

## **Joint Modeling of a Longitudinal Measurement and Parametric Survival Data with Application to Primary Biliary Cirrhosis Study**

Elif DİL<sup>1,2\*</sup>, Duru KARASOY<sup>2</sup>



\* Corresponding Author

1. TED University, Faculty of Education, Department of Mathematics and Science Education, 06000, Ankara, Turkey, [elif.dil@tedu.edu.tr](mailto:elif.dil@tedu.edu.tr)

2. Hacettepe University, Faculty of Science, Department of Statistics, 06000, Ankara, Turkey.

### **Abstract**

Although longitudinal and survival data are collected in the same study, they are usually analyzed separately. Measurement errors and missing data problems arise because of separate analysis of these two data. Therefore, joint model should be used instead of separate analysis. The standard joint model frequently used in the literature is obtained by combining the linear mixed effect model of longitudinal data and Cox regression model with survival data. Nevertheless, to use the Cox regression model for survival data, the assumption of proportional hazards must be provided. Parametric survival sub-models should be used instead of the Cox regression model for the survival sub-model of the JM where the assumption is not provided. In this article, parametric joint modeling of longitudinal data and survival data that do not provide the assumption of proportional hazards are investigated. For the survival data, the model with Exponential, Weibull, Log-normal, Log-logistic, and Gamma accelerated failure time models and the linear mixed effect model are combined with random effects and the models were applied in primary biliary cirrhosis data set obtained from Mayo Clinic. After determining the best parametric joint model according to Akaike and Bayesian information criterions, the best available model was compared with standard joint model and of separate analysis of survival data and longitudinal data. As a result, in the studies where longitudinal and survival data are obtained together, it is seen that the parametric joint model gives more better results than the standard joint model when the proportional hazard assumption is not provided.

**Key Words:** Accelerated failure time models; Longitudinal data; Parametric joint modelling; Standard joint modelling; Survival data

### **1. Introduction**

The JM is used to investigate the relationship between longitudinal and survival data and the effects of longitudinal data on survival. The standard joint model (JM), which is frequently used in the literature, is obtained by combining the linear mixed effect model of longitudinal measurement (longitudinal sub-model), and Cox regression model of survival data (survival sub-model). However, in order to analyze survival data with Cox regression, the assumption of proportional hazard is required. In studies where survival data for which this assumption is not provided, and longitudinal observation, parametric survival sub-models should be used for the survival sub-model of the JM. The JM was developed for the unbiased and effective estimates of the relationship between longitudinal and survival data after 1990s. JM was first applied in 1992 by Self and Pawitan to obtain the JM parameter estimates of the linear mixed effect (LME) model for longitudinal data and Cox regression model for survival data (Self and Pawitan, 1992).

In the literature particularly in the last 20 years, studies on JM appear. Primary studies on basic JM with one covariate survival and longitudinal data were performed by Self and Pawitan (1992), DeGruttola and Tu (1994), Tsiatis,

DeGruttola and Wulfsohn (1995), Faucett and Thomas (1996) and Wulfsohn and Tsiatis (1997). Henderson, Diggle, and Dobson (2000) have developed a general model that adopts two stationary Gaussian processes, including sequential correlation and random effects for longitudinal measurements and survival times. Tsiatis and Davidian (2001) proposed a simple model for estimating JM parameters that do not require distribution assumptions over random effects. Tseng, Hsieh and Wang (2005) studied the maximization of JM likelihood function in the presence of longitudinal data including measurement errors and accelerated failure time. Elashoff, Li and Li (2008) and Hu, Li and Li (2009) have done studies on JM in the event of competing risks and multiple failures. Rizopoulos (2010) developed the “JM” package in the R program to obtain the JM estimates of the longitudinal and survival data. Su and Wang (2012) investigated the JM of left truncation survival data and longitudinal observations.

In the next section, we introduce notation and describe the parametric joint models. In addition, we have defined how model parameters are obtained. In Section 3, parametric and standard JM have been applied to primary biliary cirrhosis (PBC) data obtained from Mayo Clinic in order to determine the effect of longitudinal observation on the survival time of patients with liver cirrhosis. And finally in section 4, we provide final comments for compared to standard JM, parametric JM and separate analysis for longitudinal and survival data and investigated the effects of using the standard JM on parameter estimation when the assumption of proportional hazards is not provided.

## 2. Material and Method

JM are two fundamental compact the formed longitudinal sub-model and survival sub-model. Standard JM is composed of LME model for longitudinal sub-model and Cox regression model for survival sub-model.

Standard JM can be obtained as follows;

$$\mathbf{y}_i(t) = \mathbf{x}'_i(t)\boldsymbol{\beta} + \mathbf{z}'_i(t)\mathbf{b}_i + \boldsymbol{\varepsilon}_i(t)$$

$$m_i(t) = x'_i(t)\boldsymbol{\beta} + z'_i(t)b_i \quad (1)$$

$$h_i(t | M_i(t), \alpha_i) = h_0(t) \exp\{\gamma' \alpha_i + \alpha m_i(t)\}, t > 0 \quad (2)$$

where Equation (1) shows the longitudinal sub-model of the joint model.  $\mathbf{x}'_i(t)\boldsymbol{\beta}$  and  $\mathbf{z}'_i(t)\mathbf{b}_i$  shows the design vectors for fixed effect  $\boldsymbol{\beta}$  and random effect  $\mathbf{b}_i$ , respectively.  $\boldsymbol{\varepsilon}_i(t)$  are errors of  $m_i \times 1$  dimension that is independent of other variables. Errors have been normal distribution in  $N(0, \sigma^2_\varepsilon)$  form (Verbeke and Molenberghs 2000; Rizopoulos, Verbeke and Molenberghs 2010). Equation (2) shows the survival sub-model of the joint model obtained by Cox regression. Where,  $m_i(t)$  indicates the unobserved longitudinal data at time  $t$  and includes measurement errors.  $\alpha_i$  and  $\gamma$  shows the covariates (e.g. treatment indicator, disease history etc) and regression coefficient vector corresponding to covariates, respectively.  $h_0(\cdot)$  is the basic risk function and can be defined in a parametric and non-parametric way as in the Cox model (Cox 1972; Bahçecitapar 2018).  $\alpha$  in Equation (2) is the association parameter for longitudinal sub-model and survival sub-model. There is no relationship between two process when  $\alpha = 0$  (Rizopoulos, Verbeke and Molenberghs 2010; Rizopoulos 2012).

In order to apply Cox regression model for survival data, the assumption of proportional hazard (PH) must be provided. Parametric models should be used in cases where the assumption is not provided, and the survival time corresponds to a certain distribution. Parametric models are divided into PH and accelerated failure time (AFT) models. Parametric PH models are parametric version of Cox regression, and commonly used PH models; Exponential, Weibull and Gompertz (Cleves, Gould and Marchenko, 2016). Although parametric PH models are widely used for survival data, there is little distribution that can be used in modelling. In this case, AFT models are used an alternative to parametric PH models. In addition, AFT models are also used as an alternative to PH models to overcome the statistical problems caused by the violation of the assumption of PH (Faruk, 2018). Therefore, survival sub-model for JM should be done by parametric regression methods where survival data do not provide the assumption of PH and have a certain parametric distribution. Let we redefine the survival sub-model given in Equation (2) for parametric AFT models as follows (Pericleous 2016);

$$h(t); \begin{cases} \log(t), \text{accelerated failure time} \\ t, \text{other parametric models} \end{cases}$$

From this form of  $h(t)$ , risk function  $h(t)$  i. for unit can be written as follows;

$$h(t_i) = \mu_i(t_i) + \sigma \times \varepsilon_i \quad (3)$$

where  $\sigma$  and  $\varepsilon_i$  indicates the scale parameter and errors from appropriate survival distribution, respectively and  $\mu_i(t_i)$  can be defined as follows,

$$\mu_i(t_i) = \alpha m_i(t) + \alpha'_i \gamma \quad (4)$$

Here, parametrical survival sub-models can be written in general form as follows;

$$h_i(t) = \alpha [x'_i(t) \beta + z'_i b_i] + \alpha'_i \gamma + \sigma \times \varepsilon_i \quad (5)$$

The survival and basic hazard functions for the parametric sub models of the JM are given in Table 2.1.

**Table 2.1.** Parametric sub-models for JM

Parametric Survival Models	S(t)	f(t)
Exponential AFT sub-model	$\exp[-(\text{texp}[\mu_i(t_i)])]$	$(\text{texp}[\mu_i(t_i)])^{\exp(-(\text{texp}[\mu_i(t_i)]))}$
Weibull AFT sub-model	$\exp[-(\text{texp}[\mu_i(t_i)])^{\exp(1/\sigma)}]$	$(\text{texp}[\mu_i(t_i)])^{\exp(1/\sigma - (\text{texp}[\mu_i(t_i)])^{\exp(1/\sigma)})}$
Log-Logistic AFT sub-model	$[1 + \exp\{\sigma^{-1}(y_i - \mu_i)\}]^{-1}$	$\exp\{\sigma^{-1}(y_i - \mu_i)\} \sigma^{-1} (1 + \exp\{\sigma^{-1}(y_i - \mu_i)\})^{-2}$
Log-Normal AFT sub-model	$\left[1 - \Phi\left(\frac{(y_i - \mu_i)}{\sigma}\right)\right]$	$\exp\left[\frac{-(T_i(t_i) \sigma (2\pi)^{\frac{1}{2}})^{-1} (y_i - \mu_i)}{2\sigma^2}\right]$

## 2.1. Parameter estimates methods for joint modelling

Two methods are used for the parameter estimation of the JMs, namely the two-stage procedure and the solution of joint probability. In literature, different two-stage methods have been suggested (Self and Pawitan 1992; Tsiatis, DeGruttola and Wulfsohn 1995; Wu et al. 2012). Two-stage method has mainly been defined as follows;

- 1) The longitudinal covariates are modeled with the LME, and the unit-specific values of the covariates are estimated.
- 2) The survival model is estimated by the values obtained in stage 1.

This approach reduces bias in parameter estimation of the Cox regression model (Wulfsohn and Tsiatis 1997). The approach is simple to implement and allows parameter estimation using existing software (Wu et al. 2012). However, this method cannot simultaneously use information from the survival and longitudinal process at the estimation stage of each model (Wu et al. 2012). The use of only longitudinal results in the first stage may comprise biased estimates in the LME, and as a result, biased and ineffective results may arise in parameter estimates of survival analysis

(Ibrahim, Chu and Chen 2010). Methods based on the joint probability function allow for unbiased parameter estimates. The JM paradigm includes different model strategies such as pattern-mixture models and selection models (Pericleous 2016). When random effects are attached to these model structures, random mixed model, random selection model and random effect model are obtained. There are the differences and similarities in the definition of each model. A simple and detailed study was conducted by Sousa and McCrink to examine this distinction (Sousa 2011). Y, T and B denote longitudinal, survival time and random effects, respectively, and can be written follow as;

$$\begin{aligned}\text{Selection Model: } [Y, T, B] &= [B] [Y | B] [T | Y] \\ \text{Mixed Model: } [Y, T, B] &= [B] [T | B] [Y | T] \\ \text{Random Mixed Model: } [Y, T, B] &= [B] [Y | B] [T | B]\end{aligned}\quad (6)$$

Among these methods, a random effect model (also is named shared parameter model) that connects survival and longitudinal process with random effect is preferred dominantly (Pericleous 2016). Therefore, in this article is used the random effect model for linked longitudinal and survival sub model.

The likelihood function of longitudinal measurement and survival data for i-th unit obtained from the random effect model can be defined as follows;

$$\begin{aligned}\log p(T_i, \delta_i, y_i; \theta) &= \log \int p(T_i, \delta_i, y_i, b_i; \theta) db_i \\ &= \log \int p(T_i, \delta_i | b_i; \theta_t, \beta) [\prod_j p\{y_i(t_{ij}) | b_i; \theta_y\}] p(b_i; \theta_b) db_i\end{aligned}\quad (7)$$

where  $\theta = (\theta'_t, \theta'_y, \theta'_b)$  has all parameter vector, and  $\theta_t$ ,  $\theta_y$  and  $\theta_b$  shows survival data parameters, longitudinal data parameters, and covariance matrix parameters of random effects, respectively.  $p(T_i, \delta_i | b_i; \theta_t, \beta)$  shows the conditional density of survival sub-model, and  $p(y_i | b_i; \theta)p(b_i; \theta)$  indicates joint density function for longitudinal measurement and random effects. Integral approaches, such as, Gauss-Hermite, Laplace, EM algorithm are needed because the integrals of shared parameter models (Equation 7) are very complex and difficult to calculate analytically to obtain parameter estimates (Rizopoulos 2010).

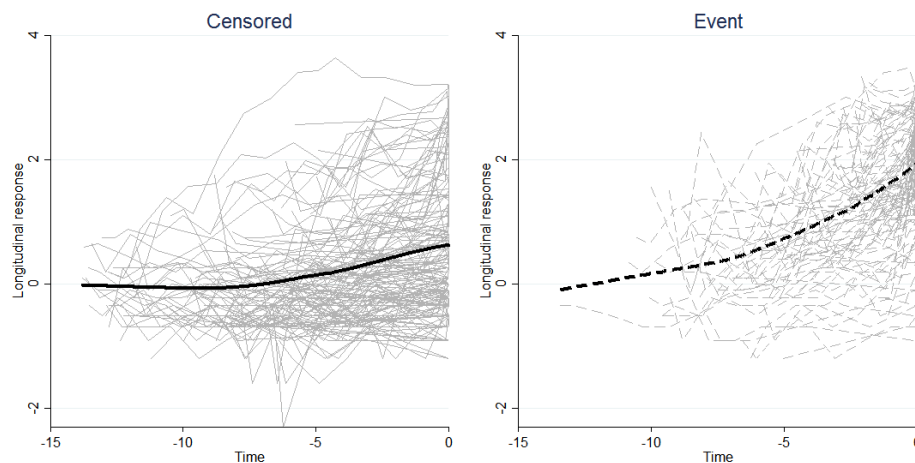
### 3. Results and Discussion

PBC data in the literature were used to examine and compare different parametric JM structures, standard JM structure, and separate analysis of two data. The data were collected to examine the progress of PBC in 312 patients in the Mayo Clinic in 1924-1984 years (Murtaugh et al. 1994). This data was obtained from "JM" package in the R program. PBC is a chronic liver disease with no effective treatment other than liver transplantation (Gordon 1987). Of the 312 patients in the PBC data set, 158 received the D-penicillin drug and the other 154 patients were identified as the placebo group. Serum bilirubin levels were measured repeatedly at specific time intervals (6 months and after each year) and total 1945 consist of unbalanced longitudinal measurements. The main aim of the study was to research the effect of treatments on the relationship between serum bilirubin (longitudinal measurement) and time of death.

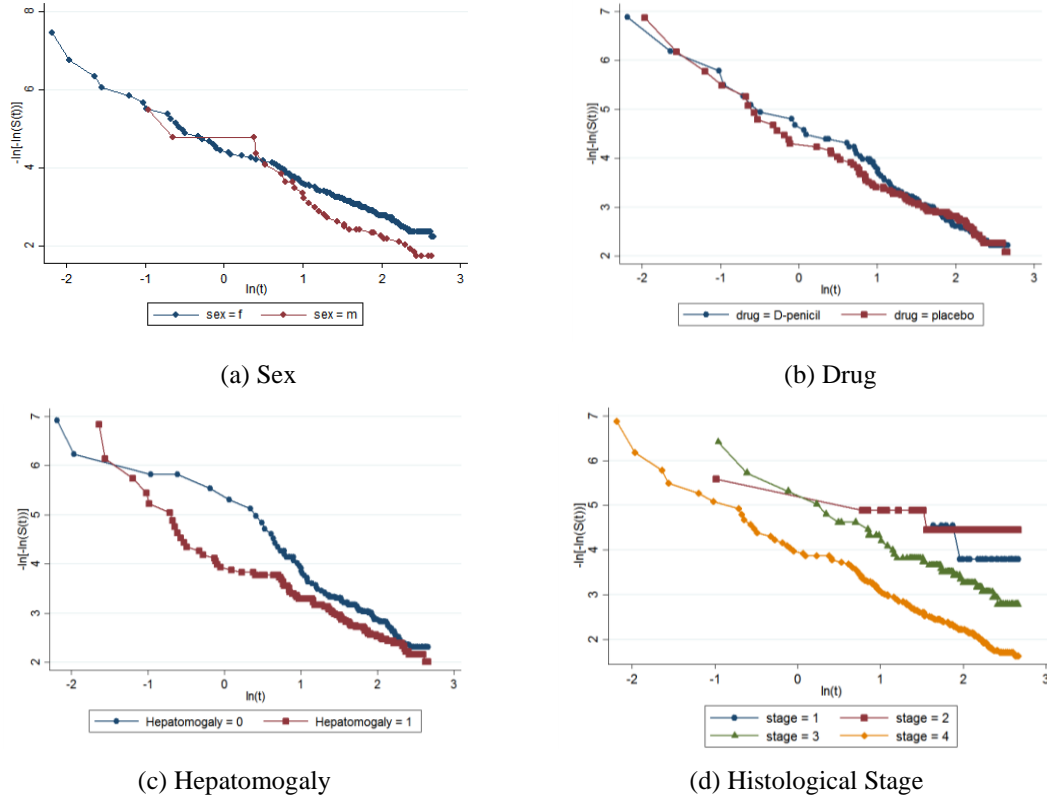
Since the patients were started to be followed, the survival time (year) were taken as the period until death. Here the failure is considered as death and other patients expressed as censored. At the end of study 140 of the 312 patients (44.9%) died and 172 (55.1%) was censored. The data set includes clinical, biochemical, and demographic risk factors for each patient. Demographic factors; age and sex of patients, Biochemical factors; drug (D-penicillin and placebo group), ascites (accumulation of water in the abdomen due to liver failure-no/yes), hepatomegaly (liver growth status-no/ yes), skin disorder (blood vessel disorders in the skin-no/yes), edema (swelling condition in hands and feet-no/yes, edema despite drug use/ yes) and consists of histological stage. Serum bilirubin (mg/dl) values were taken as biochemical properties and because of having a left skewed distribution, the logarithmic transformations of bilirubin. The covariates and levels used are given in Table 3.1.

**Table 3.1.** Used covariates and levels

Covariates	Levels	n	%
Sex	1- Female	276	88.5
	2- Male	36	11.5
Drug	1- D-penisilin	158	50.6
	2- Placebo	154	49.4
Ascites	1-No	247	79.2
	2- Yes	41	13.1
Hepatomogaly	1- No	131	42.0
	2- Yes	157	50.3
Skin Disorder	1- No	188	60.3
	2- Yes	100	32.1
Edema	1- No	43	13.8
	2- Edema despite drug	64	20.5
	3- Yes	205	65.7
Histological stage	1- Stage1	11	3.5
	2- Stage2	45	14.4
	3- Stage3	102	32.7
	4- Stage4	154	49.4

**Figure 3.1.** Temporal graphics of log (serbilir) values for patients

When Figure 3.1 examined, it is seen that the patients who died have been more sharply trend than the censored patients. This is a positive relationship between longitudinal observation and survival time and as a result, it can be said that the increase in the longitudinal observation levels increases the risk of death.



**Figure 3.2.** Graphs for the assumption of PH of some covariates

The assumption of PH required for Cox regression model before the combined model was examined and graphically analyzed in Figure 3.2. When the graphs are investigated, assumption has not been provided since there is overlap in the levels of some variables. Therefore, Cox regression model should not be used for survival analysis of the JM. However, the results of the standard JM to compare with the parametric JM results are given in Table 3.2.

**Table 3.2.** Standard JM results of LME model and Cox regression model

		$\hat{\beta}$	Std. error	z	p
Longitudinal Sub Model	Fixedterm	0.4352	0.2714	1.6033	0.1089
	Age	-0.0038	0.0051	-0.7409	0.4588
	Sex2	0.2369	0.1704	1.3904	0.1644
	Drug2	0.1283	0.1053	1.2185	0.2230
	Ascites2	0.1043	0.0415	2.5118	<b>0.0120</b>
	Hepatomo2	-0.0183	0.0247	-0.7427	0.4577
	SkinDisorder2	0.0640	0.0287	2.2273	<b>0.0259</b>
	Edema2	0.1798	0.0344	5.2289	<b>0.0000</b>
	Edema3	0.2706	0.0537	5.0434	<b>0.0000</b>
	HistStage2	0.0340	0.0646	0.5256	0.5992
	HistStage3	0.1021	0.0669	1.5259	0.1270
	HistStage4	0.1642	0.1642	2.2939	<b>0.0218</b>
Cox Sub Model	Age	0.0306	0.0097	3.1460	<b>0.0017</b>
	Sex2	0.3957	0.2537	1.5596	0.1188
	Drug2	0.2125	0.1840	1.1552	0.2480
	Ascites2	0.2925	0.2292	1.2758	0.2020
	Hepatomo2	-1.6163	0.2042	0.2042	<b>0.0000</b>
	SkinDisorder2	-0.6230	0.1994	-3.1249	<b>0.0018</b>

Edema2	0.8477	0.2421	3.5012	<b>0.0005</b>
Edema3	2.0730	0.2692	7.7003	<b>0.0000</b>
HistStage2	-1.7986	0.7371	-2.4401	<b>0.0147</b>
HistStage3	-1.4441	0.5745	-2.5138	<b>0.0119</b>
HistStage4	-0.4095	0.5468	-0.7488	0.4540
Assoct	1.3913	0.1161	11.9833	<b>0.0000</b>

Since the data did not fit the Cox regression model, parametric survival models were examined for JM. Accelerated Failure Time (AFT) models with more probability distribution for parametric regression models have been used and to determine which of the models are best appropriate have been used the model comparison criteria (AIC and BIC). The AIC and BIC values of the JM of the linear mixed effect model with Exponential, Weibull, Log-normal, Log-logistic and Gamma regression models from parametric models are given in Table 3.3.

**Table 3.3.** Model comparison criteria results for parametric JMs

Model	AIC	BIC
Exponential	4738.236	4894.281
Weibull	<b>4721.314</b>	<b>4882.982</b>
Log-normal	4742.735	4904.352
Log-logistic	4727.060	4888.678
Gamma	4722.846	4884.464

According to the model comparison criteria given in Table 3.3, the best model was determined as parametric JM of LME model with Weibull regression model.

**Table 3.4.** Parametric JM result of LME model and Weibull regression model

		$\hat{\beta}$	Std. error	z	p
Longitudinal Sub Model	Fixedterm	0.6535	0.2943	2.22	<b>0.026</b>
	Age	-0.0079	0.0055	-1.42	0.155
	Sex2	0.4905	0.1781	2.75	<b>0.006</b>
	Drug2	0.0945	0.1125	0.84	0.401
	Ascites2	0.1408	0.0506	2.78	<b>0.005</b>
	Hepatomo2	-0.0205	0.0288	-0.71	0.475
	SkinDisorder2	0.1139	0.0334	3.41	<b>0.001</b>
	Edema2	0.3977	0.0393	10.11	<b>0.000</b>
	Edema3	0.6492	0.0569	11.40	<b>0.000</b>
	HistStage2	-0.1207	0.0741	-1.63	0.104
	HistStage3	0.0241	0.0763	0.32	0.752
	HistStage4	0.2189	0.0818	2.68	<b>0.007</b>
Weibull Sub Model	Age	-0.0431	0.0092	-4.64	<b>0.000</b>
	Sex2	-0.4022	0.2682	-1.50	0.134
	Drug2	-0.3059	0.1809	-1.69	0.091
	Ascites2	-0.3434	0.1627	-2.11	<b>0.035</b>
	Hepatomo2	0.7130	0.1491	4.78	<b>0.000</b>
	SkinDisorder2	0.1397	0.1392	1.00	0.315
	Edema2	-0.6588	0.1725	-3.82	<b>0.000</b>
	Edema3	-1.5090	0.1957	-7.71	<b>0.000</b>
	HistStage2	0.9656	0.6554	1.47	0.141
	HistStage3	0.2690	0.5366	0.50	0.616
	HistStage4	-0.3743	0.5217	-0.72	0.473
Assoct		-1.1569	0.0943	-12.27	<b>0.000</b>

According to the results given in Table 3.4, gender, acid, skin defect, edema and histological stage4 covariates in longitudinal sub-model and age, ascites, hepatomegaly and edema covariates were found statistically significant in Weibull sub-model. According to Weibull sub-model, survival time is shortened as age progresses, survival time in patients with hepatic insufficiency as fluid accumulation in the abdomen is 1.41 times shorter than those without. Survival time in patients with liver growth is 2 times longer than in those without liver growth and in patients with hand and foot swelling despite the use of the drug, survival time is 1.9 times shorter than those without hand and foot swelling, survival time in patients with hand and foot swelling is 4.5 times shorter than those without. The Assoct coefficient shows the association parameter between the longitudinal and survival data. Accordingly, it can be said that the increase in serum bilirubin measurements will result in a shortening of survival time.

To compare the results obtained from the JM, the results of separate analysis of the longitudinal and survival data are given in Table 3.5 and Table 3.6, respectively.

**Table 3.5.** LME model results

	$\hat{\beta}$	Std. error	z	p
Fixedterm	-0.0679	0.3925	-0.17	0.863
Age	-0.0132	0.0049	-2.67	<b>0.008</b>
Sex2	0.4760	0.1539	3.09	<b>0.002</b>
Drug2	-0.0166	0.0978	-0.17	0.866
Ascites2	1.1179	0.2717	4.11	<b>0.000</b>
Hepatomo2	0.7778	0.1539	5.05	<b>0.000</b>
SkinDisorder2	0.4243	0.1468	2.89	<b>0.000</b>
Edema2	0.2944	0.1786	1.65	0.100
Edema3	0.4256	0.2645	1.61	0.109
HistStage2	0.0515	0.3805	1.35	0.177
HistStage3	0.6385	0.3459	1.85	0.066
HistStage4	0.7478	0.3483	2.15	<b>0.033</b>

**Table 3.6.** Weibull regression model results

	$\hat{\beta}$	Std. error	z	p
Fixedterm	7.8327	0.9469	8.27	<b>0.000</b>
Age	-0.0309	0.0100	-3.08	<b>0.002</b>
Sex2	-0.4469	0.2399	-1.86	0.062
Drug2	-0.1999	0.1761	-1.13	0.256
Ascites2	-0.7177	0.2333	-3.08	<b>0.002</b>
Hepatomo2	0.5186	0.1986	2.61	<b>0.009</b>
SkinDisorder2	-0.0140	0.1945	-0.07	0.943
Edema2	-0.8790	0.2462	-3.57	<b>0.000</b>
Edema3	-1.7537	0.2927	-5.99	<b>0.000</b>
HistStage2	0.6030	0.9179	0.66	0.511
HistStage3	-0.3688	0.7476	-0.49	0.622
HistStage4	-1.0932	0.7361	-1.49	0.138
Log(serBilir)(t)	-0.0001	0.0000	-2.92	<b>0.004</b>

According to the results of Weibull parametric model, it is seen that the hazard ratios and standard errors of parameter estimations are higher than the Weibull parametric JM. In addition, the relationship between time-dependent serum bilirubin values and survival was found to be lower than the Weibull parametric JM.



The hazard ratios and standard errors of parameter estimates in the standard JM were found to be higher than the Weibull parametric JM. In addition, the relation between serum bilirubin values and life expectancy was higher than the Weibull parametric model according to standard combined model results.

#### 4. Conclusion

Survival and longitudinal data obtained in the same study should be analyzed by JM in case there is a relationship between them. The standard JM construction which is generally used is based on combining linear mixed effect model of longitudinal observations and Cox regression model of survival data by using shared parameter models. Nevertheless, on account of using the Cox regression model for survival data, the assumption of PH should be provided. In cases where the assumption is not provided, to obtain unbiased and reliable results Exponential, Weibull, Log-normal, Log-logistic and Gamma parametric models should be used for the JM survival sub-model.

In this study, standard JM, parametric JMs and separate analysis of longitudinal and survival data were performed by the data obtained from Mayo Clinic in order to investigate the effect of treatments on the relationship between serum bilirubin and time of death in PBC patients. Parametric joint model has the indicators that yield better results than separate analysis of longitudinal and survival data and the standard joint model when assumption of PH is not provided. The standard error and hazard ratio of parameter obtained by from parametric joint model are slower than the standard error and hazard ratio of parameter obtained from standard joint model and separate analysis of longitudinal and survival data. Furthermore, when the result of separate analyzes are examined, it is seen that the effect of longitudinal observation on survival data is calculated as very low. Especially on important issues such as human health, it is important to accurately calculate the relationship between longitudinal measurement and survival data and the risk factors of covariates. Therefore, in cases where the assumption of proportional hazards is not provided, parametric survival models should be used for survival partial of the joint model in order to obtain unbiased and effective results.

#### References

1. Bahçecitapar MK (2018) Some Factor Affecting Statistical Power of Approximate Tests in The Linear Mixed Model for Longitudinal Data. *Communications in Statistics – Simulation and Computation* 47: 294-314.
2. Cleves M, Gould W, Marchenko Y (2016) *An Introduction to Survival Analysis Using Stata*. Revised 3rd Ed, Stata Press.
3. Cox DR (1972) Regression Models and Life-Tables (with discussion). *J R Stat Soc Series B* 34: 187-220.
4. DeGruttola V, Tu XM (1994) Modelling Progression of CD4-Lymphocyte Count and its Relationship to Survival Time. *Biometrics* 50: 1003-14.
5. Elashoff R, Li G, Li N (2008) A Joint Model For Longitudinal Measurements and Survival Data in the Presence of Multiple Failure Types. *Biom* 64: 762–771.
6. Faruk A (2018) The Comparison of Proportional Hazards and Accelerated Failure Time models in Analyzing the First Birth Interval Survival Data. *IOP Conf. Series: Journal of Physics: Conf. Series* 974.
7. Faucett C, Thomas D (1996) Simultaneously Modelling censored survival data and repeatedly measured covariates: A gibbs sampling approach. *Stat Med* 15: 1663–1685.
8. Gordon RD, Iwatsuki S, Tzakis AG, et al (1987) The Denver-Pittsburgh Liver Transplant Series. *National Library of Medicine National Institutes of Health* 43: 43-50.
9. Henderson R, Diggle P, Dobson A (2000) Joint modeling of longitudinal measurements and event time data. *Biostatistics* 1: 465-480.
10. Hu WH, Li G, Li N (2009) A bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. *Stat Med* 28: 1601- 1619.
11. Ibrahim JG, Chu H, Chen LM (2010). Basic Concepts and Methods for Joint Models of Longitudinal and Survival Data. *Journal of Clinical Oncology: Statistics in Oncology* 28: 2796-2801.

12. Murtaugh PA, Dickson ER, Van Dam GM, et al (1994) Primary Biliary Cirrhosis: Prediction of Short-Term Survival Based on Repeated Patient Visits. *Hepatology* 20: 126–134.
13. Pericleous P, (2016) Parametric Joint Modelling for Longitudinal and Survival Data. Phd Thesis, University of East Anglia, UK.
14. Rizopoulos D (2010) JM: An R Package for the Joint Modelling of Longitudinal and Time-to Event Data. *Journal of Statistical Software* 35: 1-33.
15. Rizopoulos D, Verbeke G, Molenberghs G (2010) Multiple-Imputation-Based residuals and Diagnostic Plots for Joint Models of Longitudinal and Survival Outcomes. *Biom* 66: 20-29.
16. Rizopoulos D (2012) Joint Models for Longitudinal and Time-to-Event Data with Applications in R. 1th ed. London: Boca Raton.
17. Self S, Pawitan Y (1992) Modeling a marker of disease progression and onset of disease. *AIDS epidemiology: Methodological issues*, eds. N. P. Jewell, K. Dietz and V. T. Farewell, 231-255. 1th ed. Boston: Birkhuser.
18. Sousa I (2011) A Review on Joint Modelling of Longitudinal Measurements and Time-to-Event. *REVSTAT-Statistical Journal* 9: 57-81.
19. Su YR, Wang JL (2012) Modeling left-truncated and right-censored survival data with Longitudinal Covariates. *The Annals of Statistics* 40: 1465-1488.
20. Tseng Y, Hsieh F, Wang J (2005) Joint Modelling of Accelerated Failure Time and Longitudinal Data. *Biometrika* 92: 587-603.
21. Tsiatis AA, DeGruttola V, Wulfsohn MS (1995) 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: 27-37.
22. Tsiatis AA, Davidian M (2001) A Semiparametric Estimator for the Proportional Hazards Model with Longitudinal Covariates Measured with error. *Biometrika* 88:447- 458.
23. Verbeke G, Molenberghs G (2000) A Model for Longitudinal Data, Linear Mixed Models for Longitudinal Data. 1th ed. New York.
24. Wu L, Liu W, Yi GY, et al (2012) Analysis of Longitudinal and Survival Data: Joint Modeling, Inference Methods and Issues. *J Probab Stat* 2012: 1-17.
25. Wulfsohn MS, Tsiatis AA (1997) A Joint Model for Survival and Longitudinal Data Measured with Error. *Biom* 53: 330-339.