbf{y}, \mathbf{s}, \mathbf{b}, \boldsymbol{\theta})$ can be factorised as the product of the joint conditional distribution $f(\mathbf{y}, \mathbf{s} | \mathbf{b}, \boldsymbol{\theta})$, the conditional distribution of the random effects $f(\mathbf{b} | \boldsymbol{\theta})$, and the prior distribution $\pi(\boldsymbol{\theta})$. Equationally,

$$f(\mathbf{y}, \mathbf{s}, \mathbf{b}, \boldsymbol{\theta}) = f(\mathbf{y}, \mathbf{s} | \mathbf{b}, \boldsymbol{\theta}) f(\mathbf{b} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \quad (3)$$

There are different proposals for the specification of the conditional distribution $f(\mathbf{y}, \mathbf{s} | \mathbf{b}, \boldsymbol{\theta})$ [37]. However, the most widely used approach is the shared-parameter specification [38], in which it assumes that the longitudinal process is conditionally independent of the survival process given the shared information:

$$f(\mathbf{y}, \mathbf{s} | \mathbf{b}, \boldsymbol{\theta}) = f(\mathbf{y} | \mathbf{b}, \boldsymbol{\theta}) f(\mathbf{s} | \mathbf{b}, \boldsymbol{\theta}), \quad (4)$$

where $f(\mathbf{y} | \mathbf{b}, \boldsymbol{\theta})$ and $f(\mathbf{s} | \mathbf{b}, \boldsymbol{\theta})$ are commonly specified according to the joint models (1) and (2).

From a joint approach, the inferential procedure to estimate $(\mathbf{b}, \boldsymbol{\theta})$ based on Equations (3) and (4) should be performed simultaneously. In addition, this joint modelling is usually quite complex due to the high number of parameters and potential integrations with no closed-form derived from the calculation of the survival function obtained from Equation (2). Hence, as expected, the processing of the inferential procedure is very time-consuming.

4. Standard Two-Stage (STS) Approach

Two-stage strategies are very useful for reducing the complexity of joint models and speeding up the inferential process. From a frequentist point of view, Tsiatis et al. [20] proposed one of the most popular two-stage approaches. The first stage is to fit the longitudinal submodel (1) and then the trajectory function $\mu_i(t)$ is calculated using the estimated parameters and random effects. In the second stage, this trajectory function estimated is considered as an endogenous time-varying covariate when fitting the survival submodel (2).

As a potential competitor, we use the Tsiatis et al. [20] approach adapted to the Bayesian framework. Specifically, in the first stage, we calculate the posterior mean of the longitudinal submodel parameters and random effects shared with the survival submodel, that is, $\hat{\boldsymbol{\beta}} = E(\boldsymbol{\beta} | \mathbf{y})$ and $\hat{\mathbf{b}} = E(\mathbf{b} | \mathbf{y})$. In the second stage, we incorporate the trajectory function into the survival submodel considering $\hat{\mu}_i(t) = \mathbf{x}_{L,i}^\top(t) \hat{\boldsymbol{\beta}} + \mathbf{z}_i^\top(t) \hat{\mathbf{b}}_i$, for $i = 1, \dots, n$, and then the posterior distribution of $(\boldsymbol{\gamma}, \alpha, h_0)$ is calculated.

5. Novel Two-Stage (NTS) Approach

The first part of our two-stage proposal is similar to the STS approach, that is, the posterior distributions of the longitudinal submodel parameters and random effects are calculated. However, we propose the following modification to the survival submodel:

$$h_i(t) = w_i h_0(t) \exp\left\{\mathbf{x}_{S,i}^\top \boldsymbol{\gamma} + \alpha \hat{\mu}_i(t)\right\}, \quad (5)$$

where $w_i > 0$ denotes a multiplicative fixed-effect for individual i and $\hat{\mu}_i(t)$ is calculated in the same way as the standard two-stage approach.

The role of w_i is essential to satisfactorily correct the estimation bias by ignoring the potential joint nature between both processes. In addition, this term can also correct problems of model misspecification and unobserved heterogeneity [39]. Specifically, what

we propose is a very small perturbation using an individual fixed-effect. Hence, to do that, we specify a highly informative prior distribution for w_i , given by:

$$w_i \sim \text{Gamma}(\eta, \eta), \quad (6)$$

where $E(w_i) = 1$ and η is a known parameter and must be specified such that $\text{Var}(w_i) = 1/\eta$ is small. Interpretatively, if w_i is not perturbed (i.e., $\text{Var}(w_i) = 0$), then we turn to the standard two-stage approach presented in Section 4. Moreover, note that if we assume that η is an unknown parameter and so a hyperprior should be set for it, then the specification (5) becomes a Bayesian frailty model [40]. In practice, the latter option is convergently unstable and therefore will not be addressed in this paper.

In the context of frailty models, w_i is typically modelled through a Gamma distribution [41]. For this reason, we chose such distribution in (6). However, other non-negative continuous distributions could be used as long as $E(w_i) = 1$ is satisfied.

6. Simulation Study

To evaluate whether the novel two-stage approach reduces the bias with low computational time, we perform a simulation study that compares our proposal with the joint specification (see Section 3) and standard two-stage (see Section 4) approaches.

The joint formulation that is considered here is based on submodels (1) and (2). In particular, the longitudinal specification for individual i at time t is given by:

$$\begin{aligned} y_i(t) &= \mu_i(t) + \epsilon_i(t) = \beta_0 + b_{0i} + (\beta_1 + b_{1i})t + \beta_2 x_i + \epsilon_i(t), \\ \mathbf{b}_i &= (b_{0i}, b_{1i})^\top \stackrel{\text{i.i.d.}}{\sim} \text{N}(\mathbf{0}, \Sigma) \quad \text{and} \quad \epsilon_i(t) \stackrel{\text{i.i.d.}}{\sim} \text{N}(0, \sigma^2), \end{aligned} \quad (7)$$

where the covariate x_i is a binary group indicator simulated from a Bernoulli distribution with probability 0.5 and will be called group parameter.

Based on the simulation scenarios proposed by Furgal et al. [8], we adopt the following hazard specification for individual i :

$$h_i(t) = \exp \{ \gamma_0 + \gamma_1 x_i + \alpha \mu_i(t) \}, \quad (8)$$

where the baseline hazard function has an exponential specification, $h_0(t) = \exp(\gamma_0)$. Note that other options for this function could be preferred, such as Gamma, Weibull, Gompertz, log-normal, log-logistic, piecewise, splines, and so forth [42,43].

6.1. Simulating Data for Joint Models

As a preliminary simulation step, all parameters and hyperparameters $\theta = (\beta, \Sigma, \sigma, \gamma, \alpha)$, number of individuals (n), minimum number of longitudinal observations (m_{\min}), and maximum observational time (t_{\max}) must be set. Then, the covariate x_i and the random effects \mathbf{b}_i , for $i = 1, \dots, n$, are simulated.

The true event time for individual i is simulated using the well-known inverse transform sampling [44], where $T_i^* = S_i(u)$, u is generated from a standard uniform distribution, and S_i denotes the survival function derived from Equation (8). The censoring time for each individual, C_i , is generated from a uniform distribution on the interval $(0, t_{\max})$ and then the observed time is set as $T_i = \min\{T_i^*, C_i\}$ and the event indicator as $\delta_i = I(T_i^* \leq C_i)$.

The number of longitudinal observations of individual i , n_i , is set as m_{\min} plus the largest integer less than T_i (i.e., $\lfloor T_i \rfloor$). The recording times of the repeated measurements are equispaced set from 0 to $\lfloor T_i \rfloor$. The random errors $\epsilon_i(t_1), \dots, \epsilon_i(t_{n_i})$ are simulated from a normal distribution with mean at zero and variance σ^2 . Finally, the longitudinal observations of individual i , $y_i(t_1), \dots, y_i(t_{n_i})$, are computed according to the submodel (7).

The simulation scheme to jointly generate longitudinal and survival data is summarised in Algorithm 1.

Algorithm 1: Simulation scheme

-
- 0 INITIALISATION: Set θ , n , m_{\min} , and t_{\max} .
 - 1 SURVIVAL SUBMODEL:
 - Simulate $x_i \sim \text{Bern}(0.5)$ and $\mathbf{b}_i \sim \text{N}(\mathbf{0}, \Sigma) \forall i$.
 - Simulate T_i^* based on the survival submodel (8) and sample $C_i \sim \text{U}(0, t_{\max}) \forall i$.
 - Set $T_i = \min\{T_i^*, C_i\}$ and $\delta_i = \mathbf{I}(T_i^* \leq C_i) \forall i$.
 - 2 LONGITUDINAL SUBMODEL:
 - Set $n_i = m_{\min} + \lfloor T_i \rfloor \forall i$.
 - Set $0 = t_1, \dots, t_{n_i} = \lfloor T_i \rfloor \forall i$ equispaced.
 - Simulate $\epsilon_i(t) \sim \text{N}(0, \sigma^2)$, $t = t_1, \dots, t_{n_i} \forall i$.
 - Compute $y_i(t_1), \dots, y_i(t_{n_i}) \forall i$ based on the longitudinal submodel (7).
-

6.2. Scenarios

We present simulation scenarios generated from the *prothro* dataset, which is available in the R-package JM (version 1.4-8) from the CRAN at <http://cran.r-project.org/package=JM>. This dataset includes 488 patients with histologically verified liver cirrhosis, where 251 patients were randomised to a treatment with prednisone and the remaining received placebo [45]. The longitudinal variable *pro* is used on a logarithmic scale and the treatment variable (*treat*) is defined as x_i in both submodels.

First, we fit the joint models (7) and (8) for *prothro* data using the function *jointModel* from the R-package JM. Then, the estimates are used as “true parameter values” in the generation of simulated data, also using the joint formulations (7) and (8). The jointly estimated parameters are $\hat{\beta}_0 = 4.274$, $\hat{\beta}_1 = -0.004$, $\hat{\beta}_2 = -0.097$, $\hat{\sigma} = 0.262$, $\hat{\Sigma}_{11} = 0.094$, $\hat{\Sigma}_{22} = 0.005$, and $\hat{\Sigma}_{12} = \hat{\Sigma}_{21} = -0.001$ for the longitudinal submodel (7); and $\hat{\gamma}_0 = 8.671$, $\hat{\gamma}_1 = -0.172$, and $\hat{a} = -2.447$ for the survival submodel (8). Finally, we simulate 100 datasets with $n=200, 500, 1000$, $m_{\min} = 3$, and $t_{\max} = 15$. Other simulation scenarios are presented in Appendix B.

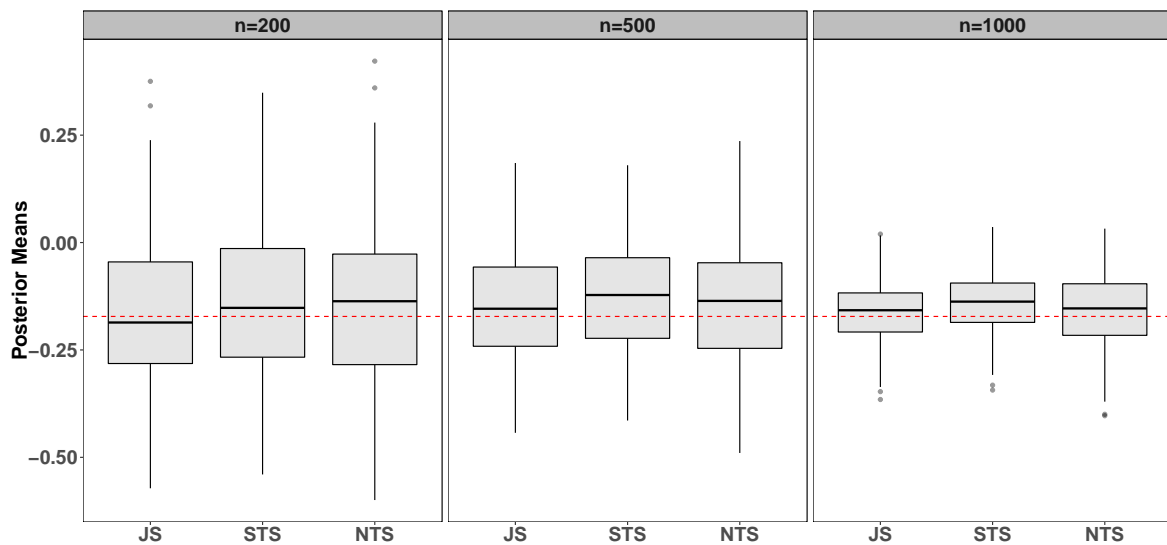
The Bayesian joint model specification (7) and (8) with the prior distributions presented in Section 2.3 is used for the three estimation strategies. The MCMC configuration is defined as follows: 2000 iterations with warm-up of 1000 for the joint model using the JS approach and for the longitudinal submodel from both two-stage approaches. Additionally, 1000 iterations with warm-up of 500 are set to run the survival submodel from both two-stage approaches. All models were implemented using *rstan* (<http://mc-stan.org>) and the codes are provided in a Supplementary Material. Simulations were performed on a Dell laptop with 2.6 GHz Intel Core i7, 16 GB RAM, OS Windows.

Here, the η parameter is set to 1.5 and so the prior variance of w_i is equal to $1/1.5 \approx 0.67$. Of course, this variance value is small and informative for the scale of simulated data in this paper, but it can still be very large for other problems. A sensitivity analysis for the choice of η is presented in Appendix A.

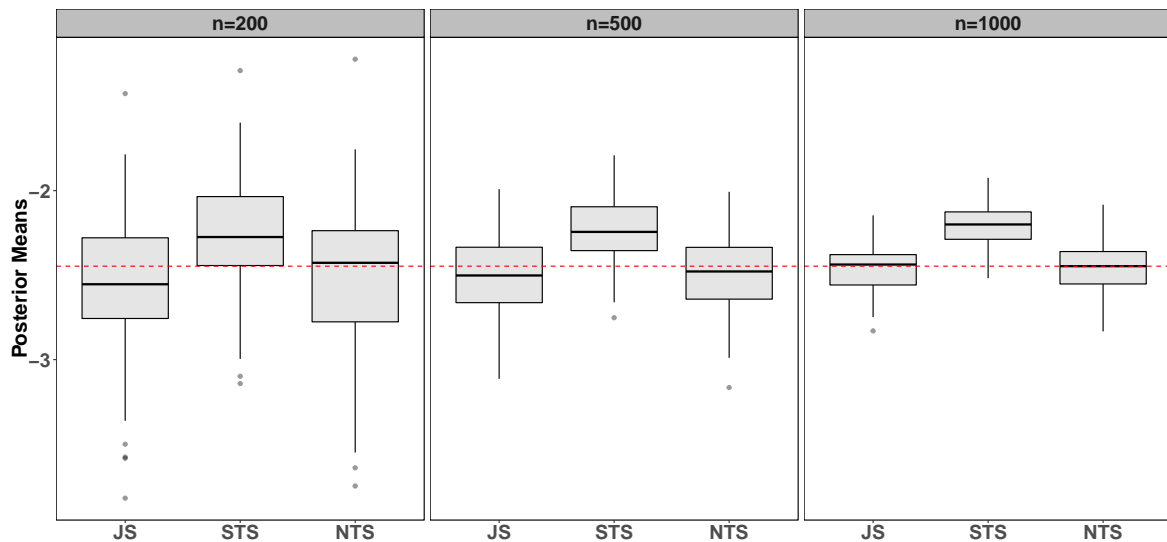
Table 1 and Figure 1 show the comparative results among joint specification (JS), standard two-stage (STS), and novel two-stage (NTS) approaches for 100 simulated datasets from the joint models (7) and (8) using the parameters set above.

Table 1. Posterior summary and computational time (in minutes) from each estimation approach.

Posterior	Parameter (True Value)	JS			STS			NTS		
		<i>n</i> = 200	<i>n</i> = 500	<i>n</i> = 1000	<i>n</i> = 200	<i>n</i> = 500	<i>n</i> = 1000	<i>n</i> = 200	<i>n</i> = 500	<i>n</i> = 1000
Mean	γ_1 (−0.172)	−0.172	−0.152	−0.165	−0.148	−0.132	−0.144	−0.141	−0.148	−0.160
	α (−2.447)	−2.574	−2.505	−2.460	−2.271	−2.237	−2.204	−2.537	−2.491	−2.456
SD	γ_1	0.193	0.120	0.084	0.184	0.115	0.082	0.241	0.149	0.106
	α	0.345	0.214	0.148	0.295	0.186	0.131	0.391	0.245	0.171
Average Comp. Time		4.291	11.055	23.428	1.885	4.692	11.906	2.224	5.836	13.845



(a) Group parameter (γ_1).



(b) Association parameter (α).

Figure 1. Simulation results from 100 datasets comparing the joint specification (JS), the standard two-stage (STS), and the novel two-stage (NTS) for *n* = 200, 500, 1000. The panels show posterior means from the 100 datasets for the survival submodel group (a) and association (b) parameters. The dashed horizontal line indicates the true parameter value.

We can see, both in Table 1 and in Figure 1, that the group parameter (γ_1) is satisfactorily estimated using the three approaches. On the other hand, in all scenarios, our approach also estimates the association parameter (α) very well. These results are better than the

STS approach and similar to the JS, which in theory is the correct way to deal with the estimation process. However, as expected, the standard deviation of posterior distributions using our methodology is slightly higher than others. Furthermore, the computational time of the NTS is a little higher than of the STS and much less than that of the JS approach.

It is worth noting that theoretically the joint specification approach is always preferable. The other two approaches are recommended when the complexity of the joint model makes the inferential procedure highly time-consuming or when there are problems of convergence of the Markov chains due to the high-dimensional parameter space. In the model selection framework (e.g., variable selection problems or model selection from different hazard function proposals), Bayesian selection criteria can be applied in the usual way. In particular, leave-one-out cross-validation (LOO) and the widely applicable information criterion (WAIC) can be easily calculated using the R-package `loo` [46] as well as Bayes factors and posterior model probabilities using the R-package `bridgesampling` [47].

7. Discussion

In this paper, we presented a novel two-stage (NTS) method for fitting Bayesian joint models of longitudinal and survival data using fixed-effects with informative prior to correct the estimation bias caused by ignoring the joint nature of both processes. We demonstrated in different scenarios that our proposal is more accurate than the standard two-stage (STS) approach and its processing time is much less than the joint specification (JS) approach.

In our simulation studies, we found that the group parameter estimation from the survival submodel is robustly estimated regardless of the estimation approach. This result was expected since this parameter does not depend on shared information. On the other hand, the association parameter is sensitive when using the STS strategy.

The specification of the informative prior variance for the fixed-effects can be critical drawback of our approach. In our simulation study, the set value produced quite satisfactory results (see the sensitivity analysis of this parameter in Appendix A). However, we would like to reinforce to the reader that this choice depends on the scale of the problem, in which the value used in this paper may not be appropriate in other applications.

It would be interesting to apply the NTS in more complex longitudinal (e.g., skewed or multiple longitudinal data) and survival (e.g., competing-risks or multistate data) submodels than those employed here. Hence, we would be able to try determining the limits of the methodology. Furthermore, our proposal could also be combined with sequential methods for Bayesian joint models [48].

Supplementary Materials: The codes are available online at <https://www.mdpi.com/1099-4300/23/1/50/s1>.

Author Contributions: All authors contributed equally to this paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Fund for Scientific and Technological Development (FONDECYT, Chile) grant number 11190018.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Sensitivity Analysis for η

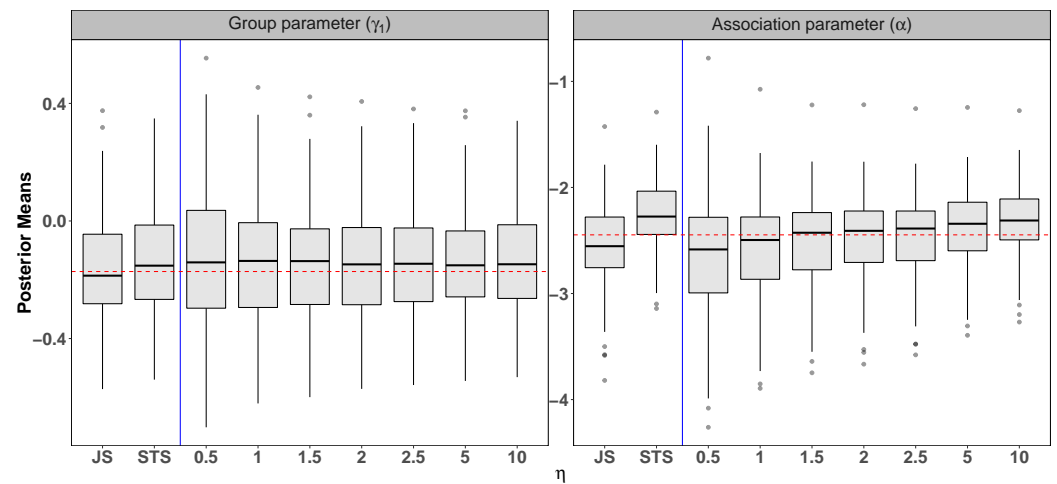


Figure A1. Sensitivity analysis for η from 100 datasets with $n = 200$. Since $\text{Var}(w_i) = 1/\eta$ (for the Gamma prior), larger values of η lead to small perturbation of w_i 's and therefore the results resemble the standard two-stage (STS) approach. The dashed horizontal line indicates the true parameter value.

Appendix B. Other Simulation Studies

Appendix B.1. Scenario 1

Setting: $\beta_0 = 5, \beta_1 = -0.02, \beta_2 = -0.1, \sigma = 0.25, \Sigma_{11} = 1, \Sigma_{22} = 1, \Sigma_{12} = \Sigma_{21} = 0.2, \gamma_0 = 5, \gamma_1 = -0.1, \alpha = -1, n = 200, m_{\min} = 3, \text{ and } t_{\max} = 15$.

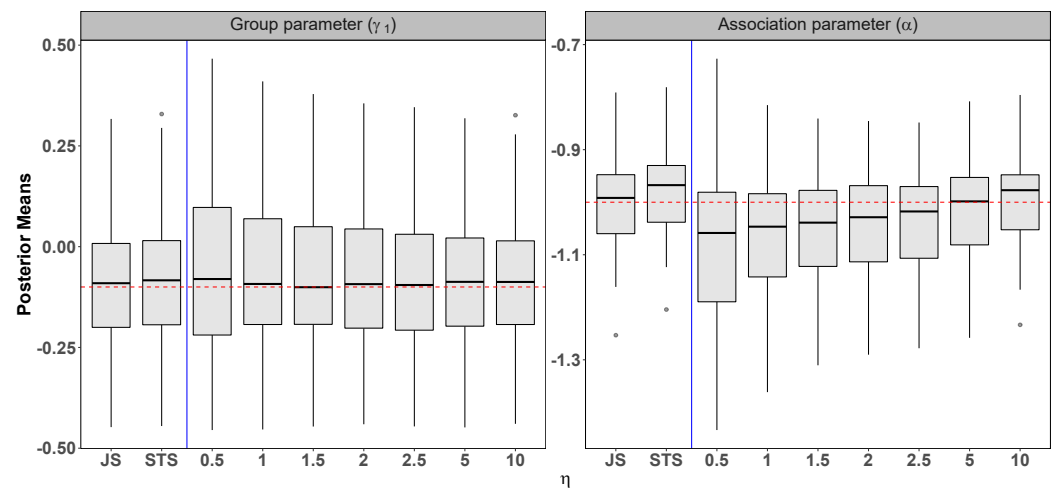


Figure A2. Scenario 1: Simulation results from 100 datasets comparing the joint specification (JS), the standard two-stage (STS), and the novel two-stage (NTS) for $\eta = 0.5, 1, 1.5, 2, 2.5, 5, 10$. The dashed horizontal line indicates the true parameter value.

Appendix B.2. Scenario 2

Setting: $\beta_0 = 1, \beta_1 = -0.1, \beta_2 = -0.01, \sigma = 0.5, \Sigma_{11} = 1, \Sigma_{22} = 1, \Sigma_{12} = \Sigma_{21} = 0.2, \gamma_0 = -3, \gamma_1 = 0.1, \alpha = 1, n = 200, m_{\min} = 3, \text{ and } t_{\max} = 15$.

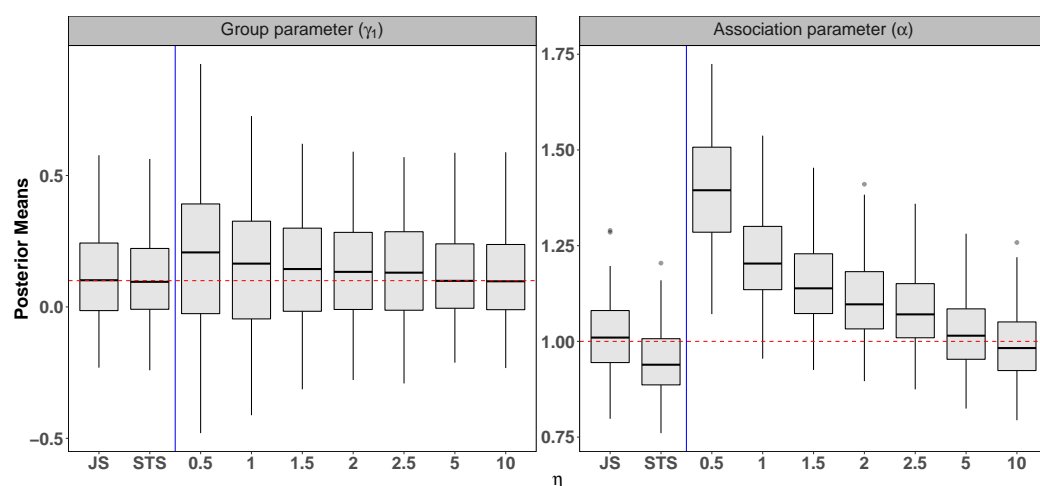


Figure A3. Scenario 2: Simulation results from 100 datasets comparing the joint specification (JS), the standard two-stage (STS), and the novel two-stage (NTS) for $\eta = 0.5, 1, 1.5, 2, 2.5, 5, 10$. The dashed horizontal line indicates the true parameter value.

References

- Rizopoulos, D.; Verbeke, G.; Molenberghs, G. Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival outcomes. *Biometrics* **2010**, *66*, 20–29. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wu, L.; Wei, L.; Yi, G.Y.; Huang, Y. Analysis of longitudinal and survival data: Joint modeling, inference methods, and issues. *J. Probab. Stat.* **2011**, *2012*, 1–17. [\[CrossRef\]](#)
- Muthén, B.; Asparouhov, T.; Boye, M.E.; Hackshaw, M.; Naegeli, A. *Applications of Continuous-Time Survival in Latent Variable Models for the Analysis of Oncology Randomized Clinical Trial Data Using Mplus*; Technical Report; Muthén & Muthén: Los Angeles, CA, USA, 2009.
- Ibrahim, J.G.; Chu, H.; Chen, L.M. Basic concepts and methods for joint models of longitudinal and survival data. *J. Clin. Oncol.* **2010**, *28*, 2796–2801. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, P.; Shen, W.; Boye, M.E. Joint modeling of longitudinal outcomes and survival using latent growth modeling approach in a mesothelioma trial. *Health Serv. Outcomes Res. Methodol.* **2012**, *12*, 182–199. [\[CrossRef\]](#) [\[PubMed\]](#)
- Elashoff, R.; Li, G.; Li, N. *Joint Modeling of Longitudinal and Time-to-Event Data*, 1st ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2016.
- Papageorgiou, G.; Mauff, K.; Tomer, A.; Rizopoulos, D. An overview of joint modeling of time-to-event and longitudinal outcomes. *Annu. Rev. Stat. Its Appl.* **2019**, *6*, 223–240. [\[CrossRef\]](#)
- Furgal, K.C.; Sen, A.; Taylor, J.M.G. Review and comparison of computational approaches for joint longitudinal and time-to-event models. *Int. Stat. Rev.* **2019**, *87*, 393–418. [\[CrossRef\]](#) [\[PubMed\]](#)
- Alsefiri, M.; Sudell, M.; García-Fiñana, M.; Kolamunnage-Dona, R. Bayesian joint modelling of longitudinal and time to event data: A methodological review. *BMC Med. Res. Methodol.* **2020**, *20*, 1–17. [\[CrossRef\]](#)
- Henderson, R.; Diggle, P.; Dobson, A. Joint modelling of longitudinal measurements and event time data. *Biostatistics* **2000**, *1*, 465–480. [\[CrossRef\]](#)
- Wu, L. *Mixed Effects Models for Complex Data*, 1st ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2009.
- Gould, A.L.; Boye, M.E.; Crowther, M.J.; Ibrahim, J.G.; Quartey, G.; Micallef, S.; Bois, F.Y. Joint modeling of survival and longitudinal non-survival data: Current methods and issues. Report of the DIA Bayesian joint modeling working group. *Stat. Med.* **2015**, *34*, 2181–2195. [\[CrossRef\]](#)
- Wu, L.; Yu, T. Joint modeling of longitudinal and survival data. In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons: Hoboken, NJ, USA, 2016; pp. 1–9.
- Lesaffre, E.; Spiessens, B. On the effect of the number of quadrature points in a logistic random effects model: An example. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **2001**, *50*, 325–335. [\[CrossRef\]](#)
- Pinheiro, J.C.; Chao, E.C. Efficient Laplacian and adaptive Gaussian quadrature algorithms for multilevel generalized linear mixed models. *J. Comput. Graph. Stat.* **2006**, *15*, 58–81. [\[CrossRef\]](#)
- Rizopoulos, D.; Verbeke, G.; Lesaffre, E. Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2009**, *71*, 637–654. [\[CrossRef\]](#)
- Wu, L.; Liu, W.; Hu, X.J. Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics* **2010**, *66*, 327–335. [\[CrossRef\]](#) [\[PubMed\]](#)

18. Barrett, J.; Diggle, P.; Henderson, R.; Taylor-Robinson, D. Joint modelling of repeated measurements and time-to-event outcomes: Flexible model specification and exact likelihood inference. *J. R. Stat. Soc. Ser. B (Methodol.)* **2015**, *77*, 131–148. [[CrossRef](#)] [[PubMed](#)]
19. Self, S.; Pawitan, Y. Modeling a marker of disease progression and onset of disease. In *AIDS Epidemiology*; Springer: Berlin, Germany, 1992; pp. 231–255.
20. Tsiatis, A.A.; DeGruttola, V.; Wulfsohn, M.S. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *J. Am. Stat. Assoc.* **1995**, *90*, 27–37. [[CrossRef](#)]
21. Wulfsohn, M.S.; Tsiatis, A.A. A joint model for survival and longitudinal data measured with error. *Biometrics* **1997**, *53*, 330–339. [[CrossRef](#)] [[PubMed](#)]
22. Ye, W.; Lin, X.; Taylor, J.M.G. Semiparametric modeling of longitudinal measurements and time-to-event data—A two-stage regression calibration approach. *Biometrics* **2008**, *64*, 1238–1246. [[CrossRef](#)]
23. Albert, P.S.; Shih, J.H. On estimating the relationship between longitudinal measurements and time-to-event data using a simple two-stage procedure. *Biometrics* **2010**, *66*, 983–987. [[CrossRef](#)]
24. Huong, P.T.T.; Nur, D.; Pham, H.; Branford, A. A modified two-stage approach for joint modelling of longitudinal and time-to-event data. *J. Stat. Comput. Simul.* **2018**, *88*, 3379–3398. [[CrossRef](#)]
25. Murawska, M.; Rizopoulos, D.; Lesaffre, E. A two-stage joint model for nonlinear longitudinal response and a time-to-event with application in transplantation studies. *J. Probab. Stat.* **2012**, *2012*, 1–18. [[CrossRef](#)]
26. Viviani, S.; Alfó, M.; Rizopoulos, D. Generalized linear mixed joint model for longitudinal and survival outcomes. *Stat. Comput.* **2014**, *24*, 417–427. [[CrossRef](#)]
27. Mauff, K.; Steyerberg, E.; Kardys, I.; Boersma, E.; Rizopoulos, D. Joint models with multiple longitudinal outcomes and a time-to-event outcome: A corrected two-stage approach. *Stat. Comput.* **2020**, *30*, 999–1014. [[CrossRef](#)]
28. Faucett, C.L.; Thomas, D.C. Simultaneously modelling censored survival data and repeatedly measured covariates: A Gibbs sampling approach. *Stat. Med.* **1996**, *15*, 1663–1685. [[CrossRef](#)]
29. Tsiatis, A.A.; Davidian, M. Joint modeling of longitudinal and time-to-event data: An overview. *Stat. Sin.* **2004**, *14*, 809–834.
30. Rizopoulos, D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*, 1st ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2012.
31. Verbeke, G. Linear mixed models for longitudinal data. In *Linear Mixed Models in Practice*; Springer: Berlin, Germany, 1997; pp. 63–153.
32. Pinheiro, J.C.; Bates, D.M. Linear mixed-effects models: Basic concepts and examples. In *Mixed-Effects Models in S and S-Plus*; Springer: Berlin, Germany, 2000; pp. 3–56.
33. Kumar, D.; Klefsjö, B. Proportional hazards model: A review. *Reliab. Eng. Syst. Saf.* **1994**, *44*, 177–188. [[CrossRef](#)]
34. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2013.
35. Gelman, A. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.* **2006**, *1*, 515–534. [[CrossRef](#)]
36. Schuurman, N.K.; Grasman, R.P.P.P.; Hamaker, E.L. A comparison of inverse-Wishart prior specifications for covariance matrices in multilevel autoregressive models. *Multivar. Behav. Res.* **2016**, *51*, 185–206. [[CrossRef](#)]
37. Alvares, D. *Sequential Monte Carlo Methods in Bayesian Joint Models for Longitudinal and Time-to-Event Data*. Ph.D. Thesis, University of Valencia, Valencia, Spain, 2017.
38. Wu, M.C.; Carroll, R.J. Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* **1988**, *44*, 175–188. [[CrossRef](#)]
39. Wienke, A. *Frailty Models in Survival Analysis*, 1st ed.; Chapman & Hall/CRC: Boca Raton, FL, USA, 2010.
40. Ibrahim, J.G.; Chen, M.H.; Sinha, D. *Bayesian Survival Analysis*, 1st ed.; Springer: Berlin, Germany, 2001.
41. Balan, T.A.; Putter, H. A tutorial on frailty models. *Stat. Methods Med Res.* **2020**, *29*, 3424–3454. [[CrossRef](#)]
42. Lee, E.T.; Go, O.T. Survival analysis in public health research. *Annu. Rev. Public Health* **1997**, *18*, 105–134. [[CrossRef](#)]
43. Lázaro, E.; Armero, C.; Alvares, D. Bayesian regularization for flexible baseline hazard functions in Cox survival models. *Biom. J.* **2020**. [[CrossRef](#)] [[PubMed](#)]
44. Crowther, M.J.; Lambert, P.C. Simulating biologically plausible complex survival data. *Stat. Med.* **2013**, *32*, 4118–4134. [[CrossRef](#)] [[PubMed](#)]
45. Andersen, P.K.; Borgan, O.; Gill, R.D.; Keiding, N. *Statistical Models Based on Counting Processes*, 1st ed.; Springer: Berlin, Germany, 1993.
46. Vehtari, A.; Gelman, A.; Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **2016**, *27*, 1413–1432. [[CrossRef](#)]
47. Gronau, Q.F.; Singmann, H.; Wagenmakers, E.J. bridgesampling: An R package for estimating normalizing constants. *J. Stat. Softw.* **2020**, *92*, 1–29. [[CrossRef](#)]
48. Alvares, D.; Armero, C.; Forte, A.; Chopin, N. Sequential Monte Carlo methods in Bayesian joint models for longitudinal and time-to-event data. *Stat. Model.* **2020**. [[CrossRef](#)]