

Primary Biliary Cholangitis (PBC) Survival Analysis

PSTAT 175, Spring 2025

Richard Zhang, Randy Van Waes, Vincent Agrella

Introduction:

In this new era of body positivity in society, it is important for many to be aware of their own health. An aspect of measuring one's health is identifying for any chronic/genetic illness that one might have. Primary biliary cholangitis (PBC) is one of these health issues. Where it is a chronic, progressive autoimmune liver disease that affects an estimated 1.76 to 14.60 individuals per 100,000 people globally. It predominantly affects women more than men. In the United States, the prevalence is higher, with 33.5 to 57.8 per 100,000 women and 7.7 to 15.4 per 100,000 men diagnosed. PBC is characterized by inflammation and gradual destruction of the bile ducts, which are vital for transporting bile from the liver. If left untreated, PBC can lead to liver failure and death. Given its progressive nature and higher frequency in the United States, understanding the course of PBC is crucial for optimizing treatment and improving patient outcomes.

Research Question

In this study we will utilize the Mayo Clinic's Primary Biliary Cholangitis dataset, available in R's survival package. The dataset includes detailed clinical information on 418 patients, encompassing 20 variables that allow for a comprehensive analysis of the progression of PBC. With Mayo Clinic's data, we will be answering the following research question:

- How effective is D-penicillamine in slowing the progression of primary biliary cholangitis (PBC), and what role do other confounding clinical factors (e.g., cholesterol levels, histologic stage, etc.) play in influencing survival time in PBC patients?

By uncovering these questions, our study aims to produce a comprehensive understanding of relevant confounding factors that influence the progression and survival of primary biliary cholangitis. To help us find the most effective treatment strategies for people with PBC.

Key Covariates for our Research:

In our research, we use the following variables/covariates in our Data matrix to answer our questions above:

- time: Patient Follow-up time (days)
- status: Health status of patient (0= censored, 1= liver transplant, 2= dead)
- age: Patient age (years)
- sex: Patient gender at birth (m=male, f=female)
- bili: Bilirubin produced when red blood cell breakdown (mg/dl)
- albumin: Albumin protein produced (g/dl)
- copper: Hepatic copper concentration (mg/g)
- protime: Prothrombin time, time blood takes to clot (sec)
- stage: Histologic stage of PBC (1= minimal fibrosis, 2= moderate fibrosis, 3 = advanced fibrosis, 4= cirrhosis)
- trt: Treatment group (1= placebo, 2= D-penicillamine)
- ast: Aspartate Aminotransferase, liver enzyme to identify liver inflammation or damage (U/L)

Understanding Our Data

To help us find out how effective Mayo Clinic's treatment of PBC is, we must first load and observe a few data entries of our dataset. This will help us spot some noticeable features of our data as well as see if there are null entries that we might have to remove. As the inclusion of null values can lead to biased results.

Loading Our Dataset and the First Few Entries of our Data:

```
##      time status trt      age sex ascites hepato bili chol albumin copper   ast
## 1   400      2   1 58.76523   f      1      1 14.5  261   2.60   156 137.95
## 2  4500      0   1 56.44627   f      0      1  1.1  302   4.14    54 113.52
## 3  1012      2   1 70.07255   m      0      0  1.4  176   3.48   210  96.10
## 4  1925      2   1 54.74059   f      0      1  1.8  244   2.54    64  60.63
## 5  1504      1   2 38.10541   f      0      1  3.4  279   3.53   143 113.15
## 6  2503      2   2 66.25873   f      0      1  0.8  248   3.98    50  93.00
##      trig protime stage
## 1   172    12.2     4
## 2    88    10.6     3
## 3    55    12.0     4
## 4    92    10.3     4
## 5    72    10.9     3
## 6    63    11.0     3
```

When looking over the first few data entries of our data, we can see something significant. In which the entries we have under the *status* variable. As our *status* variable has 3 different categories. With 0 being censored observations, 1 being patients who got a liver transplant, and 2 being patients who died. This is significant because this tells us that there are competing risks in our dataset. We will get into more detail when we begin getting initial insight on the effectiveness of our treatment.

Last Few Data Entries:

```
##      time status trt      age sex ascites hepato bili chol albumin copper ast
## 413  989      0  NA 35.00068   f      NA      NA  0.7   NA   3.23    NA  NA
## 414  681      2  NA 67.00068   f      NA      NA  1.2   NA   2.96    NA  NA
## 415 1103      0  NA 39.00068   f      NA      NA  0.9   NA   3.83    NA  NA
## 416 1055      0  NA 56.99932   f      NA      NA  1.6   NA   3.42    NA  NA
## 417  691      0  NA 58.00137   f      NA      NA  0.8   NA   3.75    NA  NA
## 418  976      0  NA 52.99932   f      NA      NA  0.7   NA   3.29    NA  NA
##      trig protime stage
## 413  NA    10.8     3
## 414  NA    10.9     3
## 415  NA    11.2     4
## 416  NA     9.9     3
## 417  NA    10.4     3
## 418  NA    10.6     4
```

We can notice that we will need to be aware of cleaning our data properly as we have many entries that have “NA” values. Which can affect our results for our research. Hence we will need to remove entries that have “NA” values.

Intial Insights on the Effectiveness of Our Treatment

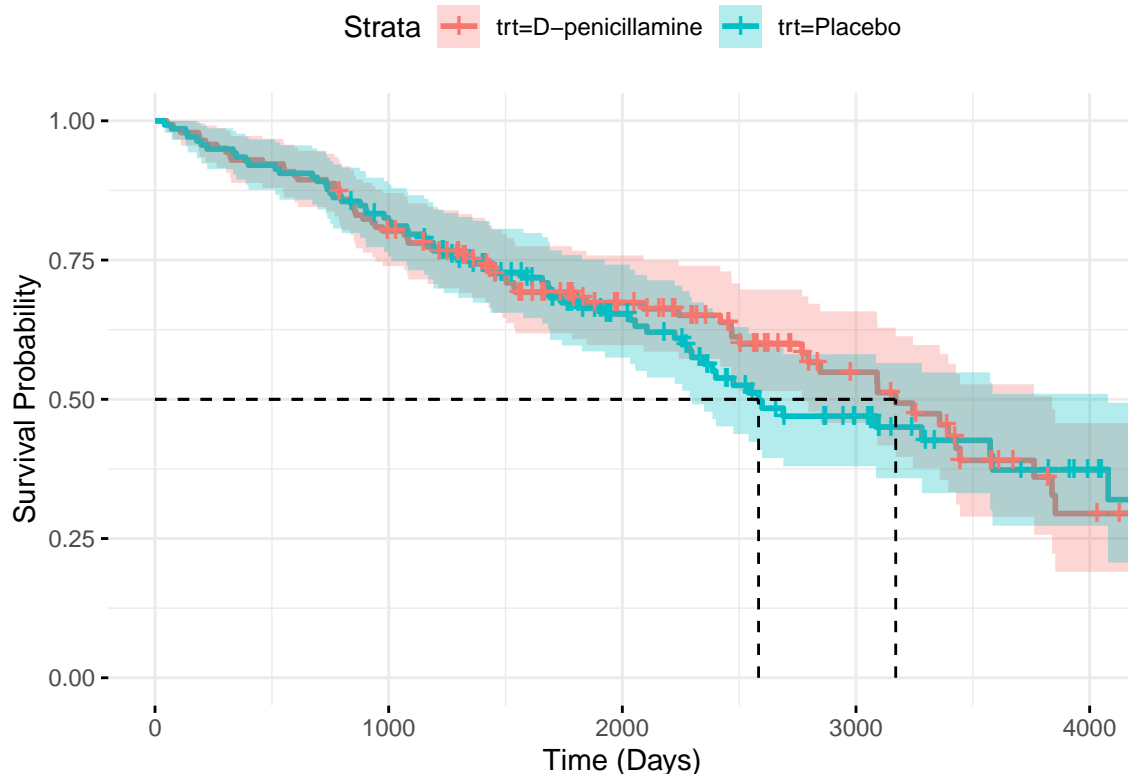
As discussed when we were observing a few entries of our dataset, we notice that there are competing risks in our data. In our situation, where we are trying to see the effectiveness of our treatment, we will be utilizing a composite endpoint approach. In which we will be treating death and liver transplant as a adverse event that patients could experience. As a patient receiving D-penicillamine and later needing a liver transplant, would indicate that our treatment is ineffective. This is also the same when a patient dies. Where there is no benefit in taking the drug.

Fitting Survival Function Based on Treatment

```
## Call: survfit(formula = pbc.surv ~ trt, data = pbc)
##
##           n events median 0.95LCL 0.95UCL
## trt=1 138      66  2583    2297    4079
## trt=2 142      64  3170    2769    3839
```

Based on the Mayo Clinic dataset, our PBC patients who took the placebo have a median survival time of 2,583 days. This means that 50% of the patients who took the placebo drug have died or have experienced a liver transplant after 2,583 days. We can also notice that patients who received D-penicillamine have a median survival time of 3,170 days. Overall, 138 patients took the placebo and 142 took D-penicillamine. With 66 of the placebo patients and 64 of the treatment patients who have died or have received a liver transplant. With the results that we get we can see that the D-penicillamine treatment is working in prolonging the need to liver transplant and life. However, we must note that we are unsure this is statistically significant, as further testing is required. To get a better understanding of the patient's survival probability, we will utilize a Kaplan-Meier Survival curve.

Kaplan Meier Survival Curve Based on Treatment



The KM Plot describes the behavior of the treatment group and the placebo group throughout the study, approximately 5000 days. Notice how the two lines in our plot are basically following the same paths. As they intersect with each other and later start to deviate a little when time passes. This tells us that there is little to no difference in survival between the two groups. Which also means that it does not matter if you take the placebo or D-penicillamine they will yield similar results.

Fiting a Cox Proportional Hazard Model Controlling Treatment

```
## Call:
## coxph(formula = Surv(pbc$time, pbc$status > 0) ~ trt, data = pbc)
##
##      n= 280, number of events= 130
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## trtD-penicillamine -0.08236   0.92094  0.17559 -0.469   0.639
##
##              exp(coef) exp(-coef) lower .95 upper .95
## trtD-penicillamine   0.9209      1.086   0.6528   1.299
##
## Concordance= 0.509  (se = 0.024 )
## Likelihood ratio test= 0.22  on 1 df,   p=0.6
## Wald test               = 0.22  on 1 df,   p=0.6
## Score (logrank) test = 0.22  on 1 df,   p=0.6
```

Based on the results produced from our Cox Proportional Hazard Model, we can notice that we get a *coef* value of -0.08236 and a *exp(coef)* value of 0.92094 . Which tells us that patients who have taken D-penicillamine are about 7.9% less at risk of dying or needing a liver transplant. With a 95% hazard ratio confidence interval between $(0.6528, 1.299)$. However, we get a p-value of 0.639 , which indicates that the effect of D-penicillamine is not statistically significant. When applying a likelihood ratio test, we get a p-value of 0.6 . Which indicate that the addition of our treatment covariate to our Cox proportional hazard model does not improve our survival prediction.

Model Fitting

Since the effect of D-penicillamine treatment was not significant, we will try fitting an improved model that incorporates the treatment variable to assess whether it can reveal a significant effect. To accomplish this, we will be utilizing the AIC forward selection method to find a improved model.

```
## Start:  AIC=1299.65
## Surv(pbc$time, pbc$status > 0) ~ trt
##
##           Df      AIC
## + bili      1 1225.9
## + copper     1 1244.2
## + stage      1 1251.7
## + albumin    1 1256.6
## + ascites    1 1259.3
## + hepato     1 1270.2
## + protime    1 1281.2
## + ast        1 1283.8
## + chol       1 1288.9
## + trig       1 1288.9
## + age        1 1292.4
## + sex        1 1298.4
## <none>       1299.7
```

```

##
## Step: AIC=1225.94
## Surv(pbc$time, pbc$status > 0) ~ trt + bili
##
##           Df      AIC
## + stage    1 1188.9
## + copper    1 1198.5
## + albumin   1 1200.1
## + ascites   1 1207.0
## + hepato    1 1214.2
## + protime   1 1217.5
## + age       1 1220.9
## + sex       1 1223.5
## + ast       1 1224.8
## <none>      1225.9
## + trig      1 1227.9
## + chol      1 1227.9
##
## Step: AIC=1188.88
## Surv(pbc$time, pbc$status > 0) ~ trt + bili + stage
##
##           Df      AIC
## + copper    1 1169.7
## + albumin   1 1176.4
## + ascites   1 1182.5
## + ast       1 1183.9
## + sex       1 1185.9
## + protime   1 1185.9
## <none>      1188.9
## + age       1 1189.0
## + trig      1 1190.2
## + hepato    1 1190.4
## + chol      1 1190.6
##
## Step: AIC=1169.7
## Surv(pbc$time, pbc$status > 0) ~ trt + bili + stage + copper
##
##           Df      AIC
## + albumin   1 1160.3
## + protime   1 1166.9
## + ast       1 1168.0
## + ascites   1 1168.5
## <none>      1169.7
## + trig      1 1170.1
## + age       1 1170.4
## + hepato    1 1171.0
## + sex       1 1171.2
## + chol      1 1171.4
##
## Step: AIC=1160.26
## Surv(pbc$time, pbc$status > 0) ~ trt + bili + stage + copper +
##      albumin
##
##           Df      AIC

```

```

## + protime 1 1158.0
## + ast 1 1159.8
## + trig 1 1160.0
## <none> 1160.3
## + sex 1 1160.3
## + age 1 1161.3
## + chol 1 1162.0
## + hepato 1 1162.0
## + ascites 1 1162.0
##
## Step: AIC=1158.03
## Surv(pbc$time, pbc$status > 0) ~ trt + bili + stage + copper +
## albumin + protime
##
## Df AIC
## + ast 1 1156.6
## <none> 1158.0
## + trig 1 1158.1
## + sex 1 1158.3
## + age 1 1159.5
## + chol 1 1159.5
## + hepato 1 1159.8
## + ascites 1 1160.0
##
## Step: AIC=1156.59
## Surv(pbc$time, pbc$status > 0) ~ trt + bili + stage + copper +
## albumin + protime + ast
##
## Df AIC
## <none> 1156.6
## + trig 1 1156.9
## + sex 1 1157.1
## + age 1 1157.5
## + chol 1 1158.5
## + hepato 1 1158.5
## + ascites 1 1158.5

```

When applying the AIC forward selection method, we notice that the following covariates: *trig*, *sex*, *age*, *chol*, *hepato*, and *ascites* are not used to give us our improved model.

Cox Proportional Hazard Model of Our Improved Model

```

## Call:
## coxph(formula = Surv(pbc$time, pbc$status > 0) ~ trt + bili +
## stage + copper + albumin + protime + ast, data = pbc)
##
## n= 280, number of events= 130
##
##      coef exp(coef) se(coef)      z Pr(>|z|)
## trtD-penicillamine -0.1696043 0.8439987 0.1811707 -0.936 0.34919
## bili 0.0924352 1.0968421 0.0178251 5.186 2.15e-07 ***
## stage 0.5270171 1.6938721 0.1317678 4.000 6.35e-05 ***
## copper 0.0036887 1.0036955 0.0008885 4.151 3.30e-05 ***
## albumin -0.7484798 0.4730852 0.2388946 -3.133 0.00173 **
## protime 0.2276050 1.2555893 0.0954061 2.386 0.01705 *

```

```
## ast          0.0031836  1.0031886  0.0016449  1.935  0.05294 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## trtD-penicillamine    0.8440    1.1848    0.5917    1.2038
## bili                 1.0968    0.9117    1.0592    1.1358
## stage                1.6939    0.5904    1.3083    2.1930
## copper                1.0037    0.9963    1.0019    1.0054
## albumin              0.4731    2.1138    0.2962    0.7556
## protime              1.2556    0.7964    1.0414    1.5138
## ast                  1.0032    0.9968    1.0000    1.0064
##
## Concordance= 0.814 (se = 0.018 )
## Likelihood ratio test= 155.3 on 7 df,  p=<2e-16
## Wald test              = 170.2 on 7 df,  p=<2e-16
## Score (logrank) test = 227.2 on 7 df,  p=<2e-16
```

When looking at our improved model, a notable thing has occurred. Which is the fact that our treatment covariate become more statistically significant with a p-value of 0.34919 compared to our previous p-value of 0.639. However, this p-value is still not statistically significant enough for us to make a confident conclusion of D-penicillamine having a positive effect on slowing the progression of PBC. We can also observe that all of our covariates' effects besides our treatment are statistically significant. With *bili*, *stage*, and *copper* being the most statistically significant. As *bili*, *stages*, *copper*, *protime*, and *ast* increases, it leads to an decrease in a patient's survival rate. Whereas *albumin* increases, it decreases the risk of a patient dying or receiving a liver transplant. Based on the application of a likelihood ratio test on our model, we have a p-value of $< 2 \times 10^{-16}$. Which indicate that the inclusions of our covariates significantly impact the improvement of our model.

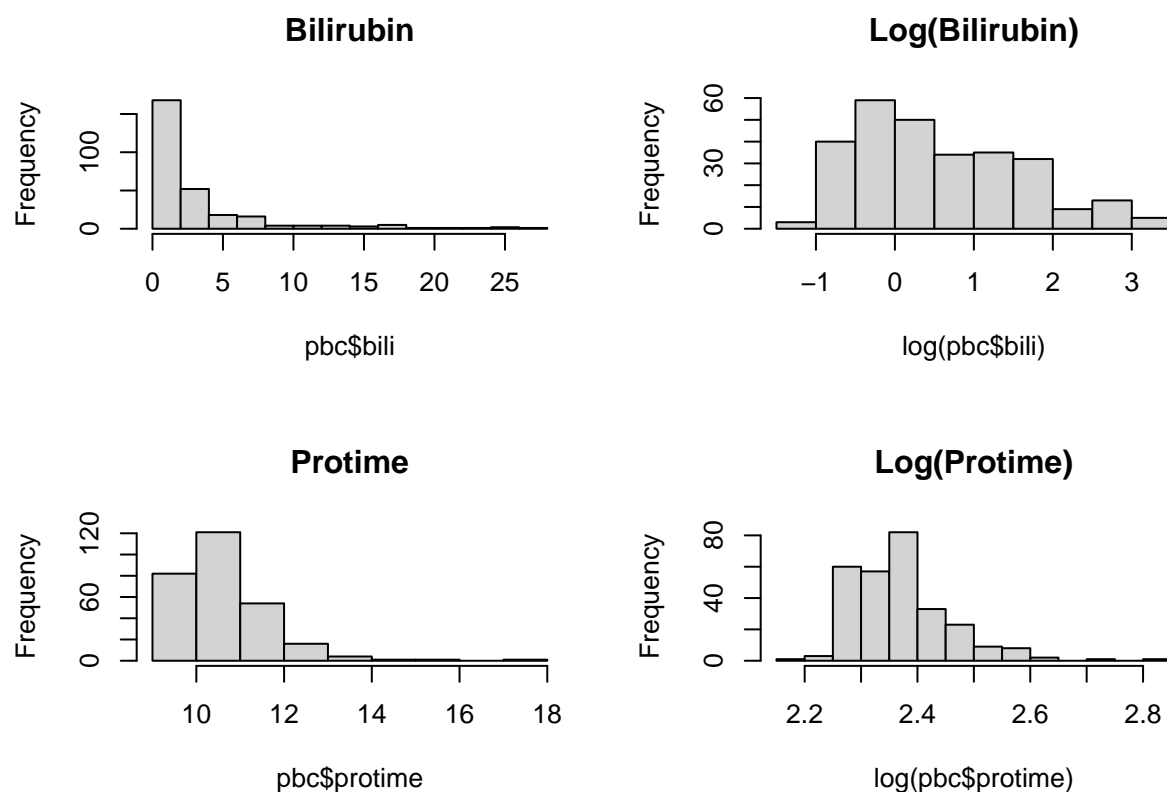
Model Assumptions

```
##          chisq df      p
## trt      0.0918  1 0.762
## bili     6.3714  1 0.012
## stage    3.3075  1 0.069
## copper    0.0448  1 0.832
## albumin  1.3694  1 0.242
## protime  4.7446  1 0.029
## ast      0.9263  1 0.336
## GLOBAL   16.4234  7 0.022
```

Since our Global has a p-value of 0.022, we will need to deal with our violating variables (i.e *bili*, *stage*, and *protime*). This might be due to the fact that we will need to adjust the non-linearity of our *bili* and *protime*. To test this we will be observing the distributions of *bili* and *protime*. It should be noted that *stage* is a catorgory variable. Which leads to us to stratify *stage*.

Checking Non-Linearity of Our Covariates

To test non-linearity of our *bili* and *protime* we will transform our variables by utilizing the log function.



We notice that our covariates are originally skewed to the right. Applying a log-transform has caused our distributions of *bili* and *protime* to be more symmetric (i.e. more of a bell-shaped curve). This indicates that we are making a non-linearity adjustment on our variables. Which help reduce skewness and make our hazard relationships more linear.

Fitting a Cox PH Model with Newly Log-Transformed Covariates

```
## Call:
## coxph(formula = Surv(time, status > 0) ~ trt + log(bili) + strata(stage) +
##       copper + albumin + log(protime) + ast, data = pbc)
##
## n= 280, number of events= 130
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## trtD-penicillamine -0.0266237  0.9737276  0.1838677 -0.145  0.88487
## log(bili)          0.7510735  2.1192739  0.1202242  6.247 4.18e-10 ***
## copper              0.0026130  1.0026164  0.0009723  2.687  0.00720 **
## albumin            -0.6668089  0.5133441  0.2325912 -2.867  0.00415 **
## log(protime)       2.3951587 10.9699388  1.1785372  2.032  0.04212 *
## ast                0.0003220  1.0003220  0.0017703  0.182  0.85567
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
```

```
## trtD-penicillamine    0.9737    1.02698    0.6791    1.3962
## log(bili)             2.1193    0.47186    1.6744    2.6824
## copper                1.0026    0.99739    1.0007    1.0045
## albumin              0.5133    1.94801    0.3254    0.8098
## log(protime)         10.9699    0.09116    1.0890   110.5037
## ast                  1.0003    0.99968    0.9969    1.0038
##
## Concordance= 0.769 (se = 0.024 )
## Likelihood ratio test= 115.3 on 6 df, p=<2e-16
## Wald test              = 116.5 on 6 df, p=<2e-16
## Score (logrank) test = 134.3 on 6 df, p=<2e-16
```

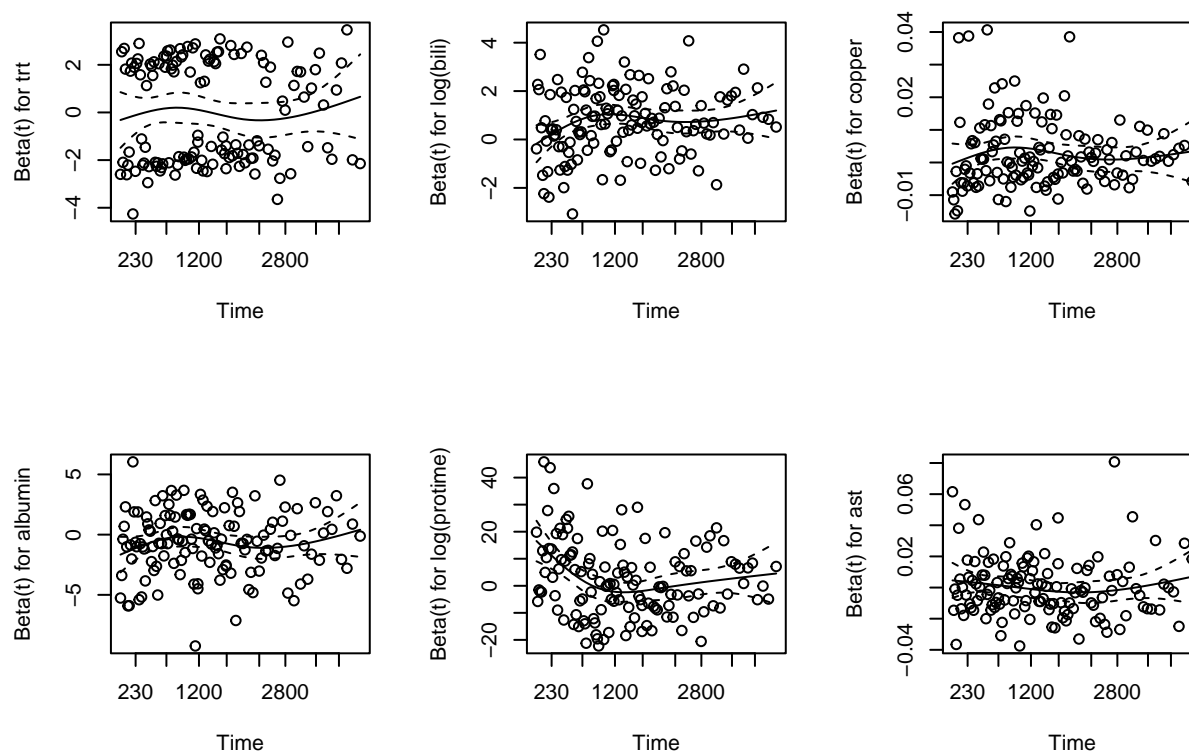
The cox proportional hazard model shows that D-penicillamine is linked to a risk reduction ($coef = -0.02662$), but has no significant treatment effect ($p = 0.88487$). In contrast, log-transformed bilirubin increases risk ($coef = 0.75107$), while having significant effect on treatment ($p < 1 \times 10^{-9}$). On the other hand, copper levels lead to a minor yet significant increase ($coef = 0.002613$, $p = 0.00720$). *Albumin* seems to lower the risk ($exp(coef) = 0.51334$, $p = 0.00415$), while log-transformed prothrombin time increases the risk ($coef = 2.39516$, $p = 0.04212$). Lastly, *ast* has little impact ($coef = 0.000322$), and is also not significant towards treatment ($p = 0.85567$). Overall, the covariates significantly improve the model's predictive ability (likelihood ratio $p < 2 \times 10^{-16}$). Interestingly, notice that our treatment covariate has become way less statistically significant with a p-value of 0.88487.

Checking our Proportional Hazards Assumptions on Our New Model

```
##           chisq df      p
## trt       0.00778  1 0.930
## log(bili) 0.71087  1 0.399
## copper     0.04315  1 0.835
## albumin   0.24163  1 0.623
## log(protime) 3.87116  1 0.049
## ast       0.00447  1 0.947
## GLOBAL    6.02161  6 0.421
```

Using `cox.zph`, we can find that most covariates have a p-value of over our significance level of 0.05, indicating that there is no significant evidence supporting the idea that they change over time. On the other hand, we can find that $\log(\text{protime})$ is significant ($p = 0.049$), implying that there is significant evidence that it changes over time. While this is true, the full model shows no significant evidence for change over time, allowing us to conclude that the assumption for Coxph is satisfied.

Plotting our Residuals



Ideally, we want a linear relationship between residuals, with slope equal to zero. Looking at the y-axis, our covariates Aspartate Aminotransferase (*ast*) and copper have the fewest increments. This means that these two covariates have the most linear residual relationship in our model. However, since the curve is not completely flat, the proportional hazard assumptions may have been slightly violated. This is especially apparent in the other covariates, which experience increasing and decreasing trends, indicating that they change over time. This violates the proportional hazards assumption, as we are looking to identify a treatment that does not falter over time. Therefore, we are less likely to violate the proportional hazard assumptions for Aspartate Aminotransferase (*ast*) and copper than for the other covariates in our model. This is because the hazard ratio of these two variables is 1, and has little effect on treatment. Thus, Aspartate Aminotransferase (*ast*) and copper experience little treatment effects over time, sustaining our proportional hazard assumption.

Finding Confounding Clinic Factors Effect Our Treatment

In order to complete our research question, we must see that there are relevant covariates in our dataset that will have an effect on our treatment. To accomplish this we will be looking into the interactions between our treatment and significant covariates that we got when we were getting the best model of fit.

```
## Call:
## coxph(formula = Surv(pbc$time, pbc$status > 0) ~ trt + trt:log(bili) +
##       trt:strata(stage) + trt:copper + trt:albumin + trt:log(protime) +
##       trt:ast, data = pbc)
##
##      n= 280, number of events= 130
##
```

```

##               coef exp(coef) se(coef)      z
## trtD-penicillamine -1.480e+01 3.746e-07 2.564e+03 -0.006
## trtPlacebo:log(bili) 5.693e-01 1.767e+00 1.731e-01 3.288
## trtD-penicillamine:log(bili) 9.672e-01 2.631e+00 1.785e-01 5.419
## trtPlacebo:strata(stage)stage=2 -1.446e+01 5.271e-07 2.564e+03 -0.006
## trtD-penicillamine:strata(stage)stage=2 NA NA 0.000e+00 NA
## trtPlacebo:strata(stage)stage=3 -1.462e+01 4.486e-07 2.564e+03 -0.006
## trtD-penicillamine:strata(stage)stage=3 NA NA 0.000e+00 NA
## trtPlacebo:strata(stage)stage=4 -1.434e+01 5.893e-07 2.564e+03 -0.006
## trtD-penicillamine:strata(stage)stage=4 NA NA 0.000e+00 NA
## trtPlacebo:copper 2.760e-03 1.003e+00 1.207e-03 2.286
## trtD-penicillamine:copper 1.704e-03 1.002e+00 1.702e-03 1.001
## trtPlacebo:albumin -7.006e-01 4.963e-01 3.265e-01 -2.146
## trtD-penicillamine:albumin -7.538e-01 4.706e-01 3.478e-01 -2.167
## trtPlacebo:log(protime) 2.234e+00 9.340e+00 1.847e+00 1.210
## trtD-penicillamine:log(protime) 2.603e+00 1.351e+01 1.567e+00 1.661
## trtPlacebo:ast 3.045e-03 1.003e+00 2.779e-03 1.096
## trtD-penicillamine:ast -2.043e-03 9.980e-01 2.600e-03 -0.786
## Pr(>|z|)
## trtD-penicillamine 0.99540
## trtPlacebo:log(bili) 0.00101 **
## trtD-penicillamine:log(bili) 6e-08 ***
## trtPlacebo:strata(stage)stage=2 0.99550
## trtD-penicillamine:strata(stage)stage=2 NA
## trtPlacebo:strata(stage)stage=3 0.99545
## trtD-penicillamine:strata(stage)stage=3 NA
## trtPlacebo:strata(stage)stage=4 0.99554
## trtD-penicillamine:strata(stage)stage=4 NA
## trtPlacebo:copper 0.02226 *
## trtD-penicillamine:copper 0.31663
## trtPlacebo:albumin 0.03188 *
## trtD-penicillamine:albumin 0.03023 *
## trtPlacebo:log(protime) 0.22644
## trtD-penicillamine:log(protime) 0.09672 .
## trtPlacebo:ast 0.27324
## trtD-penicillamine:ast 0.43188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95
## trtD-penicillamine 3.746e-07 2.669e+06 0.0000
## trtPlacebo:log(bili) 1.767e+00 5.659e-01 1.2585
## trtD-penicillamine:log(bili) 2.631e+00 3.801e-01 1.8541
## trtPlacebo:strata(stage)stage=2 5.271e-07 1.897e+06 0.0000
## trtD-penicillamine:strata(stage)stage=2 NA NA NA
## trtPlacebo:strata(stage)stage=3 4.486e-07 2.229e+06 0.0000
## trtD-penicillamine:strata(stage)stage=3 NA NA NA
## trtPlacebo:strata(stage)stage=4 5.893e-07 1.697e+06 0.0000
## trtD-penicillamine:strata(stage)stage=4 NA NA NA
## trtPlacebo:copper 1.003e+00 9.972e-01 1.0004
## trtD-penicillamine:copper 1.002e+00 9.983e-01 0.9984
## trtPlacebo:albumin 4.963e-01 2.015e+00 0.2617
## trtD-penicillamine:albumin 4.706e-01 2.125e+00 0.2380
## trtPlacebo:log(protime) 9.340e+00 1.071e-01 0.2500

```

```

## trtD-penicillamine:log(protime)      1.351e+01  7.403e-02   0.6259
## trtPlacebo:ast                       1.003e+00  9.970e-01   0.9976
## trtD-penicillamine:ast               9.980e-01  1.002e+00   0.9929
##                                     upper .95
## trtD-penicillamine                   Inf
## trtPlacebo:log(bili)                 2.4809
## trtD-penicillamine:log(bili)         3.7326
## trtPlacebo:strata(stage)stage=2      Inf
## trtD-penicillamine:strata(stage)stage=2  NA
## trtPlacebo:strata(stage)stage=3      Inf
## trtD-penicillamine:strata(stage)stage=3  NA
## trtPlacebo:strata(stage)stage=4      Inf
## trtD-penicillamine:strata(stage)stage=4  NA
## trtPlacebo:copper                    1.0051
## trtD-penicillamine:copper            1.0051
## trtPlacebo:albumin                   0.9411
## trtD-penicillamine:albumin            0.9305
## trtPlacebo:log(protime)              348.8593
## trtD-penicillamine:log(protime)      291.5376
## trtPlacebo:ast                       1.0085
## trtD-penicillamine:ast               1.0031
##
## Concordance= 0.773 (se = 0.023 )
## Likelihood ratio test= 120.1 on 14 df,  p=<2e-16
## Wald test                = 70.29 on 14 df,  p=2e-09
## Score (logrank) test = 138.2 on 14 df,  p=<2e-16

```

Based on the above table, we can find that alone, D-penicillamine doesn't have a significant impact on survival time. On the other hand, Bilirubin seems to have a significant impact on survival time. Copper seems to have a significant effect only with the placebo, while on the other hand albumin seems to have a significant effect on both D-penicillamine and the placebo. We can thus infer that bilirubin, copper, and albumin seem to have a significant effect on the patient's survival time. On the other hand, log-transformed prothrombin time and AST are not statistically significant, implying that they don't have much effect on the treatment. Our likelihood ratio test ($p < 2 \times 10^{-16}$) seems to confirm that these interaction terms have a significant impact on the model.

Hazard Ratios and 95% Confidence Intervals

```

##                                     HR lower_95_CI upper_95_CI
## trtD-penicillamine 0.9737276 0.6790927 1.3961944
## log(bili)          2.1192739 1.6743724 2.6823913
## copper              1.0026164 1.0007075 1.0045290
## albumin             0.5133441 0.3254068 0.8098237
## log(protime)       10.9699388 1.0890092 110.5037090
## ast                 1.0003220 0.9968573 1.0037988

```

The presence of penicillamine is not statistically significant, with a hazard ratio of 0.97; penicillamine has a protective effect. However, since the confidence interval is 0.68 to 1.40 with 95% confidence, it is not statistically significant in our model. However, bilirubin is highly significant with a hazard ratio of 2.12 and a confidence interval between 1.67 and 2.68 with 95% confidence. Therefore, bilirubin is strongly associated with an increase in death. Copper, in our model has a hazard ratio of 1, with a confidence interval centered around 1, suggesting a minimal effect. However, this may be misleading because we did not account for the scale of the variable. Since the hazard ratio reflects the effect per one-unit increase, and copper may be measured in small units, meaningful effects from larger changes could be obscured. Albumin significantly

decreases the hazard ratio with a value of 0.51, and a confidence interval between 0.33 and 0.81 with 95% confidence. Prothrombin time has significant adverse effects in our model, with a hazard ratio of 10.97. However, the wide confidence interval between 1.089 and 110.50 suggests high uncertainty, possibly due to outliers in our dataset. Finally, Aspartate Aminotransferase (ast) has a hazard probability of 1, indicating an uncertain effect on patient survival time. However, with a confidence interval between 0.997 and 1.004, Aspartate Aminotransferase is not statistically significant in our model. In which, we should remove Aspartate Aminotransferase from our model.

Conclusion

To answer whether D-penicillamine is effective in slowing the progression of primary biliary cholangitis, and how other confounding clinical factors interact and influence survival time, we created a cox proportional hazard model. Our model shows that D-penicillamine alone does not significantly slow the progression of PBC, as its effect on survival is little to none. Instead, key clinical factors have been found to substantially influence patient outcomes: higher levels of log-transformed bilirubin and prothrombin time are associated with a marked increase in risk. In contrast, elevated albumin levels result in a decrease in risk. A minor yet significant increase in risk is observed with an increase in copper levels, whereas AST does not significantly affect survival times. Therefore, while the main treatment shows little to no effect, other key clinical factors have been shown to affect the survival times for PBC patients.

Advanced Methods

As mentioned in the beginning of our analysis, we note that our dataset provides us with competing risks. To handle the competing risks to answer our research question, we treated both death and liver transplant as an event. Which we treat death and the need to a liver transplant negative events as needing a transplant means that our treatment is not working on our patient. Now, we are interested in finding the effect of our treatment based on our two competing risks death and liver transplant. To accomplish this we will be utilizing a competing risk analysis.

Number of Event Occurrences

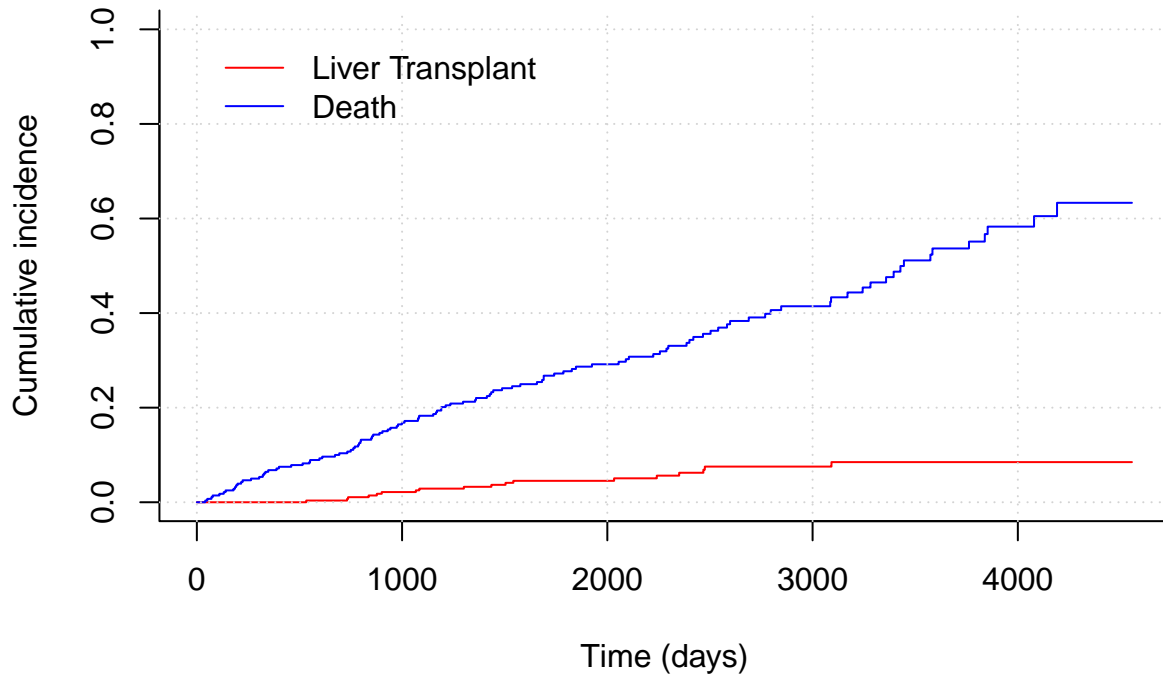
##	Censored	Liver Transplant	Death
##	150	18	112

Based on the results, we have 18 patients who received liver transplants and 112 patients dying. We also notice that we have 150 censored patients. Which indicates that 150 patients are either still alive or have dropped out of the study.

Cumulative Incidence Function

In our analysis, we will be studying the probability of one of our competing risks occurring over time while accounting for the other competing risk. Meaning that we will look at the probability of receiving a liver transplant over time, accounting for the risk of death and vice versa. To accomplish this we will use a cumulative incidence function to observe this trend.

Cumulative Incidence Function Plot

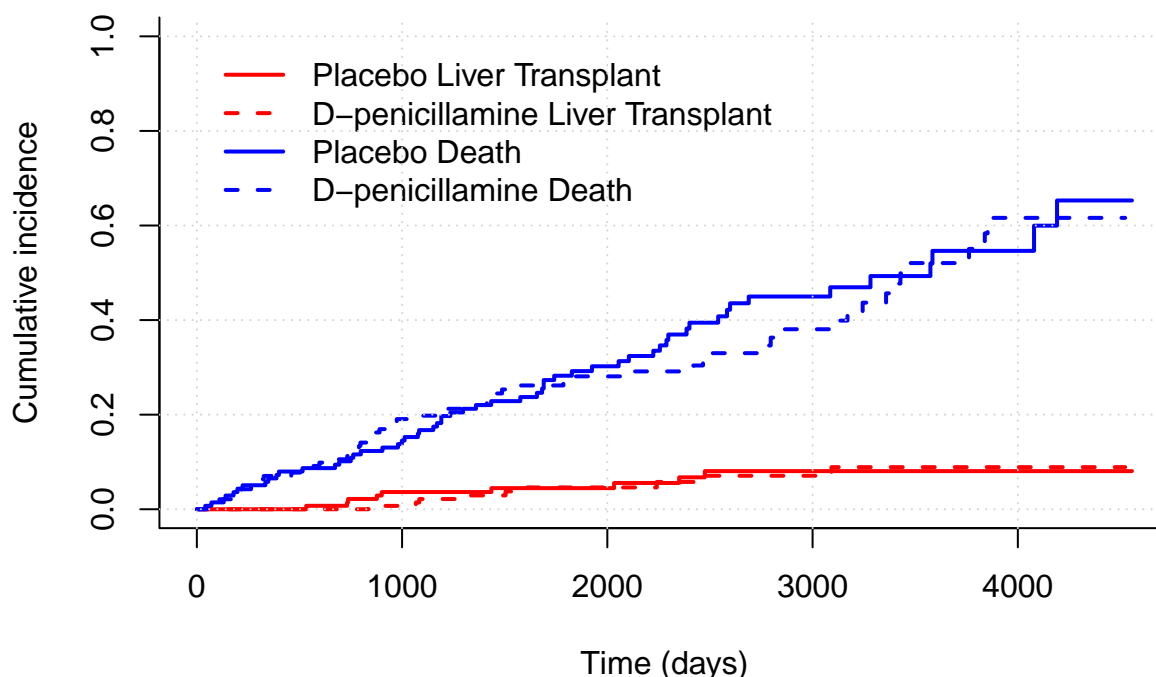


Based on our plot above, we notice that patients in our dataset are more likely to die compared to getting a liver transplant. We can also observe that our cumulative incidence curves of both competing risks follow a linear trend. Which suggests that there is a constant hazard rate of patients experiencing either risk.

Cumulative Incidence Function Based on Treatment

Now, we will be looking into the effects of our treatment based on each competing risk accounting for the other competing risk by using Gray's test for testing our treatment effect.

Cumulative Incidence Function Based on Treatment Plot



With our new CIC(Cumulative Incidence Curve) based on treatment, we notice that our dotted lines follow similar paths are its corresponding solid line. As we see that the red lines that tell us about patients who got liver transplants based on treatments are pretty much a match. Which indicate that D-penicillamine has the same effects as the placebo. Meaning that our treatment is not effective. We can also see a similar outcome with our death risk in the aspect of treatments. As the lines with patients who died with different treatments follow similar trends. This means that D-penicillamine does not do anything in helping with curing or slowing the progression of Primary Biliary Cholangitis. Which also further confirms our conclusions of our research question.

Gray's Test Result

##	stat	pv	df
## 1	0.003230754	0.9546729	1
## 2	0.168344105	0.6815877	1

Based on the results from our Gray's test, we see that for both risks our we get extremely small positive test statistics. Which indicate that our placebo group has a very slightly higher risk rate than our D-penicillamine patient groups for both competing risks. However, it should be noted that this effect is not statistically significant enough for us to conclude the positive effect of our treatment. With our liver transplant group having a p-value of 0.9546729 and our death group having a p-value of 0.6815877.

Fine-Gray Model for Transplant

It is always important to find out how other variables that will affect our patient's health. With that being said, in the case of our competing risk regression, we will utilize the Fine-gray model to first access our transplant patients. In which it will let us compile the effect of covariates in our cumulative incidence function previously mentioned.


```
## Competing Risks Regression
##
## Call:
## crr(ftime = pbc$time, fstatus = pbc$status, cov1 = model.matrix(~trt +
##   age + sex + copper + stage + albumin + log(protime) + ast +
##   log(bili), data = pbc)[, -1], failcode = 1)
##
##               coef exp(coef) se(coef)      z p-value
## trtD-penicillamine -0.232436  0.792601  0.50816 -0.457 6.5e-01
## age                -0.120325  0.886632  0.02335 -5.154 2.6e-07
## sexf               -1.287022  0.276092  0.75549 -1.704 8.8e-02
## copper              -0.000495  0.999505  0.00235 -0.210 8.3e-01
## stage               0.739886  2.095696  0.39964  1.851 6.4e-02
## albumin             0.137432  1.147323  0.80804  0.170 8.6e-01
## log(protime)       -7.518175  0.000543  3.61724 -2.078 3.8e-02
## ast                -0.005999  0.994019  0.00616 -0.974 3.3e-01
## log(bili)           0.647418  1.910601  0.39249  1.650 9.9e-02
##
##               exp(coef) exp(-coef)      2.5% 97.5%
## trtD-penicillamine  0.792601      1.262 2.93e-01 2.146
## age                0.886632      1.128 8.47e-01 0.928
## sexf               0.276092      3.622 6.28e-02 1.214
## copper              0.999505      1.000 9.95e-01 1.004
## stage              2.095696      0.477 9.58e-01 4.587
## albumin            1.147323      0.872 2.35e-01 5.591
## log(protime)       0.000543    1841.204 4.53e-07 0.651
## ast                0.994019      1.006 9.82e-01 1.006
## log(bili)          1.910601      0.523 8.85e-01 4.123
##
## Num. cases = 280
## Pseudo Log-likelihood = -81.3
## Pseudo likelihood ratio test = 30.5 on 9 df,
```

We can find from the above competing risks regression model at a 0.05 significance level that *age* and *log(Protime)* are significant predictors in determining the survival times for those who will undergo transplants. In this case, we can find that younger people and those with higher prothrombin time have a higher chance of getting a transplant. We can also observe that the effects of our treatment is still insignificant (with a p-value of 0.65).

```
## [1] "P-Value: 0.000360527284203569"
```

By the Pseudo Log-likelihood ratio test, we can find that this model has a significant influence on predicting the survival time of patients till they have a transplant as we notice a p-value of 0.00036.

Fine-Gray Model for Death

Now, we will be looking into a Fine-Gray model for our subjects who have passed away and see what relevant covariates effect the probability of our patients dying over time, while accounting for the risk of a transplant.

```
## Competing Risks Regression
##
## Call:
## crr(ftime = pbc$time, fstatus = pbc$status, cov1 = model.matrix(~trt +
##   age + sex + copper + stage + albumin + log(protime) + ast +
##   log(bili), data = pbc)[, -1], failcode = 2)
```

```
##
##               coef exp(coef) se(coef)      z p-value
## trtD-penicillamine 0.01993    1.020  0.20082  0.0992 9.2e-01
## age                0.03624    1.037  0.01192  3.0407 2.4e-03
## sexf               0.00247    1.002  0.31675  0.0078 9.9e-01
## copper              0.00167    1.002  0.00129  1.2918 2.0e-01
## stage              0.28031    1.324  0.12392  2.2621 2.4e-02
## albumin            -0.75097    0.472  0.23355 -3.2154 1.3e-03
## log(protime)       4.18280   65.549  1.02503  4.0807 4.5e-05
## ast                0.00375    1.004  0.00183  2.0455 4.1e-02
## log(bili)          0.67197    1.958  0.11559  5.8132 6.1e-09
##
##               exp(coef) exp(-coef)  2.5%  97.5%
## trtD-penicillamine    1.020    0.9803 0.688  1.512
## age                   1.037    0.9644 1.013  1.061
## sexf                   1.002    0.9975 0.539  1.865
## copper                  1.002    0.9983 0.999  1.004
## stage                  1.324    0.7555 1.038  1.687
## albumin                 0.472    2.1191 0.299  0.746
## log(protime)          65.549    0.0153 8.791 488.737
## ast                    1.004    0.9963 1.000  1.007
## log(bili)              1.958    0.5107 1.561  2.456
##
## Num. cases = 280
## Pseudo Log-likelihood = -480
## Pseudo likelihood ratio test = 169 on 9 df,
```

We can find that the above competing risks regression model at a 0.05 significance level, that *age*, *stage*, *albumin*, *log(protime)*, *ast*, and *lob(bili)* are all significant predictors in determining the survival probability for those who will die. We can find that those who are older, have a higher disease stage, lower albumin, higher *log(protime)*, higher *ast*, and higher *log(bilirubin)* all increase risk and chance for death. In addition, we can find that *log(protime)* has a much more significant impact on the survival time than other covariates, significantly increasing the risk. What is interesting is that we notice that there are additional covariate effects in our death model are more statistically significant than our transplant model. This signals that our covariates have stronger associations with the progression of PBC, which lead to death. We must account for the fact that we have significantly more cases of death than liver transplants.

```
## [1] "P-Value: 9.96819034784342e-32"
```

By the pseudo log-likelihood ratio test, we can find that the model has a significant influence on predicting the survival time for those who will die with a p-value of 9.97×10^{-32} .

Overall Thoughts

From our competing risks analysis, which separately examines liver transplant and death, we can find a similar result from our earlier Cox model, further reinforcing our findings. We once again found that D-penicillamine does not significantly slow the progression of PBC. The cumulative incidence curves for both events show similar trajectories between the D-penicillamine and placebo groups, and we can also find a higher baseline risk for death due to a faster rising cumulative incidence curve than with transplants. Gray's tests ($p = 0.9547$ for liver transplant and $p = 0.6816$ for death) confirm that there are no significant differences in risk between the treatments. This indicates that, even when accounting for the competing nature of death and transplant as negative outcomes, D-penicillamine does not give a survival benefit. Instead, other covariates such as log-transformed bilirubin and prothrombin time, which increase the hazard rate, and higher albumin levels, which decrease the hazard rate, still result as the key influencers in predicting survival in PBC patients. With all that is said, with both analyses that we applied have yield similar results. In that our D-penicillamine treatment does not have significant effects on the progression of patients with PBC that would lead to adverse events (i.e. death or getting liver transplant).

Even with controlling for our confounding variables, our treatment model remains statistically insignificant. Interestingly we notice that there are some covariates that will impact death and not transplant. It should be noted that we have significantly more cases of death than liver transplants. Females are more susceptible to PBC than males, and this was a defining factor in our analysis of PBC. Despite the inclusion of sex and other clinical predictors of PBC, the effect of treatment does not meet statistically relevant conditions to be an effective treatment option. Furthermore, the effect of these variables may change with time, violating our Cox PH Assumption. The lack of statistical relevance in the final Cox PH Model and the Competing Risk Regression proves that the effect of treatment is inherently insignificant. The fact that similar results were yielded across multiple tests in our analysis is evidence of this. Thus, D-penicillamine and other confounding clinical factors do not significantly impact survival time in PBC patients, regardless of sex.

Reference

Gray, Bob. 2024. cmprsk: Subdistribution Analysis of Competing Risks. Version 2.2-12, May 19. CRAN. <https://www.R-project.org>

Kleinbaum, D. G., & Klein, M. (2012). Survival Analysis: A Self-Learning Text, Third Edition (3rd ed. 2012.). Springer New York. 430-455, 468-473 <https://doi.org/10.1007/978-1-4419-6646-9>

Mayo Clinic Staff. “Primary Biliary Cholangitis (PBC) – Symptoms and Causes.” Mayo Clinic, April 4, 2023. <https://www.mayoclinic.org/diseases-conditions/primary-biliary-cholangitis/symptoms-causes/syc-20376874>

R Core Team. “PBC Data Set.” CRAN - R Project for Statistical Computing. Accessed May 6, 2025. <https://search.r-project.org/CRAN/refmans/compound.Cox/html/PBC.html>

Therneau, Terry M., and Patricia M. Grambsch. 2010. Modeling Survival Data: Extending the Cox Model. Edited by Terry M. Therneau and Patricia M. Grambsch. N.p.: Springer. 10.1007.

Trivella, J., John, B. V., & Levy, C. (2023). Primary biliary cholangitis: Epidemiology, prognosis, and treatment. Hepatology communications, 7(6), e0179. <https://doi.org/10.1097/HC9.0000000000000179>