Filomat 38:23 (2024), 8265–8281
https://doi.org/10.2298/FIL2423265Y

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: `http://www.pmf.ni.ac.rs/filomat`

A two-stage approach for Bayesian parametric joint modelling of longitudinal and Weibull accelerated failure time data

Elif Yıldırım[a,b,*], **Duru Karasoy**[c]

[a] *Graduate School of Science and Engineering, Hacettepe University, Beytepe, Ankara, Turkey*
[b] *Department of Statistics and Quality Coordinator, Konya Technical University, Konya, Turkey*
[c] *Department of Statistics, Hacettepe University, Beytepe, Ankara 06800, Turkey*

Abstract. Since longitudinal and survival data are often obtained together in applications, studies on joint modelling that reveal the relationship between these two data have increased considerably in recent years. These models are generally defined by a linear mixed-effects model for longitudinal data and Cox regression model for survival data, which are connected by a shared random effect. However, in order to use the Cox regression model in the analysis of survival data, the proportional hazards assumption must be satisfied. In cases where the proportional hazards assumption is not satisfied and survival data have a certain distribution, parametric joint models should be used to obtain more unbiased parameter estimates in the analysis of the relationship between two data. In this paper, we propose a two-stage approach in a Bayesian framework to obtain parameter estimates in parametric joint modelling. To examine the performance of the proposed method, we perform a simulation study in scenarios with different censoring and sample sizes and compare the method with classical approaches. In addition, to demonstrate the applicability of the proposed method, we perform an application on the aortic valve replacement surgery data, which is frequently used in the literature, and test the methods on real data. Simulation studies from different scenarios show that in all cases our approach, which consists of a longitudinal measurement and of survival data that does not require the assumption of proportional hazards, gives more unbiased estimates compared to the other two classical approaches. As a result of the application of the real data set to demonstrate the applicability of the simulation findings, we obtained results similar to the simulation results according to the AIC and BIC.

1. Introduction

Longitudinal models are frequently used, especially in medical research, to determine the associations of disease progression. Generally, in this type of models, the effects of strong biomarkers on the disease or treatment are investigated within the survival time of patient. These types of data sets include both survival information of individuals and information from longitudinal markers obtained repeatedly over time. In such studies, the effects of repeated measurements on the survival time of individuals, the changes in patients whose longitudinal measurements cannot be followed, and the relationships between biomarkers

---

2020 *Mathematics Subject Classification*. Primary 62N02; Secondary 62F15.
Keywords. Weibull accelerated failure time model; Longitudinal data; Parametric joint modelling; Bayesian two-stage approach.
Received: 31 January 2024; Revised: 05 April 2024; Accepted: 26 April 2024
Communicated by Aleksandar Nastić

* Corresponding author: Elif Yıldırım
Email addresses: `elifyildirim3123@gmail.com` (Elif Yıldırım), `durdu@hacettepe.edu.tr` (Duru Karasoy)

and the time until a certain event occurs are investigated. In order to obtain the complete inference of the relationship between these two data and the appropriate estimates unbiasedly, joint models were developed in which longitudinal models and survival time models were connected by common latent random effects.

Joint modelling consists of a longitudinal sub-model and survival sub-model, and different random effect structures are used to combine these two sub-models. Although the linear mixed effect model (LMM) for the longitudinal sub-model and the Cox regression model for the survival sub-model are generally used in the literature, studies on different joint models have increased with the increasing interest in the last twenty years. The first studies based on the two-stage approach for joint modelling of univariate survival and longitudinal data were done by Self and Pawitan [32], DeGruttola and Tu [9], Tsiatis, DeGruttola and Wulfsohn [36], Faucett and Thomas [12] and Wulfsohn and Tsiatis [40]. Elashoff et al. [23], Williamson et al. [37] and Hu et al. [18] studied on joint modelling in case of competing risks and multiple failures. Albert and Shih [1] proposed a new joint model for the combined modelling of multiple longitudinal observations and discrete survival data. Su and Wang [34] examined the joint modelling of left truncation survival data and longitudinal observations. In their study, they showed that the joint model methods available in the literature have low performance in the presence of left-truncated data and proposed a new likelihood approach to obtain joint model parameter estimates from the presence of such data. Efendi et al. [10] proposed a joint model that combines continuous and binary longitudinal observations with survival data. Tang and Tang [35] proposed a semi-parametric multivariate skew-normal joint model for multivariate survival and longitudinal data. Hickey et al. [17] developed the "*JoineRML*" package in the R program to obtain parameter estimates in the joint modelling of survival data with multivariate longitudinal observations. Mwanyekange et al. [26] proposed the Bayesian joint model approach for longitudinal data and multi-state survival data. In their proposed model, they used a LMM for the longitudinal sub-model and a proportional density function for the multi-state process as the survival analysis sub-model, and obtained the parameter estimates with the Bayesian approach. Yıldırım and Karasoy [42] developed the "*gsem*" code in the Stata program for the joint modelling of LMM and parametric survival data. Alsefri et al. [3] examined the Bayesian estimation methods used in joint modelling of longitudinal and survival data in the literature. They examined 89 articles consisting of different model structures, model assumptions, estimation methods and dynamic estimates. They stated that the most commonly used joint model in the literature is the linear mixed effect model for longitudinal data and the Cox regression model for survival data. They also found that the Markov Chain Monte Carlo (MCMC) algorithm was used for parameter estimates in Bayesian joint models. Furgal [15] used the Bayesian approach in the joint modelling of survival data containing different competing risk situations and longitudinal data. Additionally, they conducted a simulation study to examine joint model structures consisting of different random effects structures. Xu [41] developed a multivariate mixed-effects joint modelling for survival data consist of skew distribution and skewed-longitudinal. Alkhathami [2] proposed a new joint model structure for the LMM and the Weibull Frailty model. They used random effects structures with different distributions to examine the parameter estimates of the model he developed.

Due to the complex structure of joint models, another focus of researchers has been to obtain parameter estimates. Joint model parameter estimates are generally obtained by the maximum likelihood approach based on maximizing the likelihood function. However, the likelihood function of the joint model has a complex structure and high-dimension integrals in the likelihood function must be taken to obtain parameter estimates with the maximum likelihood approach. Due to the complex structure of the likelihood function and the lack of a closed solution, different parameter estimation approaches have been proposed in the literature. Wulfsohn and Tsiatis [40] simultaneously obtained the parameter estimates of the joint model for longitudinal and survival data by using the full likelihood approach in parameter estimates. Faucett and Thomas [12] obtained joint model parameter estimates using the MCMC method and Gibbs sampling. Song et al. [33] used the EM algorithm approach to obtain joint model parameter estimates. Li et al. [23] proposed joint modelling consisting of competing risks and longitudinal ordinal measurements. They used the EM algorithm to obtain parameter estimates of the joint model they created. Wu et al. [38] used the Laplace approach and Monte Carlo Expectation Maximization (MCEM) methods in joint model parameter estimates and compared the performance of the two methods with real data and simulation study. Huang et al. [19] worked on a model that formulates model errors with skewed-normal distributions for joint modelling of longitudinal dynamical processes and time-to-event data.To obtain the parameter estimates of the model they proposed, they used a Bayesian approach that focuses on the time-to-event process associated with response, covariate, and random effects. Rizopoulos [30] proposed an integration method based on rescaling the adaptive Gauss-Hermite (AGH) approach as an alternative for solving the joint modelling likelihood. Baghfalaki [8] proposed a new model for joint modelling of multivariate longitudinal data containing mixed continuous and ordinal responses and the time-to-event variable. They modeled the relationship structure between longitudinal mixed data and time-to-event data with a multivariate zero-mean Gaussian process. In the joint model, they used the Gaussian mixed effect model for longitudinal data and the accelerated failure time model with log-normal and Weibull distribution for survival data. They considered a Bayesian approach using the MCMC algorithm to obtain parameter estimates. Rizopoulos [31] developed the "*JMBayes*" package in the R program, which obtains Bayesian estimates using the Monte Carlo algorithm for joint model parameter estimation. Dagne [8] proposed a joint model consisting of a bent-cable Tobit model for longitudinal data and an accelerated failure time model for survival data. A Bayesian approach with multivariate skew-t distributions is proposed to obtain parameter estimates of this model. Furgal et al. [14] conducted a simulation study to compare the performance of statistical programming packages developed for the classical approaches and the Bayesian approach used to obtain parameter estimates in the joint model. Atlı [4] compared the classical approach with the Bayesian approach to obtain parameter estimates in the joint modelling of the LMM and the Cox regression model. Random effect structures with different distributions were used to examine the parameter estimates of the model developed in the study. Leiva-Yamaguchi and Alvares [22] proposed a new two-stage approach to obtain more unbiased parameter estimates in joint modelling of longitudinal and survival data. They compared their proposed approach with the joint specification approach and standard two-stage approaches in the Bayesian framework and examined the performance of the methods. Lin and Luo [24] proposed a new artificial neural network architecture to obtain joint model parameter estimates of survival data and multiple longitudinal data and examined the performance of their proposed approach with a simulation study. Murray and Philipson [25] proposed an alternative EM algorithm approach, addressing the difficulty in obtaining parameter estimates with classical approaches in multidimensional joint modelling of survival data with one or more longitudinal data. Khan and Basharat [21] worked on a joint model consisting of a linear mixed effect model for longitudinal data and an accelerated failure time model for survival data. They proposed a Bayesian two-stage approach based on the full likelihood approach to obtain parameter estimates for the joint model consisting of log-logistic and Weibull sub-models, and examined the performance of the two models on simulation study and real data application.

In this paper, we examine parameter estimation approaches for the parametric joint modelling consist of LMM for longitudinal data and Weibull accelerated failure time (AFT), which can be used when the proportional hazards assumption is not satisfied for survival data. To obtain parametric joint modelling parameter estimates, we proposed a two-stage approach using the parameters estimated in the Weibull AFT sub-model after estimating the longitudinal sub-model under the Bayesian approach. We examined the parameter estimation performance of the proposed approach with the Gauss-Hermite (GH) and AGH approaches available in the literature through simulation studies consisting of the different sample sizes and censoring rate. In this paper, unlike the literature, we proposed a new Bayesian two-stage approach on the joint modelling of Weibull AFT and LMM models and examined the performance of the developed method with classical approaches such as AGH and GH at different sample sizes and censoring rates.

The paper is organized as follow: Section 2 describes a general formulation of parametric joint models. Sections 3 introduces the parameter estimation methods for parametric joint modelling. Section 4 presents simulation study and compares the performance of our proposal against the other standard approaches. Section 5 presents a real data analysis. Finally, Section 6 discusses the advantages and extensions of our methodology.

2. The Joint Model

We assume that there are n individuals with repeated observation values and an associated event of interest. We also assume the event time does not satisfy the assumption of proportional hazards and has a

Weibull distribution. We will first define the standard parametric joint model structure and then define the parametric joint model structure consisting of the Weibull AFT model for the survival sub-model.

2.1. Longitudinal Sub-Model

To model repeated measures over time for the i-th subject, we use LMM that allow for both fixed effects and subject-specific random effects. We can express the response variable $y_i\left(t\right)$ for subject i at time t as follows:

$$y_i\left(t\right)=x_i^{'}\left(t\right)\beta+z_i^{'}\left(t\right)b_i+\varepsilon_i\left(t\right)$$

$$m_i\left(t\right)=z_i^{'}\left(t\right)b_i \tag{1}$$

where the $x_i'\left(t\right)$ and $z_i'\left(t\right)$ are the design vectors for fixed effects β and random effects b_i, respectively, and $m_i\left(t\right)$ refers to linear combinations between the random effect. The random effects b_i have a KxK variance-covariance matrix G, where K is the number of random effects and a multivariate normal distribution with zero mean. $\epsilon_i\left(t\right)$ is the normally distributed measurement error term with mean 0 and variance σ^2 and is assumed to be independent of each other ($cov\left(\epsilon_i,\epsilon_j\right)=0$ for $i=j$). Moreover, since there is $cov\left(b_i,\epsilon_j\right)=0$ for $\forall i,j$, the error terms and random effects are assumed to be independent of each other [13].

2.2. AFT Sub-Model

Let T_i denote the observed survival time for the i-th subject ($i=1,2,\ldots,n$), T_i^* the true survival time, and C_i the censoring time. Accordingly, we can write $T_i=\min(T_i^*,C_i)$, the event indicator $\delta_i=I(T_i^*\leq C_i)$ and the observed survival time and the event indicator for the i-th subject as (T_i,δ_i). So, we can express the AFT survival sub-model with repeated measurements over as follows [6], [21]:

$$h_i(t|\,\mathrm{M}_i(t)\,,x_i)=h_0\left(g_i\left(t\right)\right)exp\left\{-\mathrm{x}_{\mathrm{i}}^{'}\gamma-\alpha m_i(t)\right\},t>0 \tag{2}$$

where $M_i\left(t\right)=\left\{m_i\left(s\right),0\leq s\leq t\right\}$ is the history of the set of $m_i(t)$ values of the follow-up until t for i-th subject and $g_i\left(t\right)=\int_0^{t_i}\exp\left(-x_i'\gamma-\alpha m_i\left(u\right)\right)du$. x_i' shows baseline covariates with coefficients γ and $h_0(.)$ denotes the baseline hazard function. α denotes the effect of the association. If $\alpha=0$, there is no association between longitudinal and survival data. In such cases, it is recommended to employ separate standard models, including a LMM and survival models, instead of a joint model.

Additionally, we can write the probability density function and survival function for AFT models as follows, respectively:

$$f_i(t|\,\mathrm{M}_i(t)\,,x_i)=f_0\left(g_i\left(t\right)\right)exp\left\{-\mathrm{x}_{\mathrm{i}}^{'}\gamma-\alpha m_i(t)\right\},t>0,$$

$$S_i(t|\,\mathrm{M}_i(t)\,,x_i)=\mathrm{S}_0\left(g_i\left(t\right)\right),t>0, \tag{3}$$

It is very important to choose the basic hazard function $h_0(t)$ to analyze the survival sub-model given in Equation 2. In cases where the proportional hazards assumption, which assumes that the hazard is constant over time, is provided, the Cox regression model can be chosen as the hazard function. However, AFT models should be used when the PH assumption is not satisfied, and the hazard function has a certain distribution. There are different AFT models that can be used in the joint model which consists of a longitudinal measurement and of survival data that does not require the assumption of PH. These models can be listed as exponential, log-logistic, logistic, log-normal and gamma sub-models [21]. In this paper, we focused on the Weibull AFT sub-model since the situations where survival data do not satisfied the PH assumption and are suitable for the Weibull distribution will be discussed.

2.3. Weibull AFT Sub-Model

In this paper, we will study the Weibull AFT sub-model for the joint modelling of longitudinal, and survival data that does not provide the PH assumption and has a Weibull distribution. Therefore, we will first redefine the hazard function for *i*-th subject defined in Equation 2 with κ and λ shape and scale parameters, respectively, as follows.

$$h_i(t_i) \; = \lambda\kappa\left(\lambda g_i(t_i)^{\kappa-1}\right)\exp\left(-\mathrm{x}_{\mathrm{i}}'\gamma-\alpha m_i(t)\right) \tag{4}$$

Similarly, we can obtain the survival functions and probability density distribution for the Weibull AFT sub-model, respectively, as follows.

$$S_i\left(t_i\right) = exp\;\left[-\left(\lambda g_i\left(t_i\right)\right)^{\kappa}\right],$$

$$f\left(\mathrm{t_i},\delta_i|\,\theta,b_i\right) = \left\{\lambda\kappa\left(\lambda g_i(t_i)^{\kappa-1}\right)\exp\left(-\mathrm{x}_{\mathrm{i}}'\gamma-\alpha m_i(t)\right)\right\}^{\delta_i} exp\;\left[-\left(\lambda g_i\left(t_i\right)\right)^{\kappa}\right] \tag{5}$$

where θ is full parameter vector and since $g_i\left(t\right) = \int_0^{t_i}\exp\left(-x_i'\gamma-\alpha z_i'\left(u\right)b_i\right)du$ does not have a closed formula, the Gauss-Kronrod or Gauss-Legendre quadrature rule can be used to calculate this integral [21].

3. Parameter Estimates for Parametric Joint Modelling

In joint modelling, two-stage approaches and joint likelihood methods are generally used to obtain parameter estimates [39]. We will propose a two-stage approach in the Bayesian perspective to obtain parameter estimates in parametric joint modelling and compare it with the likelihood approaches frequently used in the literature.

3.1. Gaussian Hermite and Adaptive Gaussian Hermite Approaches

The likelihood function for longitudinal and survival data described by time-independent random effects b_i for the i-th subject is,

$$logp\;\left(T_i,\delta_i,y_i;\theta\right) = log\int p\left(T_i,\delta_i,y_i,b_i;\theta\right)db_i \; = log\int p\left(T_i,\delta_i|b_i;\theta_t,\beta\right)\left[\prod_j p\left\{y_i(t_{ij})|b_i;\theta_y\right\}\right]p\left(b_i;\theta_b\right)db_i \tag{6}$$

Where $\theta = \left(\theta_t',\theta_y',\theta_b'\right)$ is the full parameter vector θ_t, θ_y and θ_b are survival data parameters, longitudinal data parameters and covariance matrix parameters of random effects, respectively. $p\left(T_i,\delta_i|b_i;\theta_t,\beta\right)$ is the conditional probability density function of the survival sub-model and $p\left(y_i|b_i;\theta\right)p\left(b_i;\theta\right)$ is the joint probability density function of longitudinal measurements and random effects. To obtain the parameter estimates of the joint model, researchers have proposed different integration methods such as Gauss Hermite, Laplace, and EM approach due to the complex structure of the likelihood function and the difficulty of calculation ([40]; [16]; [33]; [29]). Since our focus is on Gauss Hermite and Adaptive Gauss Hermite approaches, we only explained these two approaches in detail.

Let's rewrite the score vector of the joint likelihood in Equation 6 in the following form:

$$S(\theta) = \frac{\partial\ell(\theta)}{\partial\theta'} = \sum_i\frac{\partial}{\partial\theta'}log\int p\left(T_i,\delta_i|b_i;\theta\right)p\left(y_i|b_i;\theta\right)p\left(b_i;\theta\right)db_i = \sum_i\int A\left(\theta,b_i\right)p\left(b_i|T_i,\delta_i,y_i;\theta\right)db_i \tag{7}$$

Here, $A(\cdot)$, $A(\theta,b_i) = \partial\left\{\log p(T_i,\delta_i|b_i;\theta) + \log p(y_i|b_i;\theta) + \log p(b_i;\theta)\right\}/\partial\theta'$ refers to the score vector obtained from all data. Since the integrals do not have a closed form solution in the formulas of this vector, the integral solution can be made by the Gauss Hermite rule, which approximates the integral by the weighted sums of integrant measurements in predetermined abscissas [28]. Under the quadratic rule and random effects, the score vector for each form of A(·) is calculated as follows [30]:

$$E\left\{A\left(\theta,b_i\right)|T_i,\delta_i,y_i;\theta\right\}\int A\left(\theta,b_i\right)p\left(b_i|T_i,\delta_i,y_i;\theta\right)db_i$$

$$\approx 2^{\frac{q}{2}}\sum_{t_1,\dots,t_q}\pi_t A(\theta,b_t\sqrt{2})p(\,b_t\sqrt{2}\,|T_i,\,\delta_i,y_i;\,\theta)exp(\,\|b_t\|^2) \tag{8}$$

Since the Gauss-Hermite rule requires integrant calculations on the Cartesian product of abscissas for each random effect, the computational burden increases as the size of the q random effects vector increases.

The adaptive Gauss Hermite rule, developed to solve the problems in the Gauss Hermite rule, appropriately averages and measures the integrator in each iteration of the optimization algorithm:

$$E\left\{A\left(\theta,b_i\right)|T_i,\delta_i,y_i;\theta\right\}\approx 2^{\frac{q}{2}}\sum_{t_1,\dots,t_q}\pi_t A(\theta,\hat{r}_t\sqrt{2})p(\,\hat{r}_t\sqrt{2}|T_i,\,\delta_i,y_i;\,\theta)exp(\,\|\hat{r}_t\|^2) \tag{9}$$

where $\hat{r}_t = \hat{b}_i + \sqrt{2}\hat{B}_i^{-1}b_t$, $\hat{b}_i = argmax_b\left\{logp\ \left(T_i,\delta_i,y_i,b;\theta\ \right)\right\}$ and it represents the Choleski factor of $\hat{B}_i$, $\hat{H}_i = \frac{-\partial^2 logp\ \left(T_i,\delta_i,y_i,b;\theta\ \right)}{\partial b\partial b'}|_{b=\hat{b}_i}$.

In order to obtain the same size approximation error in the adapted Gauss-Hermite rule, fewer square points must be used than in the standard Gauss-Hermite rule. However, the need to locate the mode b_i and the second-order derivative matrix $\hat{H}_i$ for each subject in each iteration increases the processing load [30].

3.2. Two Stage Approach with Bayesian Framework

The two-stage approach is often used in parameter estimation of the joint model because of the complex structure in the likelihood of joint model and to speed up the estimation process. The two-stage approach was first proposed by Tsiatis et al. [36]. In this approach, in the first stage, longitudinal data are modeled with a LMM and subject-specific values of covariates are calculated, and in the second stage, the survival sub-model is calculated using the estimated values obtained in the first step. We used this two-stage approach developed by Tsiatis et al. [36] to obtain parametric joint modelling parameter estimates by adapting it for the Weibull AFT sub-model with a Bayesian perspective.

In the first stage, we estimated the parameters of the longitudinal sub-model using the LMM ($\widehat{\beta} = E\left(\beta|y\right)$) model and the random effects shared with the Weibull AFT sub-model ($\hat{b} = E\left(b|y\right)$). In the second stage, we estimate the posterior distribution of (γ,α,h_0) by using the $\hat{m}_i(t)$ estimates obtained from the first stage in the Weibull AFT sub-model. In addition, we used the prior distributions assigned by Leiva-Yamaguchi and Alvares [22] to determine parameters and hyperparameters in the Bayesian parametric joint model. Accordingly, we get $\beta \sim N\left(0,100\right)$, $\alpha,\ \gamma \sim N\left(0,10\right)$, $\sigma \sim half - Cauchy\left(0,5\right)$ and $\Sigma \sim inverse - Wishart(V,r)$ where V denotes KxK matrix and $r = K$ shows the degree of freedom parameter.

4. Simulation Study

To examine the performance of the two-stage approach proposed for the parameter estimate of parametric joint modelling, we compared our proposed approach with the GH and AGH methods frequently used in the literature for joint modelling parameter estimates. AGH and GH methods are frequentist approaches and are obtained by solving the likelihood given in Equation (6) with the integral approaches given in section 3.1. In this paper, we applied the AGH and GH methods with the integration approaches given in section 3.1 and the proposed two-stage approach using the prior distributions given in section 3.2. We first establish a parametric joint model for the i-th subject at time t using the sub-models in Equations (1) and (4) as follow.

$$y_i\left(t\right) = x_i^{'}\left(t\right)\beta + z_i^{'}\left(t\right)b_i + \varepsilon_i\left(t\right) = \beta_0 + b_{0i} + \left(\beta_1 + b_{1i}\right)t_{ij} + \beta_2 x_i + \epsilon_i(t)$$

$$h_i(t_i) \ = \lambda\kappa\left(\lambda g_i(t_i)^{\kappa-1}\right)\exp\left(-\mathrm{x}_1^{'}\gamma_1 - \alpha m_i(t)\ \right) \tag{10}$$

where $(b_{0i},\ b_{1i})\ \sim\ N(0,\Sigma)$ and $\epsilon_i(t)\ \sim\ N\left(0,\sigma^2\right)$. x is group covariate simulated from Bernoulli distribution with 0.5 probability. In addition, we used the 5-point Gauss-Legendre Quadrature rule with the R-package *"statmod"* to solve $g_i(t) = \int_0^{t_i}\exp\left(-x_i'\gamma - \alpha z_i^{'}(u)\, b_i\right)du$.

4.1. Simulating Data for Parametric Joint Modelling

To examine the performance of parameter estimates in the parametric joint model, we generate data in accordance with this model. The true parameter values were used as $\beta_0 = 2$, $\beta_1 = 1.5$, $\beta_2 = 0.5$, $\sigma = 0.8$, $\Sigma_{11} = 0.8$, $\Sigma_{12} = \Sigma_{21} = -0.025$ and $\Sigma_{22} = 0.15$ for longitudinal sub-model, and $\gamma_1 = -0.3$, $\alpha = 0.6$, $\lambda = 0.75$, $\kappa = 1.5$ for survival sub-model in Equation 10. In addition, we assigned the longitudinal time $t_{ij} = 0, 2, \ldots, 8$ (5 time points including $t_{min} = 0$ and $t_{max} = 8$), the sample size $n = 30,\ 100,\ 500$, and the censoring rate $c = 0.10,\ 0.30,\ 0.60$. The correct survival time for the i-th unit in the Weibull AFT sub-model was obtained using inverse transform sampling ($T_i^* = S_i^{-1}(u)$) in the survival function (S_i) derived from Equation 10, where u is obtained from uniform distribution. The censoring time for each unit is obtained from $C_i \sim Unif(t_{min} = 0, t_{max} = 8)$, c is a control parameter used to determine the censoring rate and T_i^* is calculated based on this determined censoring rate and then, is produced as $T_i = \min\left(T_i^*, C_i\right)$ and the event indicator as $\delta = I\left(T_i^* \le C_i\right)$ (0: lost to follow up and 1: death.). We examined scaled Schoenfeld residuals test and $\log\left(-log(S)\ \right)$ plots using the R-package *"survival"* to test whether the survival data satisfied the proportional hazards assumption after each data generated. To show that the assumption was not satisfied at each iteration ($p < \alpha = 0.05$), we recorded the p values obtained from the scaled Schoenfeld residuals test in different scenarios and displayed the results graphically. Longitudinal observations $y_i(t_i)$ for the ith subject were calculated according to the longitudinal sub-model in Equation 10. Recording times of repeated measurements were produced balanced across five time points. The random effects β_i are derived from a normal distribution with mean 0 and variance Σ, and the random error term ϵ_i is derived from a normal distribution with mean 0 and variance σ^2.

4.2. Simulation Algorithm

We can summarize the simulation steps for jointly producing longitudinal and Weibull AFT model and calculating the parameter estimation methods as follows.

1. Generate longitudinal and Weibull AFT data jointed with random effects.

- Set initial values for all parameters, censoring rate, and sample size.
- Assign longitudinal time points to each subject as $t_i = 0, 2, 4, 6, 8$.
- Generate $x_i \sim Bernoulli(0.5)$, $(b_{0i},\ b_{1i}) \sim N(0,\Sigma)$ and $\epsilon_i(t) \sim N\left(0,\sigma^2\right)$ in accordance with the initial parameters.
- Generate $y_i(t_i)$ in accordance with the longitudinal sub-model given in Equation 10.
- Generate T_i^* in accordance with the Weibull AFT sub-model given in Equation 10 using $\hat{m}_i(t)$, and set the $C_i \sim U(t_{min}, t_{max})$.
- Define $T_i = \min\left(T_i^*, C_i\right)$ and $\delta_i = I\left(T_i^* \le C_i\right)$.
- Record the p-value values obtained from the scaled Schoenfeld residuals test for the created data set.

2. Obtain parameter estimates of the created data set using methods.

- Obtain parameter estimates with the GH and AGH approaches mentioned in Section 3.1.
- Obtain parameter estimates with the Bayesian two-stage (BTS) approach mentioned in Section 3.2.

- Calculate MSE and absolute bias values of the results obtained from the methods. Here, the group parameter corresponds to the γ_1 coefficient in the Weibull AFT model (Equation 10) and the alpha parameter corresponds to the associated coefficient.

We obtained the joint model parameter estimates of the longitudinal and Weibull AFT sub-model given in Equation 10, in accordance with the prior distributions specified in Section 3.2 for the BTS approach proposed for parametric joint models in the R- package "*rstan*". For the proposed two-stage approach, we set the MCMC configuration to 2000 iterations with 1000 warm-ups for the longitudinal sub-model and 1000 iterations with 500 warm-ups for the Weibull AFT sub-model. We used R-package "*JM*" to obtain GH and AGH estimates. We compared the performances of the three methods with 1000 repetitions for all scenarios. That is, we used 1000 repetitions for n=30 and c=0.10, 0.30 and 0.60, 1000 repetitions for n=100 and c=0.10, 0.30 and 0.60, and 1000 repetitions for n=500 and c=0.10, 0.30 and 0.60.

4.3. Simulation Results

To examine the performance of the three methods at different censoring rates and samples sizes, we calculated parameter values, $MSE = \frac{1}{1000}\sum_{i=1}^{1000}\left(par_i - \widehat{par_i}\right)^2$ values and $|Bias| = \frac{1}{1000}\sum_{i=1}^{1000}|par_i - \widehat{par_i}|$ values following the above simulation steps. Table 1-2 and Figures 2-5 show these values obtained from the methods comparatively. We also showed the p-values for the scaled Schoenfeld residuals test for the proportional hazard assumption for each scenario in Figure 1.

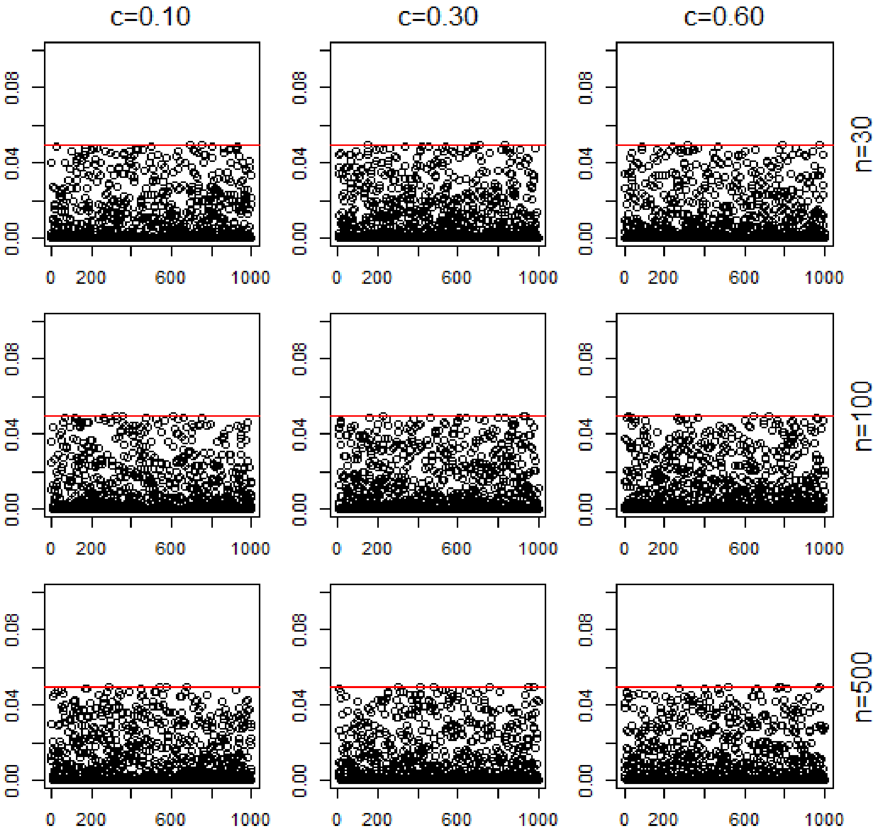

Figure 1: p-value values obtained from the scaled Schoenfeld residuals test for testing the proportional hazards assumption. The red horizontal line indicates $\alpha = 0.05$.

According to the proportional hazards assumption results obtained in Figure 1, it is seen that the p-value values obtained in all simulation results are less than $\alpha = 0.05$. Accordingly, we can say that data suitable for AFT is produced for all cases.

Table 1: The average absolute bias and MSE values of the methods in different sample and censoring situations for α parameter.

		AGH		GH		BTS	
		|Bias|	MSE	|Bias|	MSE	|Bias|	MSE
n=30	c=0.60	0.6716	0.5280	0.6638	0.5114	**0.5993**	**0.3713**
	c=0.30	0.6680	0.5351	0.6365	0.4757	**0.6027**	**0.3745**
	c=0.10	0.6066	0.4399	0.6061	0.4544	**0.5993**	**0.3712**
n=100	c=0.60	0.7909	0.6329	0.7817	0.6328	**0.5969**	**0.3682**
	c=0.30	0.7461	0.5622	0.7232	0.5335	**0.6056**	**0.3795**
	c=0.10	0.6603	0.4406	0.6262	0.4046	**0.6009**	**0.3724**
n=500	c=0.60	0.7215	0.5222	0.7359	0.5436	**0.5958**	**0.3668**
	c=0.30	0.6364	0.4054	0.6396	0.4095	**0.5988**	**0.3705**
	c=0.10	0.6041	0.3649	0.6042	0.3650	**0.5492**	**0.3112**

Table 1 and Figures 2 and 3 show that the BTS approach gives less biased estimates than the AGH and GH approaches for all censoring rates when the sample width for the α parameter is small ($n = 30$ and $n = 100$). Although the three methods give similar results when $n = 500$, especially when $c = 0.10$, it is seen that the BTS approach gives relatively less biased estimates. In addition, it was observed that the BTS approach gave less biased results for $c = 0.10$ compared to other censoring situations at all sample sizes.

Table 2: The average absolute bias and MSE values of the methods in different sample and censoring situations for group parameter(γ_1).

		AGH		GH		BTS	
		|Bias|	MSE	|Bias|	MSE	|Bias|	MSE
n=30	c=0.60	0.8524	1.1267	0.8465	1.1084	**0.5755**	**0.5205**
	c=0.30	0.7510	0.8941	0.7285	0.8528	**0.5300**	**0.4458**
	c=0.10	0.3408	0.2181	0.3652	0.2701	**0.3008**	**0.1024**
n=100	c=0.60	0.6851	0.6818	0.6915	0.6818	**0.4034**	**0.2371**
	c=0.30	0.5620	0.4313	0.5617	0.4386	**0.3749**	**0.1993**
	c=0.10	0.4688	0.3093	0.4850	0.3473	**0.3751**	**0.1947**
n=500	c=0.60	0.4409	0.2630	0.4754	0.2924	**0.3411**	**0.1323**
	c=0.30	0.3437	0.1541	0.3598	0.1650	**0.3361**	**0.1228**
	c=0.10	0.3244	0.1449	0.3215	0.1424	**0.2930**	**0.0949**

When the estimated values for the group parameter are examined in Figures 4 and 5, it is seen that the results are similar to the results of α parameter. However, when $n = 500$, it can be said that the BTS approach gives more unbiased values compared to other approaches. In addition, it is observed that the BTS approach gave less biased results for $c = 0.10$ compared to other censoring situations at all sample

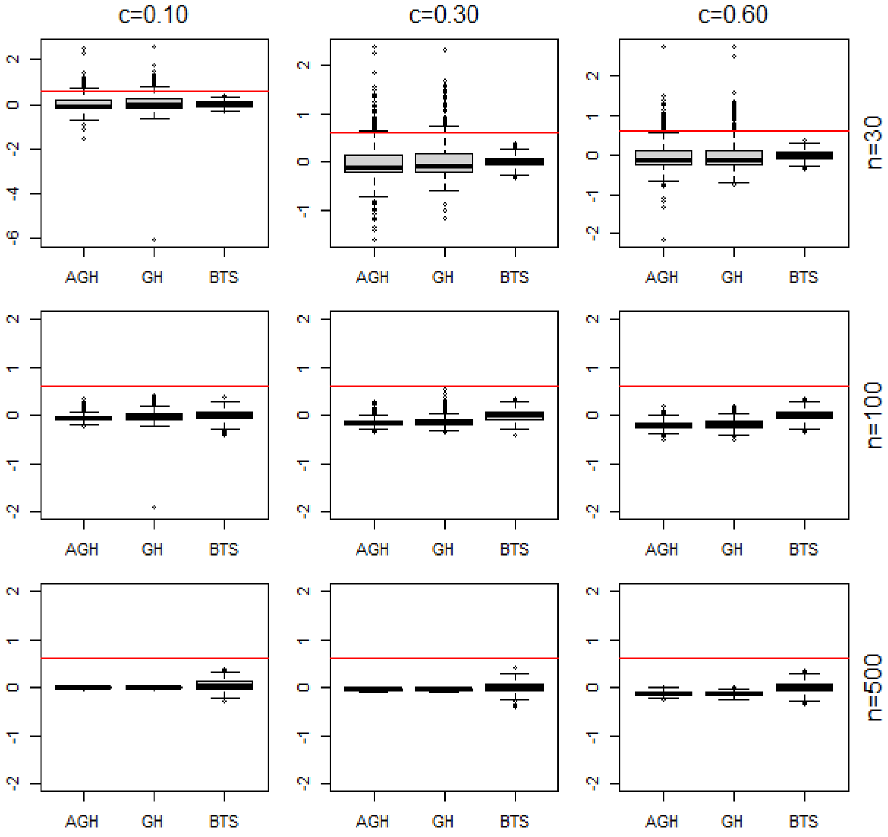

Figure 2: Graph of α parameter estimation values obtained for 1000 random samples from AGH, GH and BTS approachs. The true parameter value is shown with a red horizontal line.

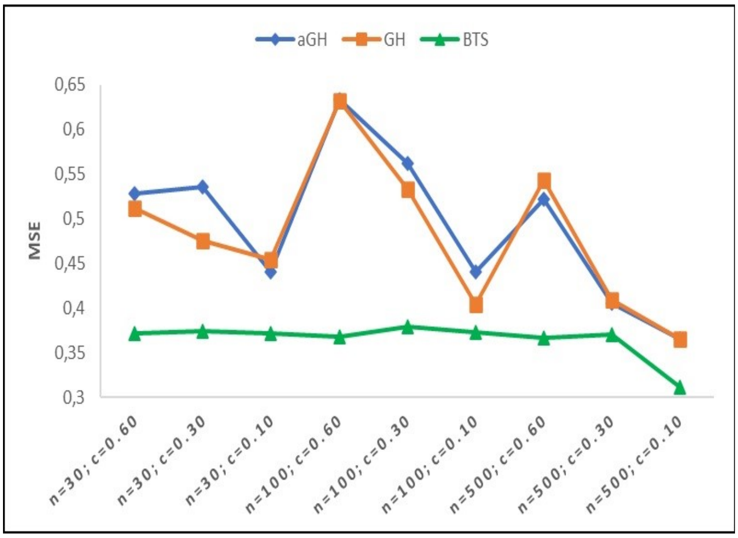

Figure 3: MSE values obtained from three methods for the α parameter.

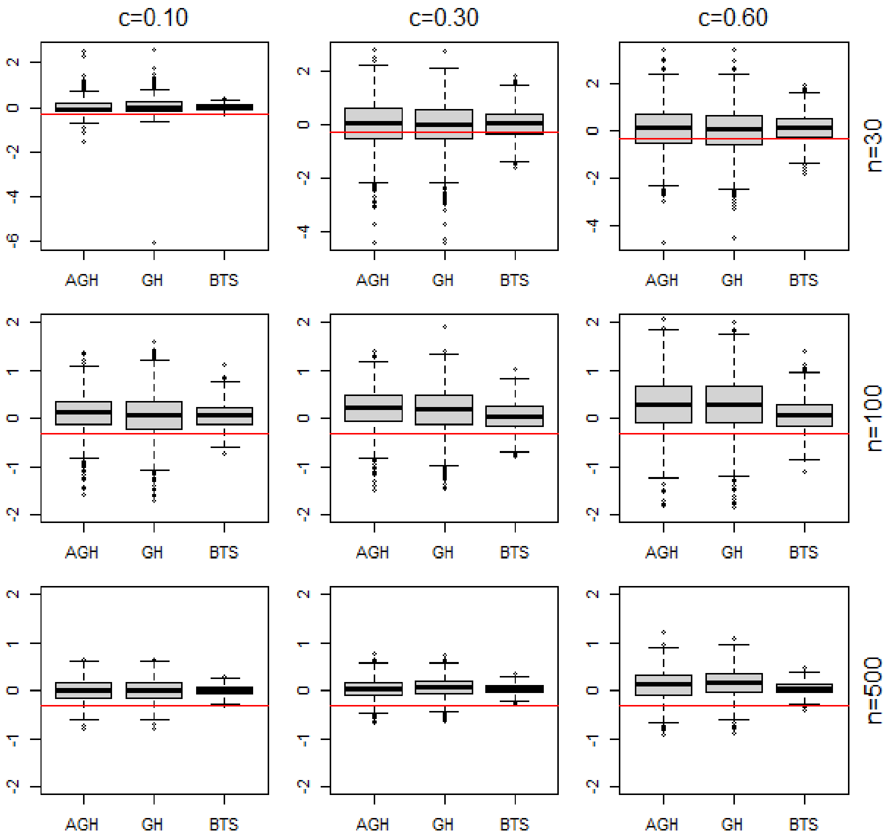

Figure 4: Graph of group parameter (γ_1) estimation values obtained for 1000 random samples from AGH, GH and BTS approachs. The true parameter value is shown with a red horizontal line.

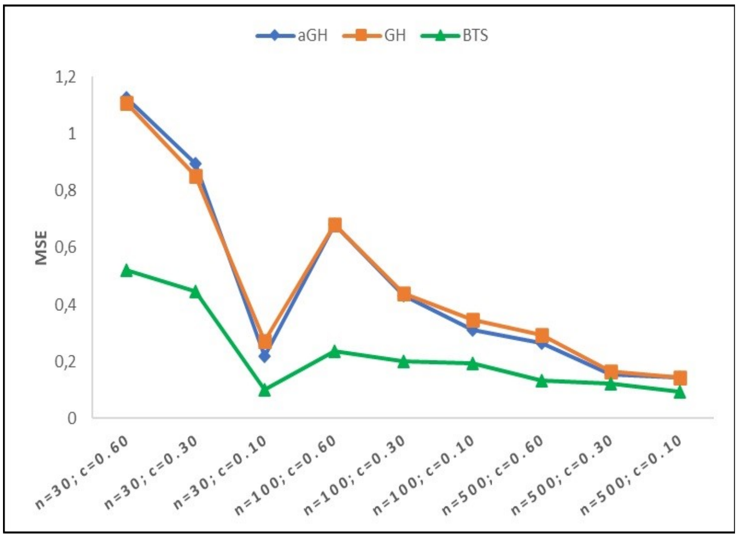

Figure 5: MSE values obtained from three methods for the group parameter.

sizes.

To examine the results of the estimation methods at different values, we conducted a different simulation study with 1000 repetitions in the case of n = 30 and c = 0.30. For this scenario, we set the parameter values as $\beta_0 = 2$, $\beta_1 = 1.5$, $\beta_2 = 0.5$, $\sigma = 0.8$, $\text{Sigma}_{11} = 0.8$, $\Sigma_{12} = \Sigma_{21} = -0.025$ and $\Sigma_{22} = 0.15$ for longitudinal sub-model, and $\gamma_1 = 1.2$, $\alpha = -1.5$, $\lambda = 0.75$, $\kappa = 1.5$. In particular, the results were compared by changing only the values of the α and group (γ_1) parameters. According to the simulation results given in Appendex A, it was seen that the results were similar to the other scenario. According to the results of this scenario, it is seen that the BTS approach gives less biased estimates than the AGH and GH methods for the group and α parameters, but this difference is lower for the α parameter. In addition, the average computation times of the GH, AGH and BTS methods for HP laptop with 1.60 GHz Intel Core i5, 16 GB RAM OS Windows were calculated as 18.619, 18.433 and 12.832, respectively. Accordingly, it can be seen that the BTS approach has less computational time compared to the other two classical approaches. The reason for this may be the integral approach used by AGH and GH methods in solving likelihood functions [7]. Additionally, sensitivity analysis for the prior distribution of α and group parameters for scenario 2 is given in Appendix B. We can say that the BTS approach gives less biased estimates when the variance values of the prior distribution are not small.

5. Application

To examine the performance of our proposed BTS approach, we apply our procedure to the aortic valve replacement surgery data set, which is publicly available in the R-package *"joineR"* [27]. The study was conducted to investigate the effect of implanting different types of heart valves in aortic positions on the survival time of patients who had aortic valve replacement surgery. The 256 patients participating in the study were followed for at least one year between 1991 and 2001, and echocardiographic measurements were taken annually. The heart valve functions examined in the study are grad, lvmi and ef. Grad, lvmi and ef show the valve gradient, left ventricular mass index (standardized) and ejection fraction at the follow-up visit, respectively. To examine the method developed in accordance with our model structure, we selected a subsample from the original data with an 83% censoring rate. Sample selection was made randomly to ensure approximately 80% censoring rate, small sample condition and other model assumptions.This new data set consists of balanced left ventricular mass index (log.lvmi) values of 43 patients, measured in 3 different years (time). Survival time (fuyrs) refers to the period from the date of surgery to the time of the event and the event (status) is considered as 1: death, 0: lost to follow up. In this study, we considered the binary hs variable as a covariate and it indicates 1: homograft, 0: stentless porcine tissue. This variable provides information about the type of heart valve implanted during surgery according to tissue types.

To investigate the effects of hs on longitudinal measurements and risk of death, we first examined the proportional hazards assumption. According to the scaled Schoenfeld residuals test results, we obtained $p = 0.032 < \alpha = 0.05$ for the hs variable and the model. We also used a plot of scaled Schoenfeld residuals against transformed time (Figure 6a) and a $\log(-\log(S))$ plot (Figure 6b) for assumption checking and concluded that the proportional hazards assumption was not satisfied for the hs variable.

We used the Kolmogorov-Smirnov test and the Q-Q plot to test whether the survival time was suitable for the Weibull distribution. According to the Kolmogorov-Smirnov test result, we can say that the data is suitable for Weibull distribution since $D = 0.134$ and $p = 0.387 > \alpha = 0.05$. In addition, according to the Q-Q plot result in Figure 7, it is seen that the survival time of the data is suitable for the Weibull distribution.

Since the hs variable does not satisfy the proportional hazards assumption and the survival time is in accordance with the Weibull distribution, we use the parametric joint model consisting of the longitudinal and Weibull AFT model to investigate the effect of hs on the longitudinal and survival data.

Parameter estimation values obtained from three different parameter estimation approaches are given in Table 3. We used $AIC = 2k - 2\ln\left(\hat{L}\right)$ and $BIC = kln\left(n\right) - 2\ln\left(\hat{L}\right)$ values to compare model performances.

According to the results obtained from real data, we can say that the BTS results are better than other methods, since the AIC and BIC values are lowest in the model obtained with the BTS method.

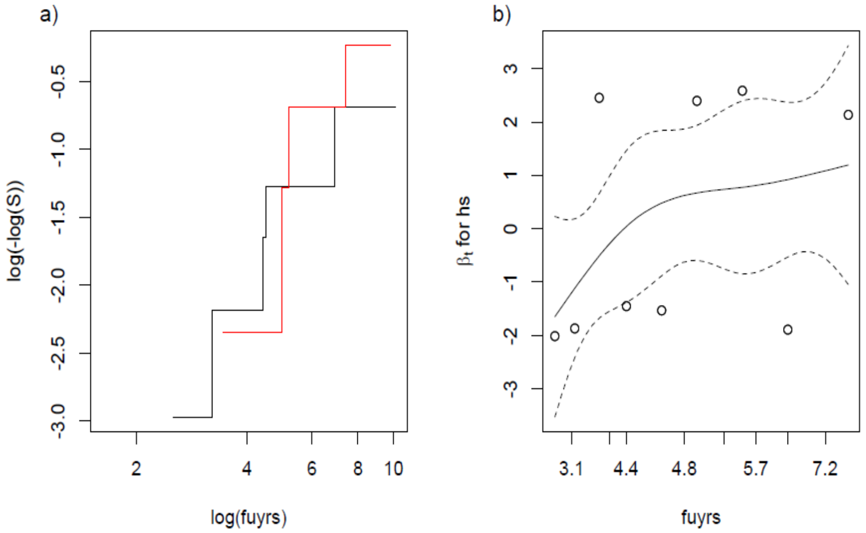

Figure 6: Proportional hazards assumption for the variable hs a) $\log(-\log(S))$ graph b) scaled Schoenfeld residuals against transformed time graph.

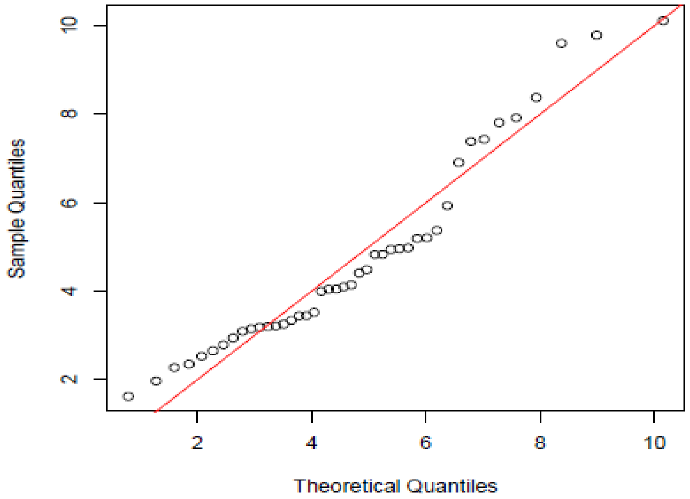

Figure 7: Q-Q plot of survival time (fuyrs)

Table 3: Parameter values of the Weibull AFT parametric joint model obtained from different approaches for the aortic valve replacement surgery data.

		GH	AGH	BTS
		Estimation(Std. Er.)	Estimation(Std. Er.)	Estimation(Std. Er.)
Longitudinal sub-model	Time	0.0158 (0.0096)	0.0125 (0.0163)	0.0220 (0.0076)
	hs	-0.0448 (0.0573)	0.0164 (0.0813)	-0.0521 (0.0011)
Survival sub-model	hs	-0.9896 (0.9049)	0.8546 (1.6384)	-0.4419 (0.1225)
	α (Assoc.)	0.0614 (0.0096)	-0.0433 (0.0256)	0.1225 (0.0074)
AIC		217.3215	405.9841	**80.3352**
BIC		236.6947	425.3573	**85.6189**

6. Discussion

There are many studies in the literature on parameter estimates in joint modelling of longitudinal and survival data ([4],[14],[18],[22]). However, there are very few studies on the AFT submodel, which is used in cases where survival data do not stasfied the proportional hazards assumption and are suitable for a certain distribution [21]. In this paper, we firtsly proposed a BTS approach to obtain parameter estimates in parametric joint modelling of survival data and longitudinal data where the proportional hazards assumption is not satisfied, and survival data follow a certain distribution. To examine the performance of our proposed approach, we conducted a simulation study in scenarios consisting of different censoring rates and samples sizes, and compared the proposed approach with the full likelihood approaches such as AGH and GH. Accordingly, in this study we showed researchers which parameter estimation approaches can obtain more unbiased results at different sample sizes and censoring rates. We also examined the applicability of the Bayesian two-stage approach as an alternative to frequentist approaches such as AGH and GH in parametric joint models.

According to the MSE and absolute bias values obtained in two different scenarios, we discovered that the group parameter of our proposed approach gives more unbiased results than the other two approaches in all scenarios. We also discovered that when the censoring rate was 0.10 for all sample sizes in the first scenario, parameter estimates were less biased than other censoring rates. Although similar results were seen for the alpha parameter jointing the two models, it was observed that all three approaches gave similar results when the sample size was 500. Similar to the other parameter, since the censoring rate is 0.10, it has been observed that our proposed method gives more unbiased results compared to the other two methods. For the second scenario, we observed that the BTS approach gave more unbiased parameter estimates for the alpha parameter compared to the AGH and GH approaches. Moreover, according to the computation times obtained for the second scenario, we obtained that the BTS approach has a shorter computation time compared to the other two approaches. The reasons why the parameter estimates of the AGH and GH approaches are biased and have a longer computation time compared to the BTS approach may be that the AGH and GH methods are solved by integration approaches.The two-stage approach has faster computational time than the full likelihood approach because it avoids high-dimensional integrals. However, since the two-stage approach consists of fitting longitudinal data in the first stage and using as a covariate these estimates into the survival model in the second stage, not using real unobserved longitudinal processes in survival models may lead to bias in parameter estimates [20]. This may be the reason why the BTS approach gives more biased estimates in the α parameter compared to the group parameter. However, we observed that the BTS approach gives less biased estimates compared to the full likelihood approaches AGH and GH.

To demonstrate the applicability of the proposed method, we compared the performance of the methods on the aortic valve replacement surgery data and concluded that our proposed approach is the best model according to the model comparison criteria (AIC and BIC). In addition, it is seen that the standard error values of the parameter estimation values of our proposed approach are lower than the other two approaches.

As a result, when the proportional hazards assumption is not satisfied for survival data and the data has a certain distribution, the parameter estimation values of the proposed BTS approach are more unbiased compared to the GH and AGH approaches when the joint model is used when testing the relationship between a single longitudinal measurement and survival data.

There are some limitations to this paper. First of all, there are different sub-models for parametric joint models such as log-logistic, logistic, log-normal and gamma. In this study, we only considered the Weibull sub-model. In future studies, other models can be examined by generating appropriate data for different models. Secondly, in our proposed Bayesian two-stage approach, some correction terms can be added to obtain less biased estimates, and the performance of the developed method can be examined through simulation studies. In addition, we only considered AGH and GH methods among the frequentist methods. However, there are different frequentist approaches in the literature. In future studies, other approaches available for joint modelling in the literature can be included in the study and it can be examined which methods give less biased estimates in different situations.

7. Acknowledgements

This study is based on Elif Yıldırım's Ph.D. thesis and supervised by Duru Karasoy.

8. Funding

The authors took on no certain funding for this study.

References

[1] Albert, P. S., & Shih, J. H, *An Approach For Jointly Modeling Multivariate Longitudinal Measurements and Discrete Time-To-Event Data*. The Annals of Applied Statistics. **4** (2010), 1517-1532.

[2] Alkhathami, A, *Joint Modeling of Longitudinal and Survival Data* PhD thesis, Carleton University, (2021).

[3] 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**94.

[4] Atlı, Z, *A study on joint modelling of longitudinal and survival data* PhD Thesis, Mimar Sinan Güzel Sanatlar University, (2020).

[5] Baghfalaki, T., Ganjali, M., Berridge, D. *Joint modeling of multivariate longitudinal mixed measurements and time to event data using a Bayesian approach.* Journal of Applied Statistics, Taylor and Francis Journals, 2014, **41**(9), 1934–1955.

[6] Cox, D. R., Oakes, D, *Analysis of survival data*. Chapman & Hall, CRC Press London. **21** (1984), 5, page 67.

[7] Crowther, M. J., Andersson, T. M., Lambert, P. C., Abrams, K. R., Humphreys, K. *Joint Öodelling of Longitudinal and Survival Data: Incorporating Delayed Entry and an Assessment of Model Misspecification.* Statistics in medicine (2016), **35** (7).

[8] Dagne, G. A. *Joint bent-cable Tobit models for longitudinal and time-to-event data.*, Journal of Biopharmaceutical Statistics, (2018), **28**(3), 385–401.

[9] DeGruttola, V.,Tu, X. M.*Modelling Progression of CD4-Lymphocyte Count and its Relationship to Survival Time*, Biometrics. **50** (1994), 1003-14.

[10] Efendi, A., Molenberghs, G., Njagi, E. D., Dendale, P, *A Joint Model for Longitudinal Continuous and Time-To-Event Outcomes with Direct Marginal Interpretation*, Biometrical Journal, **155** (2013), 572–588.

[11] Elashoff, R., Li, G., Li, N, *A Joint Model for Longitudinal Measurements and Survival Data in the Presence of Multiple Failure Types*,Biom, **164**(2008), 762–771.

[12] Faucett, C., Thomas, D., Simultaneously Modelling censored survival data and repeatedly measured covariates: A gibbs sampling approach. Stat Med, **15** (1996), 1663–1685.

[13] Fitzmaurice, G. M., Laird, N. M., Ware, J. H., *Applied Longitudinal Analysis*, John Wiley and Sons, Vol. 2, New York, (2011).

[14] Furgal, A. K. C., Sen, A., Taylor, J. M. G., *Review and Comparison of Computational Approaches for Joint Longitudinal and Time-to-Event Models*, International statistical review = Revue internationale de statistique, **87**(2), (2019), 393–418.

[15] Furgal, A., *Bayesian models for joint longitudinal and multi-state survival data*, PhD thesis, University of Michigan, (2021).

[16] Henderson, R., Diggle, P., Dobson, A, *Joint modeling of longitudinal measurements and event time data*, Biostatistics, **1** (2000), 465-480.

[17] Hickey, G. L., Philipson, P., Jorgensen, A., Dona, R. K, *joineRML: A Joint Model and Software Package for Time-To-Event and Multivariate Longitudinal Outcomes*, BMC Medical Research Methodology, **18** (2018), 1-24.

[18] Hu, W. H., Li, G., Li, N, *A bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data*, Stat Med, **28** (2009), 1601- 1619.

[19] Huang, Y., Dagne, G., Wu, L. *Bayesian inference on joint models of HIV dynamics for time-to-event and longitudinal data with skewness and covariate measurement errors.* Statistics in medicine, (2011), **30**(24), 2930–2946.

[20] Huong, Pham Thi Thu, Darfiana Nur, Hoa Pham, and Alan Branford. *A Modified Two-Stage Approach for Joint Modelling of Longitudinal and Time-to-Event Data.* Journal of Statistical Computation and Simulation,(2018) (17), 3379–98.

[21] Khan, S. A., Basharat, N, *Accelerated failure time models for recurrent event data analysis and joint modeling*, Computational Statistics, **37**(4), (2022), 1569-1597.

[22] Leiva-Yamaguchi, V., Alvares, D, *A Two-Stage Approach for Bayesian Joint Models of Longitudinal and Survival Data: Correcting Bias with Informative Prior*, Entropy (Basel, Switzerland), 23(1), (2020), 50.

[23] Li, N., Elashoff, R. M., Li, G., Saver, J, *Joint Modeling of Longitudinal Ordinal Data and Competing Risks Survival Times and Analysis of the NINDS RT-PA Stroke Trial*, Statistics in Medicine, **29** (2009), 546-557.

[24] Lin, J., Luo, S, *Deep learning for the dynamic prediction of multivariate longitudinal and survival data*, Statistics in medicine, **41**(15), (2022), 2894–2907.

[25] Murray, J., Philipson, P, *A fast approximate EM algorithm for joint models of survival and multivariate longitudinal data*, Computational Statistics & Data Analysis, **170** (2022), 107438.

[26] Mwanyekange, J., Mwalili, S., Ngesa, O, *Bayesian Joint Model for Longitudinal and Multi-state Survival Data*, International Journal of Statistics and Probability, **8** (2019), 34.

[27] Philipson, P., Sousa, I., Diggle, P., Williamson, P., Kolamunnage-Dona, R., Henderson, R., Hickey, G, *joineR: Joint Modelling of Repeated Measurements and Time-to-Event Data*, R package version 1.2.4, (2018).

[28] Pres, W., Teukolsky, S., Vetterling, W., & Flannery, B, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Vol. **3** (2007), New York.

[29] Rizopoulos, D, *Joint Models for Longitudinal and Time-to-Event Data with Applications in R.* Boca Raton, (2012a).

[30] Rizopoulos, D, *Fast fitting of Joint Models for Longitudinal and Event Time Data Using a Pseudo-Adaptive Gaussian Quadrature Rule*, Computational Statistics and Data Analysis, **56** (2012b), 491-501.

[31] Rizopoulos, D, *The R Package JMbayes for Fitting Joint Models for Longitudinal and Time-to-Event Data Using MCMC*, Journal of Statistical Software, **72** (2016), 1-45.

[32] Self, S., & Pawitan, Y. *Modeling a marker of disease progression and onset of disease. In AIDS epidemiology: Methodological issues* , eds. N. P. Jewell, K. Dietz and V. T. Farewell (pp. 231-255), (1992), Boston: Birkhuser.

[33] Song, X., Davidian, M., & Tsiatis, A, *A Semiparametric Likelihood Approach to Joint Modeling of Longitudinal and Time-To-Event Data*, Biometrics, **58** (2002), 42-753.

[34] Su, Y. R., & Wang, J. L, *Modeling left-truncated and right-censored survival data with Longitudinal Covariates*, The Annals of Statistics, **40** (2012), 1465-1488.

[35] Tang, A. M., & Tang, N. S, *Semiparametric Bayesian Inference on Skew-Normal Joint Modeling of Multivariate Longitudinal and Survival Data*, Statistics in Medicine, **34** (2015), 824-843.

[36] 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, **90** (1995), 27-37.

[37] Williamson, P. R., Kolamunnage-Dona, R., Philipson, P., Marson, A. G, *Joint Modeling of Longitudinal and Competing Risks Data*, Statistics in Medicine, **27** (2008), 6426-6438.

[38] Wu, L., Liu, W., Hu, X. J, *Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors*, Biometrics, **66** (2) (2010), 327–335.

[39] Wu, L., Liu, W., Yi, G. Y., et al. *Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods and Issues*, J Probab Stat, (2012), 1-17.

[40] Wulfsohn, M. S., Tsiatis, A. A, *A Joint Model for Survival and Longitudinal Data Measured with Error*, Biom, **53** (1997), 330-339.

[41] Xu, L, *Bayesian Multivariate Joint Modeling for Skewed-longitudinal and Time-to-event Data*, PhD thesis, University of South Florida, (2021).

[42] Yıldırım, E., & Karasoy, D, *gsem: A Stata command for parametric joint modelling of longitudinal and accelerated failure time models*, Computer methods and programs in biomedicine, **196** (2020), 105612.

9. Appendix A

Different simulation scenario results Parameter setting: $\beta_0 = 2$, $\beta_1 = 1.5$, $\beta_2 = 0.5$, $\sigma = 0.8$, $\text{Sigma}_{11} = 0.8$, $\Sigma_{12} = \Sigma_{21} = -0.025$, $\Sigma_{22} = 0.15$, $\gamma_1 = 1.2$, $\alpha = -1.5$, $\lambda = 0.75$, $\kappa = 1.5$. We performed 1000 simulation repetitions to compare the performances of the AGH, GH and BTS methods for the these parameter values.

Table 4: The average absolute bias and MSE values of the methods in different sample and censoring situations for group and α parameters for second scenario (n=30 and c= 0.30).

	AGH		GH		BTS	
	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE
α	1.3868	1.9598	1.6748	63.0880	1.0746	1.3750
Group	1.1900	2.1994	1.2427	3.8142	0.4197	0.2858
Average Comp. Time	18.619		18.433		12.832	

10. Appendix B

Sensitivity Analysis for Prior Distribution of Alpha and Group Parameter

We considered the a prior distribution for alpha and group parameters as a normal distribution with 0 mean and 100 variance for Scenario 2. Sensitivity analysis for the variance value in the prior distribution for these parameters for Scenario 2 is given in Figure 9.

In Figure 9, x-axis values express the variance values of the prior distributions for the group and α parameters for Scenario 2, respectively. For example; 0.5, 0.5: $\alpha \sim N(0, 0.5)$ and $\gamma_1 \sim N(0, 0.5)$ and 15, 5: $\alpha \sim N(0, 5)$ and $\gamma_1 \sim N(0, 15)$.

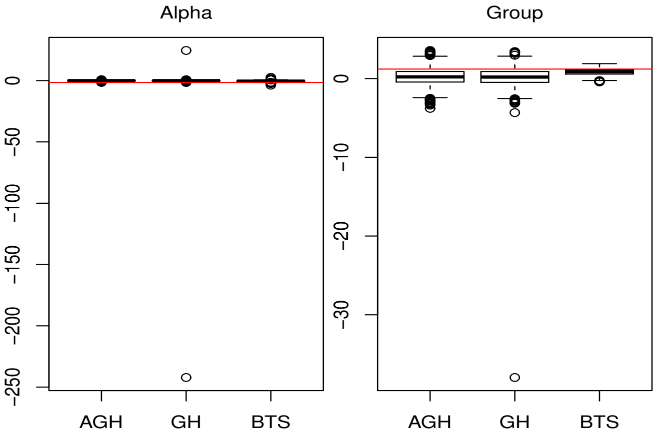

Figure 8: Graph of group (γ_1) and alpha parameter estimation values obtained for 1000 random samples from AGH, GH and BTS approachs (n=30 and c=0.30).The true parameter value is shown with a red horizontal line.

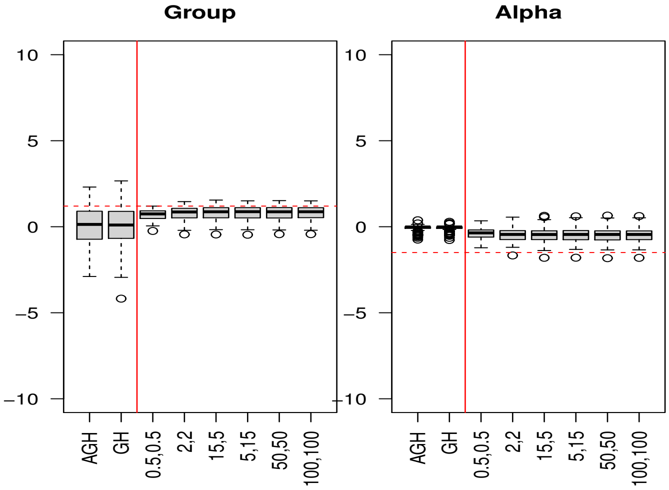

Figure 9: Sensitivity analysis for alpha and group parameter from 100 datasets with n=30, c=0.30, $\gamma_1 = 1.2$ and $\alpha = -1.5$