Joint Modeling for Primary Biliary Cholangitis

Mohammed Rahman

December 13, 2019

Background

Primary biliary cholangitis, previously called primary biliary cirrhosis (PBC) is an autoimmune, slowly progressive, liver disease with liver biopsy findings of nonsuppurative destructive cholangitis and interlobular bile duct destruction (Purohit, 2015). Primary biliary cholangitis is common among women of middle age worldwide. The disease ratio among female to male is 9:1. The diagnosis is usually made in women of age between 30 and 60 (Pandit 2019). Between January 1974 and May 1984, the Mayo Clinic conducted a double blinded randomized trial in primary biliary cholangitis of the liver (PBC), comparing the drug D-penicillamine with a placebo. There were 424 patients who met the eligibility criteria seen at the Clinic while the trial was open for patient registration. Both the treating physician and the patient agreed to participate in the randomized trial in 312 of the 424 cases. The date of randomization and many clinical, biochemical, serologic, and histologic parameters were recorded for each of the 312 clinical trial patients (Fleming & Harrington).

Research Questions

1. Does treatment affect survival? How does the Cox Model compare to the Joint Model?
2. Are the average longitudinal change in Serum Bilirubin levels different between males and females?
3. Can we increase predictivity of longitudinal log Seribilium? Is the Joint model a better model for prediction?
4. How strong is the association between Serum Bilirubin and the risk of death?

Purpose of Joint Modeling

A Joint Model is a model that incorporates two outcome variables, one longitudinal and one survival. The two processes are associated and so their joint distributions are of the form,

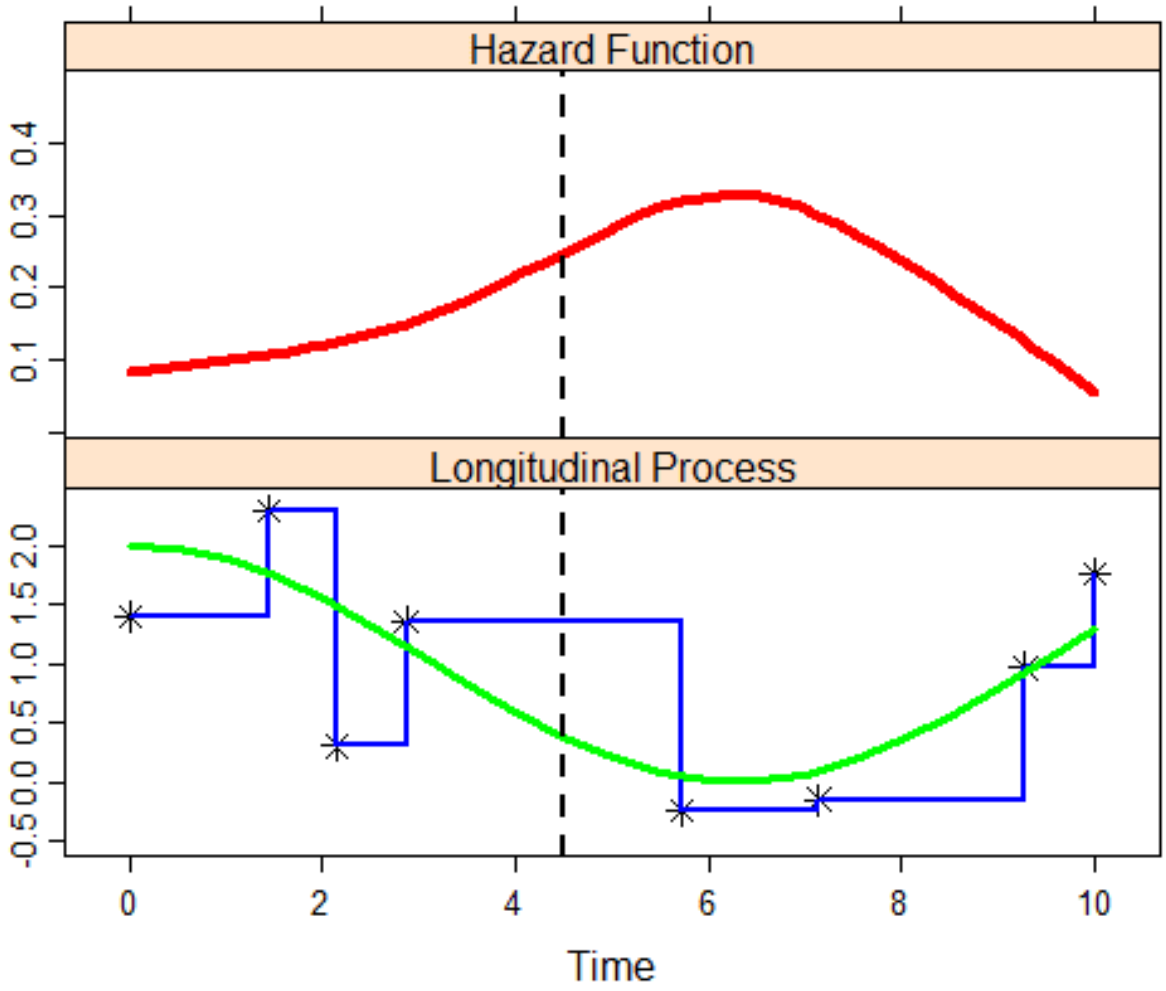
$$p(y_i, T_i, \delta_i) = \int p(y_i \mid b_i)\{h(T_i \mid b_i)^{\delta_i} S(T_i \mid b_i)\}p(b_i)db_i$$

- b_i a vector of random effects that explains the interdependencies
- $p(.)$ density function; $S(.)$ survival function
- δ_i: Event indicator, i.e., equals 1 for true events

The major assumption of this joint model is that there is full conditional independence suggesting that the random effects explain all interdependencies. The longitudinal outcome is independent of the time-to-event outcome and the repeated measurements in the longitudinal outcome are independent of each other.

Estimation of the joint model can be done through both maximum likelihood estimator and Bayesian Statistical approaches. For the maximum likelihood estimator optimization algorithms such as EM, Newton-type, hybrids are implemented to maximize the approximated log-likelihood. Under the Bayesian approach, the inference is based on the posterior distribution. Proposal distributions are then obtained from the separate fits of the two submodels. Then the inference proceeds with output from Markov Chain Monte Carlo methods.

The Joint model is a valuable model because it allows for endogenous time-varying covariates to be considered in predicting the survival outcome. Endogenous time-varying covariates are defined as a variable in which the future path of the covariate up to any time t > s is affected by the occurrence of an event at time point s. The Cox regression model only takes into consideration the baseline covariate values and does not account for the longitudinal change in the covariate.



The upper grid shows the hazard function process and assess how the instantaneous risk of an event changes over time. The bottom grid shows the longitudinal process, where the asterisk illustrates the observed longitudinal responses and the green line the underlying longitudinal process. Joint models assume that the hazard function at any time point t, denoted by the vertical dashed line, is associated with the value of the longitudinal process (green line) at the same time point. Estimation of the model is based on the joint distribution of the two outcomes and can be done either under maximum likelihood or under a Bayesian approach. The framework of joint models can be used to account for both endogenous time-varying covariates and non-random dropout. (Rizopoulos, 2012)

Statistical Analysis

Table 1. Demographic and Biomarker Summary by Treatment

Characteristic	**placebo**, N = 154 (49%)	**D-penicil**, N = 158 (51%)	**p-value**
Status			0.55
alive	68 (44%)	75 (47%)	
transplanted	17 (11%)	12 (7.6%)	
dead	69 (45%)	71 (45%)	
Sex			0.42
male	15 (9.7%)	21 (13%)	
female	139 (90%)	137 (87%)	
Ascites			0.57
no	144 (94%)	144 (91%)	
yes	10 (6.5%)	14 (8.9%)	
Hepatomegaly			0.088
no	67 (44%)	85 (54%)	
yes	87 (56%)	73 (46%)	
Spiders			0.98
no	109 (71%)	113 (72%)	
yes	45 (29%)	45 (28%)	
Edema			0.76
edema despite diuretics	10 (6.5%)	11 (7.0%)	
edema no diuretics	24 (16%)	20 (13%)	
no edema	120 (78%)	127 (80%)	
Serum Bilirubin (mg/dl)	1.3 (0.7, 3.6)	1.4 (0.8, 3.2)	0.79
Serum Cholesterol (mg/dl)	304 (254, 377)	316 (248, 417)	0.54
Albumin (gm/dl)	3.54 (3.34, 3.78)	3.56 (3.21, 3.83)	0.95
Alkaline Phosphatase (U/liter)	1283 (922, 1950)	1214 (841, 2028)	0.81
SGOT (U/ml)	117 (84, 152)	112 (77, 152)	0.46
Platelets (ml/1000)	260 (207, 322)	255 (190, 322)	0.45
Prothrombin Time (s)	10.60 (10.00, 11.38)	10.60 (10.03, 11.00)	0.64
Histologic Stage of Disease			0.20
1	4 (2.6%)	12 (7.6%)	
2	32 (21%)	35 (22%)	
3	64 (42%)	56 (35%)	
4	54 (35%)	55 (35%)	
Categorized Age			0.058
20-39	30 (19%)	28 (18%)	
40-59	104 (68%)	93 (59%)	
60+	20 (13%)	37 (23%)	

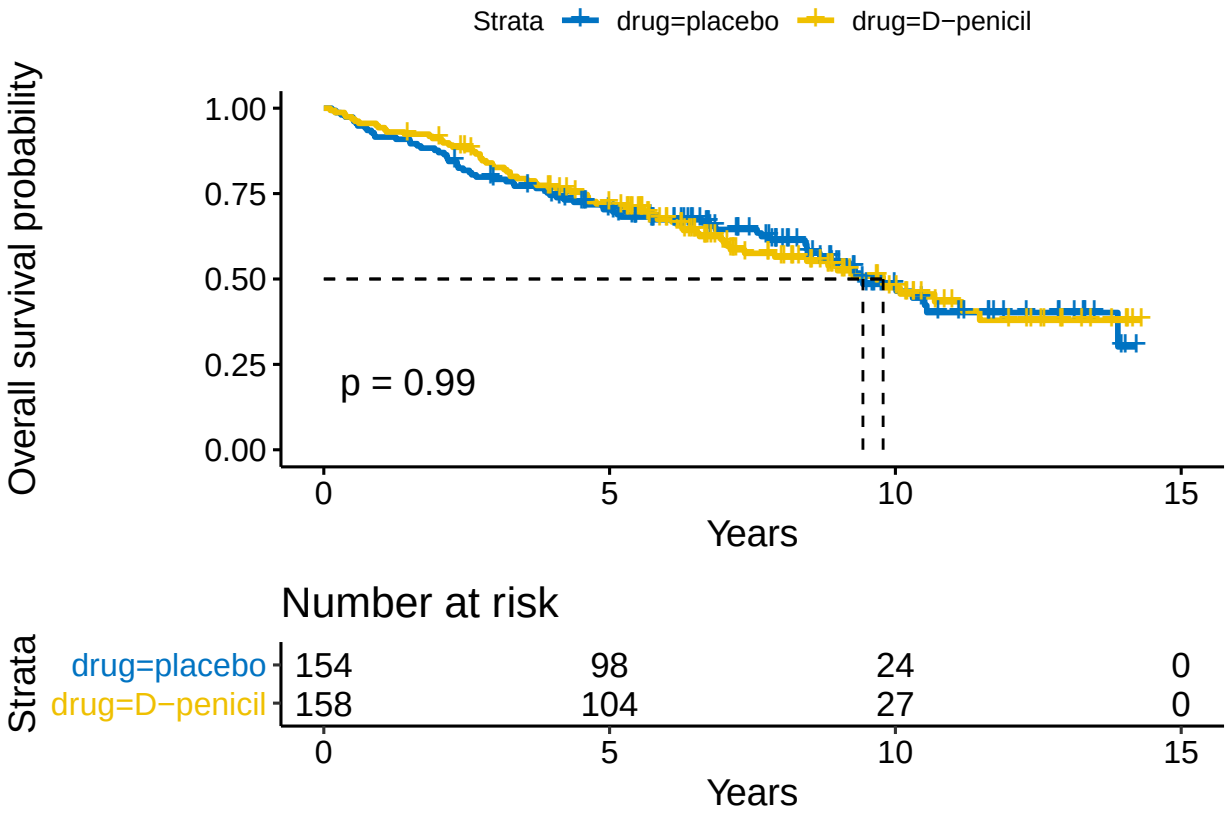
Categorical Variables Statistic: n (%), Continuous Variables Statistic: median (IQR)
Statistical tests performed: chi-square test of independence

The descriptive statistics for the data shown in **table 1** further supports prior findings of PBC being more prevalent in females than in men, with around 90% of the patients in this study being female. Many of the patients in this study were in the later histologic periods of PBC, with around 39% of all patients being in the 3rd histologic stage and 35% of all patients being in the 4th histologic stage of disease. Table 1 also highlights that the largest observed age category for this study was between 40 – 59 years old, having 64% of the study population.

Kaplan-Meier

The Kaplan–Meier estimator, also known as the product limit estimator, is a non-parametric statistic used to estimate the survival function from time-to-event data. **Figure 1** is a Kaplan-Meier curve fit using the `survfit` function from the `survival` package. Transplanted patients in this study were coded as having been censored. This curve assesses the overall survival probabilities of PBC patients grouped by their treatment drug for the nearly 15 year study period. The black dashed line represents median survival time for each strata.

Figure 1. Kaplan-Meier Curve of PBC patients stratified by drug



strata	median
drug=placebo	9.43
drug=D-penicil	9.79

Statistical Test Performed: Log-rank test

The p-value from the long-rank test suggests that there was no significant difference in survival for PBC patients according to the treatment they received. Using the `surv_median()` function, the median survival for patients assigned placebo was 9.43 years and median survival time for patients assigned D-penicil was 9.79 years.

Cox Regression Event Process

The Cox Regression Model assumes a multiplicative effect of covariates on the hazard. They are of the form,

$$h_i(t) = h_0(t)exp(\gamma_1\omega_{i1} + \gamma_2\omega_{i2} + ... + \gamma_p\omega_{ip})$$

where, $h_i(t)$ denotes the hazard of an event for patient i at time t, $h_0(t)$ denotes the baseline hazard, and $\omega_{i1}, ..., \omega_{ip}$ are a set of covariates.

The Cox Regression Model can be extended to evaluate time-dependent covariates using a counting process,

$$h_i(t|N_i(t), \omega_i) = h_0(t)R_i(t)exp\{\gamma^T\omega_i + \alpha y_i(t)\}$$

where, $N_i(t)$ is a counting process which counts the number of events for subject i by time t, $h_i(t)$ denotes the intensity process for $N_i(t)$, $R_i(t)$ denotes the at risk process ('1' if subject i still at risk at t), and $y_i(t)$ denotes the value of the time-varying covariate at t.

The model output in **Table 2** was generated in R from the `finalfit` package using the `finalfit()` function. The output was generated using the following formula,

$$h_i(t) = h_0(t)exp\{\gamma_1 drug + \gamma_2 sex_i + \gamma_3 age_i + \gamma_4 log(albumin)_i + \gamma_5 log(SGOT)_i + \gamma_6 log(prothrombin)_i +$$

$$\gamma_7 histologic_i + \gamma_8 edema_i + \gamma_9 log(serChol)_i + \gamma_{10} log(serBilir)_i\}$$

Table 2. Multivariate Extended Cox Proportional-Hazards Analysis of PBC Survival

Dependent: Surv(start, stop, status2)		`HR (95% CI, p-value)`
drug	placebo	-
	D-penicil	0.85 (0.63-1.15, p=0.289)
sex	female	-
	male	1.39 (0.95-2.05, p=0.091)
age		1.04 (1.03-1.06, **p<0.001**)
log(albumin)		0.58 (0.19-1.79, p=0.342)
log(SGOT)		1.54 (1.02-2.32, **p=0.041**)
log(prothrombin)		1.00 (0.20-5.09, p=0.997)
histologic	1	-
	2	1.46 (0.43-4.93, p=0.545)
	3	1.61 (0.50-5.25, p=0.428)
	4	1.69 (0.52-5.49, p=0.386)
edema	No edema	-
	edema no diuretics	1.02 (0.69-1.49, p=0.934)
	edema despite diuretics	1.39 (0.78-2.46, p=0.262)
log(serChol)		1.17 (0.77-1.80, p=0.461)
log(serBilir)		1.65 (1.32-2.06, **p<0.001**)

The cox regression model shown in **table 2** assesses that PBC patients that have taken the drug D-penicil have a 15% decrease in risk of death, however this is not statistically signficant. Male patients have a 39% increase in risk of death compared to female patients, at any given time. With one year increase in age the risk of death increases by 4%. A 2.718 times increase in SGOT levels increases the risk of death by 54%. Serum bilirubin has the largest significant association with the risk for death, with a 2.718 times increase in serum bilirubin levels ncreases the risk of death by 65%.

Longitudinal Process of Log Serum Bilirubin

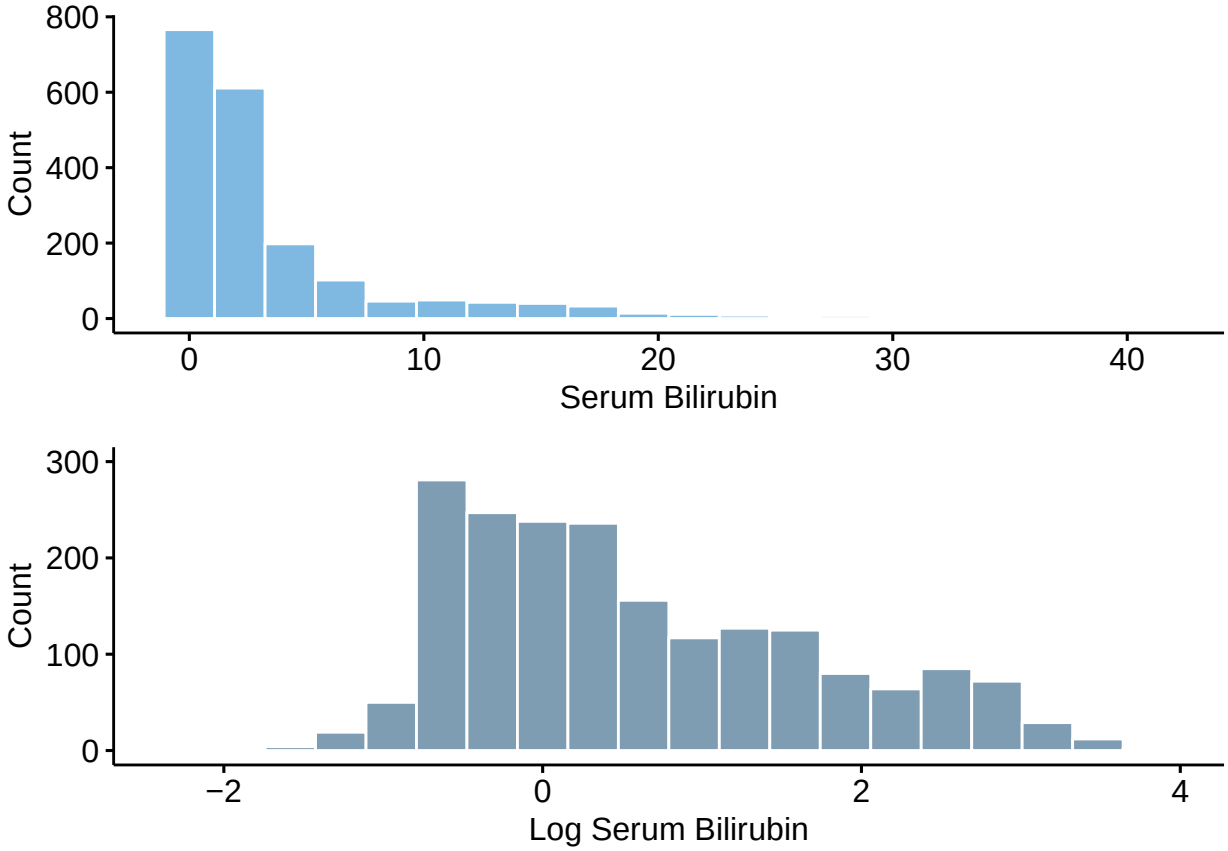
A mixed-effects model is a model used to explain a dependent variable from longitudinal data incorporating both fixed and random effects. This model is necessary in order to capture the evolution of the marker over time for each patient. The model is generally written in the form,

$$y_{ij} = (\beta_0 + b_{i0}) + (\beta_1 + b_{i1})t_{ij} + \epsilon_{ij}$$

where βs are known as the *fixed effects* and bs are known as the *random effects*

Prior research suggests that Serum Bilirubin may be associated with death in PBC patients. (Shapiro, 1979). Levels of bilirubin can predict outcomes either liver transplantation or death of patients with PBC and can be used as surrogate end points in therapy trials.(Lammers, 2014) With this in mind the outcome variable of interest for the longitudinal process is Serum Bilirubin.

Figure 2. Distribution of Serum Bilirubin and log Serum Bilirubin



The highly positively skewed distribution shown in **figure 2** for Serum Bilirubin warranted for normalization. Moving on the log serum bilirubin will be the primary outcome of interest for the longitudinal process.

The optimal mixed-effects model shown in **table 3** was generated in `R` from the `nlme` package using the `lme()` function. The following mixed-effects model formula was used,

$$y_i(t) = \beta_0 + \beta_1(t_i) + \beta_2(sex_i) + \beta_3(age_i) + \beta_4(log(prothrombin)_i) + \beta_5(histologic_i) + \beta_6(log(SGOT)_i) + \beta_7(log(platelets_i) + \beta_8(alkaline_i) + \beta_9(albumin_i) + \beta_{10}(ascites_i) + \beta_{11}(hepatomegaly_i) + \beta_{12}(edema_i) + \beta_{13}(serChol_i) + \beta_{14}(spiders_i) + b_{i0} + \beta_{i1}(t_{ij}) + \epsilon_{ij}$$

Table 3. Linear Mixed-Effects model results of log Serum Bilirubin of PBC patients

Parameter	Value	S.E.	DF	t value	p value
(Intercept)	-6.001	0.625	794	-9.61	**0.000**
year	0.044	0.010	794	4.65	**0.000**
sexfemale	-0.293	0.118	302	-2.49	**0.013**
ascitesYes	0.132	0.057	794	2.33	**0.020**
hepatomegalyYes	0.104	0.035	794	3.02	**0.003**
spidersYes	0.175	0.039	794	4.51	**0.000**
edemaedema no diuretics	0.188	0.044	794	4.26	**0.000**
edemaedema despite diuretics	0.280	0.068	794	4.13	**0.000**
log(serChol)	0.495	0.054	794	9.18	**0.000**
albumin	-0.115	0.041	794	-2.83	**0.005**
alkaline	0.000	0.000	794	1.92	0.055
log(SGOT)	0.587	0.044	794	13.45	**0.000**
log(prothrombin)	0.894	0.169	794	5.30	**0.000**
histologic2	-0.083	0.101	794	-0.82	0.413
histologic3	0.113	0.103	794	1.10	0.273
histologic4	0.234	0.108	794	2.17	**0.030**
log(platelets)	-0.155	0.052	794	-2.99	**0.003**

Random Effects		
	SD	**Corr**
Intercept	0.61	
year	0.08	-0.12
residual	0.32	

Table 3 provides information for the coefficient estimates, standard errors and p-values for the mixed effects model chosen with the lowest AIC value. Like the cox regression performed earlier, the log of prothrombin has the highest significant influence on the log serum bilirubin of all the covariates in this model. However, it is important to note that the standard error is also relatively high with 0.17. The p-value of the year variable being ($p < 0.01$) indicates that there is a significant change in log serum bilirubin as time passes. In Table 3 the sex variable having a p-value of ($p = 0.01$) assesses that female PBC patients on average have significantly lower serum bilirubin levels than male PBC patients. Similarly other variables such as log of platelets and log of serum Cholesterol show a significant negative relationship with log serum bilirubin. The key takeaway from the random effects output is that there is a negative correlation between time and the intercept. When a PBC patient's intercept increases by one unit of standard deviation, that patient's slope would decrease by 0.12 standard deviations.

Joint Model

The joint model output in **table 4** was generated in `R` from the `JMbayes` package using the `jointModelBayes()` function. Here we are interested in the association between the repeated measures of serum bilirubin and survival. Limitations of the extended cox model is that it only applies to exogenous time depending covariates, and treating endogenous covariates as exogenous can be very problematic. The joint Model is appropriate because it combines the longitudinal process and event process to more accurately assess the association between repeated measures and the risk of death. The following output is a result from the Bayesian estimation approach.

We fit the linear mixed model,

$$y_i(t) = \beta_0+\beta_1(t_{ij})+\beta_2(sex_i)+\beta_3(age_i)+\beta_4(log(prothrombin)_i)+\beta_5(histologic_i)+\beta_6(log(SGOT)_i)+\beta_7(log(platelets_i)+$$

$$\beta_8(alkaline_i)+\beta_9(albumin_i)+\beta_{10}(ascites_i)+\beta_{11}(hepatomegaly_i)+\beta_{12}(edema_i)+\beta_{13}(serChol_i)+\beta_{14}(spiders_i)+b_{i0}+$$

$$b_{i1}(t_{ij}) + \epsilon_{ij}$$

and for the cox model we included the time-dependent covariate that evaluates the true underlying profile of the serum bilirubin levels as estimated from the longitudinal model,

$$h_i(t) = h_0(t)exp\{\gamma_1 drug + \gamma_2 sex_i + \gamma_3 age_i + \gamma_4 log(albumin)_i + \gamma_5 log(SGOT)_i + \gamma_6 log(prothrombin)_i+$$

$$\gamma_7 histologic_i + \gamma_8 edema_i + \gamma_9 log(serChol)_i + \alpha m_i(t)\}$$

Table 4. Event output from Joint Survival and Longitudinal Model using Bayesian Method

Parameter	Value	S.E.	95% CI	p value
drugD-penicil	-0.391	0.092	(-0.837, 0.059)	0.110
sexfemale	-0.059	0.084	(-0.506, 0.559)	0.762
age	0.050	0.003	(0.029, 0.068)	**0.000**
log(prothrombin)	0.486	0.229	(-1.901, 1.64)	0.451
histologic2	0.701	0.297	(-1.101, 3.076)	0.507
histologic3	0.610	0.306	(-1.233, 2.924)	0.610
histologic4	0.757	0.341	(-1.245, 3.248)	0.499
log(SGOT)	0.127	0.081	(-0.367, 0.607)	0.734
log(albumin)	-1.920	0.309	(-3.532, -0.01)	**0.049**
edemaedema no diuretics	0.631	0.195	(-0.055, 1.435)	0.117
edemaedema despite diuretics	1.718	0.188	(0.638, 2.686)	**0.000**
log(serChol)	-0.738	0.115	(-1.613, -0.065)	**0.025**
Assoct	1.333	0.023	(1.029, 1.69)	**0.000**
tauBs	336.716	20.002	(61.586, 1004.19)	NA

When we account for the longitudinal effects, we can see that some predictors such as SGOT that was once significant at the 5% level of signficance and sex that was significant at the 10% level of signficance are now insignificant. This suggests that the cox model overestimated the importance of these variables in predicting survival. We also notice that covariates such as albumin and serum cholesterol are now significant. This implies that the extended cox model was underestimating the importance of these variables. The `Assoc` parameter measures the association between m_t and risk for death. Each 1 unit increase in serum bilirubin corresponds to a 3.781-fold increase in the risk for death. This association was highly significant suggesting that there is stronger evidence for there to have been nonrandom dropout. When we account

for the longitudinal process in the joint model, the event process assesses that D-penicil results in a 1.48-fold increase of the risk of death, but this treatment variable is still insignificant.`tauBs` is the smoothing parameter for the penalized B-spline approximation of the log baseline hazard function.

Conclusion

Primary biliary cholangitis (PBC) is an autoimmune disease that has a very high association with an individual's log serum bilirubin levels. The traditional method in assessing survival probabilities of diseases such as PBC is cox regression. However, one of the limitations of this method is that this model can not account for endogenous time-varying variables. A joint model of both survival and longitudinal processes helps to account for these endogenous variables and give a more accurate estimation of the true association size of markers.

While sex was significant (alpha = 10%) in the cox model, the results from the joint model suggests that when we account for the time-varying variables sex no longer has a signficant difference in risk for death. The only variables that are signficant when accounting for the longitudinal process are log serum cholesterol, log albumin, age and when the patient has edema present despite diuretic therapy. We also notice that the results from the joint model depict a large drop in the p-value of the drug D-penicil. While this variable was still insignificant, this drop when accounting for the longitudinal process sheds new potential for further research.

When considering the joint model for prediction purposes, it was not as effective in predicting log serum bilirubin levels of held out final observation records compared to the random mixed-effects model alone. Based on root mean squared error of the mixed-effects model, on average the difference between the predicted and actual log Serum Bilirubin levels was 0.26 mg/dl. For the Joint Model, on average the difference between the predicted and actual log Serum Bilirubin levels was 0.95 mg/dl. This suggests that there may be some limitations that come with the joint model. However, this limitation in predictive ability comes with a tradeoff with more accurate interpretability of the survival risk.

References

Pandit S, Samant H. Primary Biliary Cholangitis (Primary Biliary Cirrhosis) [Updated 2019 Feb 28]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing;

Purohit T, Cappell MS. Primary biliary cirrhosis: Pathophysiology, clinical presentation and therapy. World J Hepatol. 2015;7(7):926–941. doi:10.4254/wjh.v7.i7.926

Rizopoulos, D. (2012). Joint models for longitudinal and time-to-event data: with applications in R. Boca Raton, FL: Chapman & Hall/CRC.

Shapiro, J. M., Smith, H., & Schaffner, F. (1979). Serum bilirubin: a prognostic factor in primary biliary cirrhosis. Gut, 20(2), 137–140. doi:10.1136/gut.20.2.137

Lammers W.J., Van Buuren H.R., Hirschfield G.M., Janssen H.L.A., Invernizzi P., Mason A.L., Ponsioen C.Y., (. . .), Global PBC Study Group (2014),pp. 1338-1349.e5